Journal of Safety Research 78 (2021) 47–58

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A methodology for assessment of long-term exposure to whole-body vibrations in vehicle drivers to propose preventive safety measures

María L. de la Hoz-Torres [a,*], Antonio J. Aguilar [a], Mª Dolores Martínez-Aires [b], Diego P. Ruiz [a]

[a] Department of Applied Physics, University of Granada, Av. Severo Ochoa s/n, Granada, Spain
[b] Department of Building Construction, University of Granada, Av. Severo Ochoa s/n, Granada, Spain

ARTICLE INFO

Article history:
Received 26 May 2020
Received in revised form 16 January 2021
Accepted 21 April 2021
Available online 4 June 2021

Keywords:
Whole-body vibrations
ISO2631-5
Vibration assessment
Safety interventions
Heavy equipment vehicles

ABSTRACT

Introduction: The appearance of musculoskeletal disorders (MDs) in professional drivers due to exposition to whole-body vibration (WBV) makes it relevant to assess this exposure. The European Directive 2002/44/EC has two methods to evaluate exposure to WBV (defined in ISO2631-1:2008). These methods evaluate the exposure associated with an 8-hour working day; however, MDs due to WBV could also be caused by accumulated exposure to vibrations over long term, and hence, the methods defined in the European directive may be limited in their ability to ensure the safety of workers exposed to WBV throughout their years of employment. *Method:* A detailed comparison and discussion of methods defined in the European Directive and the ISO2631-5:2018 was used as a starting point of the main results of this paper. On this basis, a new methodology for the management and organization of preventive measures is proposed to consider the assessment of ISO2631-5:2018 standard and the full working life of workers. Experimental data to assess exposure to WBV in heavy equipment vehicle (HEV) drivers under different road surface conditions and range of velocities were considered to illustrate the process of the proposed methodology. *Results:* The methods defined in the standards provide different assessments leading to a different possible consideration of safe operations when the risks associated with them may actually be high. The proposed methodology can be used with the aim of ensuring safety of workers throughout their working lives and providing an easy implementation of the calculations of ISO2631-5:2018 standard. *Conclusions:* A procedure to assess the health risk probability to which the HEV worker is exposed in terms of the exposure years and a different range of operational vehicle speeds is proposed and exemplified with a study case. *Practical applications:* This study provides a practical tool for the management of WBV exposure related to work-tasks in HEV drivers. Safety managers should consider the global exposition to WBV throughout their working life, and this research provides an easy tool to accomplish it.

1. Introduction

European workers reported that musculoskeletal disorders (MDs) are one of the main causes of work-related ill-health (Nielsen, Jørgensen, & Malgorzata Milczarek, 2018), and construction, agriculture, and transportation are among the industries with higher rates of MDs. Hence, it is important to have health and safety requirements and strategies to limit, assess, and control specific risks associated with the appearance of MDs (Spielholz et al., 2008; Yazdani et al., 2018). There is epidemiological evidence that relates whole-body vibration (WBV) to MDs, such as low back pain (Bovenzi & Betta, 1994; Punnett & Wegman, 2004; Raffler et al., 2017), degenerative changes in the lumbar spine (Miyamoto, Shirai, Nakayama, Gembun, & Kaneda, 2000; Wilder et al., 1996), sciatica (Burström, Nilsson, & Wahlström, 2015), neck pain (Kim, Dennerlein, & Johnson, 2018; Milosavljevic, Bagheri, Vasiljev, Mcbride, & Rehn, 2011; Rehn, Nilsson, Lundström, Hagberg, & Burström, 2009), and disorders such as motor performance (Costa, Arezes, & Melo, 2014). In this sense, and within MDs, spine disorder is the most frequently reported group of diseases among workers in the construction industry (Bakusic et al., 2018; Health and Safety Executive, 2018). The most important single risk factor associated with low back pain is the amount of lumbar disc degeneration (Livshits et al., 2011), with overweight and obesity increasing the risk of appearance (Liuke et al., 2005; Shiri, Karppinen, Leino-Arjas, Solovieva, & Viikari-Juntura, 2009). Among individuals affected by low back pain, between 5.0% and 10.0% will develop a chronic pain problem, the prevalence of which increases linearly from 30 until 60 years (Meucci, Fassa, & Faria,

* Corresponding author.
E-mail address: mlhoz@ugr.es (M.L. de la Hoz-Torres).

https://doi.org/10.1016/j.jsr.2021.04.002

2015). These occupational diseases have a great impact on individuals and social care systems as well as high treatment costs and sick absence (Woolf & Pfleger, 2003).

Professional drivers of heavy equipment vehicles (HEVs) are often exposed to WBV and mechanical shocks (Kittusamy & Buchholz, 2004; Johnson, Dennerlein, Ramirez, Arias, & Rodríguez, 2015; de la Hoz-Torres, López-Alonso, Ruiz & Martínez-Aires, 2017). In fact, their work tasks can lead to WBV exposures among HEV operators (Blood, Rynell, & Johnson, 2012). A large number of the activities performed with HEVs are carried out on uneven surfaces, which are more likely to produce high levels of WBV and mechanical shocks as compared to activities performed on even surfaces (Kumar et al., 1999; Griffin et al., 2008; Milosavljevic, Bergman, Rehn, & Carman, 2010). Previous research has shown that the type of operation and ride conditions significantly affect the compressive stress on lumbar spine response (Singh et al., 2019). In addition, the type of seat suspension (active or passive) and seat suspension maintenance (Rahimdel & Mirzaei, 2020) significantly reduce WBV exposures. Because long periods of exposure to WBV can lead to health problems for drivers (Milosavljevic et al., 2010; Smets, Eger, & Grenier, 2010; Kia et al., 2020), in this research, we focus specifically on this sector, with the aim of proposing a methodology that provides information on the risk of adverse health effects to the vertebral endplates of the lumbar spine for seated individuals due to compression. Our procedure is based on the long-term exposition of workers to WBV and the analysis of the current standards. The proposed methodology, combined with medical, imaging, and biomechanical evaluation and health surveillance, has the potential to be a key tool to prevent possible negative effects on health.

On the basis of the above initial hypothesis, currently, the most accepted and used method for assessing WBV exposure is that defined in ISO 2631-1:2008 (2008). The methods defined in this standard are used to evaluate exposure in an 8-hour working day. However, the basic assessment method defined in this standard is only suitable for describing the severity of vibrations in relation to their effects on human beings for exposure with peak factors of the measured signal less than or equal to 9. In other cases, when the basic assessment method may underestimate the effect of vibrations, the standard refers to the use of a method based on the concept of vibration dose value (VDV).

Recently, ISO 2631-5:2018 (2018) was published based on the revision of the previous standard ISO 2631-5:2004 (2004). The new ISO2631-5:2018 standard defines two different methods in terms of the exposure regime (severe exposure regime and less severe exposure regime), and these methods assess the risk of chronic injury from exposure to repeated shock based on the predicted biomechanical response of the bony vertebral endplate (hard tissue). Now, the methods for calculating the acceleration transmitted to the spinal column is through a transfer function of a biomechanical model. Unlike the previous standard, neither method is limited by the signal crest factor. However, the limits in this new revised standard ISO2631-5:2018 remain unchanged compared to ISO2631-5:2004, although Eger et al. (2008) reported that the limits established in the standard ISO2631-5:2004 may be set possibly too high. Previous research has also concluded that evaluation of the relationship between ISO 2631-1 and ISO 2631-5 parameters deserves further investigation (Blood et al., 2012; De la Hoz-Torres, Aguilar-Aguilera, Martínez-Aires & Ruiz, 2019).

In summary, there are two relevant standards, namely ISO2631-1:2008 and ISO2631-5:2018, for the problem addressed in this research. Both are important for evaluating the operation performances with HEVs, because these activities may expose drivers to WBV with a high amount of mechanical shocks. Although there are clear differences between the two standards, given the short lapse of time since the publication of ISO 2631-5:2018, very little research has been published related to it and there is a lack of information on how and when to use them and their feasibility.

This research was performed to make a more comprehensive comparison of the evaluation methods described in ISO2631-1:2008 and ISO 2631-5:2018 in the context of HEV drivers, given their relevance. Based on this comparison, we propose a new methodology for the management of WBV exposure in vehicle drivers (or other workers also exposed to WBV), with the aim of improving their quality of life both throughout their working lives and in their retirement. The methodology is then implemented in an illustrative case study to show and illustrate how the application of the proposed steps can be done in an experimental setup and in a real case. The results for this illustrative case should not be automatically extended to other cases, as each driving activity must be analyzed on a case-by-case basis with the application of the proposed scheme. To achieve this objective, a collection of real data from a set-up field experiment (case study) was performed considering a typical variability in speed and surface conditions for HEV drivers. From this experiment, a total of 94 measured data sets were analyzed and then evaluated according to the standards. On the one hand, this data analysis allowed us to draw conclusions on the effect on the WBV magnitude calculation and the possible exceeding of the standard limits of both the different surface categories (e.g., on tarmac road) and vehicle speed. On the other hand, the data sets were also used to investigate the evolution of the risk factor over time derived from cumulative exposure to WBV. From the analysis of this behavior, the proposed methodology allows us to manage WBV exposure for HEV drivers to keep them safe in a quick and easy way. In this context, it is worth noting that processes shown in this paper could be an essential tool to support many at-risk workers for those safety and health professionals who do not have a deep knowledge of WBV (Paschold & Sergeev, 2009).

The article is structured as follows:

- Section 2: preliminary concepts, definitions, and data processing techniques used in this research are featured. Thus, in this section, the assessment parameters established by the ISO 2631-1 and ISO 2631-5:2018 standards are also reviewed, as well as the standardized limits to an 8-hour exposure reference period and the boundaries for the emergence of probable health effects derived from multiple shock vibration exposure coming from the ISO 2631-5:2018. In addition, the data processing clustering method used as part of the proposed methodology is outlined in this section.
- Section 3: proposal of a methodology for health risk prediction assessment, wherein a method is proposed to quickly assess the health risk probability to which the HEV worker is exposed in terms of the exposure years and a different range of speeds, considering the entire working life of the driver.
- Section 4: implementation of the proposed methodology to a HEV driver case, the overall process of the implementation of the proposed methodology is illustrated on the basis of a real case, from data collection to health risk prediction assessment.
- Section 5: conclusions, wherein the main findings, conclusions, and practical applications of this research are drawn.

2. Preliminary concepts, definitions, and data processing techniques

2.1. Whole-body vibration assessment

As noted in the introduction, ISO 2631-1 and ISO 2631-5 standards define methods of risk quantification. These standards use the recorded and measured acceleration on the seat surface to calculate the WBV exposure parameters. The procedure used in both ISO standards is summarized below:

2.1.1. ISO 2631-1:2008 parameters

This standard is based on the calculation of the root mean square (rms_w) of the weighted averaged acceleration (m/s^2) and the vibration dose value (vdv_w).

For calculating the first parameter, Butterworth filters are used to weigh the acceleration in frequency according to the ISO 2631-1 standard. The *x*- and *y*-axes are weighted using weights denoted as W_d, and for the *z*-axis using W_k. The root mean square (rms_w) of the weighted averaged acceleration (m/s^2) is then calculated as the second power of the acceleration time history as the basis for the averaging process (Eq. (1)):

$$rms_w = \left[\frac{1}{T}\int_0^T a_w^2(t)dt\right]^{\frac{1}{2}} \tag{1}$$

where a_w is the frequency-weighted instantaneous acceleration (W_d on *x* and *y* axes, W_k on *z* axis), and *T* is the time duration of the measurement.

The vibration dose value (vdv_w) is calculated as the fourth power of the acceleration time history (Eq. (2)); hence, this parameter is more sensitive to peaks than the rms_w:

$$vdv_w = \left[\int_0^T a_w^4(t)dt\right]^{\frac{1}{4}} \tag{2}$$

To allow comparisons between different exposures, these parameters are normalized to reflect 4 h of exposure to WBV for an 8-h work cycle. The daily exposure value ($A(8)$) (Eq. (3)) and the vibration dose value method (VDV) for each axis (Eq. (4)) are then calculated as follows:

$$A(8) = krms_w\sqrt{\frac{T_{exp}}{T_0}} \tag{3}$$

$$VDV = kvdv_w\sqrt{\frac{T_{exp}}{T_{meas}}} \tag{4}$$

where *k* denotes the multiplication factor defined for each axis ($k_{x,y} = 1.4$ and $k_z = 1$), T_{exp} is the measurement period, T_0 is the reference duration of 8 h, and T_{meas} is the daily duration of exposure to the vibrations. The calculated values can then be compared with the daily exposure action value ($A(8) = 0.50$ m/s^2 and $VDV = 9.10$ m/s$^{1.75}$) and the daily exposure limit ($A(8) = 1.15$ m/s^2 and $VDV = 21.00$ m/s$^{1.75}$) established by the EU directive (Directive, 2002/44/EC).

2.1.2. ISO 2631-5:2018 parameters

Unlike the basic evaluation method described in ISO 2631-1, this standard defines two assessment methods based on different exposure regime conditions. This research implements the method that addresses what the standard calls "less severe conditions" as the exposures do not contain free-fall events.

As established by the standard, the analysis should be accomplished assuming the most unfavorable exposure conditions, considering the exposure time periods (hours per day and days per year) and the life-time exposure history. Also, the Posture Group and the anthropometric characteristics of the drivers are used as input of the model. The intervertebral compressive forces and the daily compressive dose for the six disc levels of the lumbar spine (T12/L1, L1/L2, L2/L3, L3/L4, L4/L5, and L5/S1) were calculated as follows (Eq. (5)).

$$S^A = \left(\sum_i\left(\frac{c_{dyn,i}}{B}\right)^6\right)^{\frac{1}{6}} \tag{5}$$

where $c_{dyn,i}(N)$ is the sum of peak compressive forces acting on the vertebral endplate and B (mm^2) is the area of the vertebral endplate. The equivalent daily compressive dose is calculated considering the total duration of the exposure during a day (Eq. (6)).

$$S_d^A = \left(\sum_j S_j^{A6}\frac{t_{dj}}{t_{mj}}\right)^{\frac{1}{6}} \tag{6}$$

where S_j^A is the dynamic compressive stress of the lumbar spine due to vibration exposure, t_{dj} is the time period of the daily vibration exposure and t_{mj} is the time period over which S_j^A has been measured. The Risk Factor R^A is estimated at each vertebral level based on the S_d^A:

$$R^A = \left(\sum_{m=1}^{n}\left(\frac{S_d^A N_m^{\frac{1}{6}}}{S_{ui}^A - S_{stat,i}^A}\right)^6\right)^{\frac{1}{6}} \tag{7}$$

$$S_{stat,i}^A = 6.765MPa - 0.067MPA(b+i) \tag{8}$$

where *N* is the number of exposure days per years, *n* is the number of years of exposure, S_{ui}^A is the ultimate strength of the lumbar spine for a person of age *(b + i)* years and $S_{stat,i}^A$ is the mean value of the compressive-decompressive force divided by the area of vertebra endplate.

2.1.3. Health guidance caution zone (HGCZ)

The European Directive 2002/44/EC specifies that the methods for assessing WBV exposure are those defined in ISO 2631-1, and it determines standardized limits to an 8-h exposure reference period. In addition, ISO 2631-5:2018 defines boundaries for the emergence of probable health effects derived from multiple shocks vibration exposure (Table 1).

2.2. Data processing clustering method: unsupervised clustering

To obtain a grouping of the daily compressive dose to test the appearance of differences between different groups based on the mean velocity, a clustering process is performed as part of the proposed methodology. In our study, we used the k-means++ algorithm, a variant of the original k-means algorithm. It is an unsupervised classification algorithm (Arthur & Vassilvitskii, 2007) in which the grouping is done by minimizing the sum of distances between each object to the centroid of its group or cluster. Given an initial number of data, the algorithm follows the following steps: (1) select an initial center using a uniform random variable. This first centroid is called *c*1; (2) calculate the distances for each point *x* to the centroid *cj*. The distance between the observation m and the centroid *cj* is denoted as D(*Xm*,*C*1)/D(*x*); (3) Select the next centroid (uniform random variable) *cj*, with the probability (Eq. (9)) [using a weighted probability distribution where a point *x* is chosen with the probability proportional to $D(x)^2$]:

$$\frac{d^2(x_m, c_1)}{\sum_{j=1}^n d^2(x_j, c_1)} \tag{9}$$

(4) Repeat step 2 and 3 until the centroids k are chosen.

In the next steps, the algorithm proceeds as in the original *k-means* algorithm, i.e., (5) for each $i \in (1, \dots k)$, set the C_i cluster as the set of points in *X* that are closer to ci than to *cj* for all *j* other than *i*; (6) for each $i \in (1, \dots k)$, sets c as the center of mass of all points in C_i: $c_i = 1|C_i|\sum x \in Cix$; and finally (7) repeat steps 5–6 until there are no changes in the cluster assignment or until the maximum number of iterations is reached.

To select the optimal number of clusters, in this work, the Elbow and GAP methods (Tibshirani, Walther, & Hastie, 2001) were used. The GAP method is based on comparing the intra-grouping

Table 1
Health guidance caution zone.

Directive 2002/44/EC Exposure limit values and action values			ISO 2631-5:2018 Probability of an adverse health effect		
Exposure Action Value (EAV)	$A(8) = 0.50\frac{m}{s^2}$	$VDV = 9.1\frac{m}{s^4}$	Low	$S_d^A < 0.5\ MPa$	$R^A < 0.8$
Exposure Limits Value (ELV)	$A(8) = 1.15\frac{m}{s^2}$	$VDV = 21.0\frac{m}{s^4}$	Moderate	$S_d^A > 0.5\ MPa\ S_d^A < 0.8\ MPa$	$R^A > 0.8\ R^A < 1.2$
			High	$S_d^A > 0.8\ MPa$	$R^A > 1.2$

dispersion with the expected one under a uniform distribution of points that plays the role of the null hypothesis. The number of groups that maximizes that difference is the optimal number of clusters.

The elbow method is based on minimizing the intra cluster variation (within-cluster sum of squares). Thus, it compares the within-cluster sum of squares with its expectation under a reference null distribution. In this method, the value of the so-called Elbow index decreases as the number of clusters increases. In this case, the "elbow" point on the graph becomes the optimal number of clusters, as the slope value on the graph is no longer important.

3. Methodology for health risk prediction assessment of human exposure to long-term whole-body vibration

The proposed methodology is based on ISO 2631-5:2018 standard, and it assesses the impact of long-term exposure to WBV, unlike the methods used for the assessment of WBV in the Directive 2002/44/EC for the evaluation of vibration exposure (based on A(8) and VDV parameters), which only assesses the exposure associated with an 8-hour working day.

The methodology is articulated in six steps, and these steps are summarized in Fig. 1. This procedure is a generalization of the process followed in this research, from data acquisition to risk factor evolution modeling. The key point is the generation of a color map of the Risk Factor evolution for each HEV and activity. Note that the strength of the proposed methodology is to ensure that vibration exposure is managed over the years so that the probability of occurrence of an adverse health effect remains low. Despite the initial effort required to implement the proposed methodology, the results obtained enable to ensure the safety of the worker throughout their working life.

Step 1. WBV exposure and drivers characteristics. In the initial step, data are collected from three different categories. Category I: the characteristics of the activity (surface, velocity, and performance characteristics); Category II: the characteristics of WBV exposure (if the exposure contains multiple shock and duration); Category III: driver characteristics (age, height, weight, and posture). The information collected during the initial phase of the process will be used as input of the methodology. The analysis of these data is used to define a measurement strategy to ensure that the WBV exposure is measured in representative situations. In addition to the measurement strategic requirements stated in ISO2631-1 and ISO2631-5, additional standards could be considered in the design of the experimental set up.

Step 2. Measurement and data processing. From the data obtained in the previous step, the measurement strategy has to be established. The speed of the vehicle and the acceleration at the interface between the seat and the driver must be measured. The number and duration of the measurement should be sufficient to ensure that the measured results are representative of the exposure.

Based on the data obtained in the measurement process, the parameters defined in ISO 2631-5 (S_d^A) and ISO 2631-1 (A8 and VDV) must be estimated. The daily compressive dose (S_d^A) (most unfavorable vertebra) and the average velocity must be used as input variables in a cluster data process. The objective is to obtain a grouping of the daily compressive dose to test the appearance of differences between different groups based on the mean velocity. The k-means++ algorithm is used in this process.

Step 3. Assessment of daily exposure. The parameter A8, VDV, and S_d^A estimated in step 2 must be used to assess the daily exposure. For this purpose, these values have to be compared with the HGCZ values (Table 1). If the exposure exceeds the ELV or EAV or the probability of an adverse health effect is high, measures must be taken to limit exposure of drivers to ensure their safety.

Step 4. Assessment of cumulative WBV exposure over the years. For the data set of each cluster (estimated in step 2), the average value S_d^A must be calculated for the vertebral levels. From the value obtained for the most unfavorable level of vertebra, the probability of occurrence of an adverse health effect (R^A) is estimated. The evolution of the R^A parameter over the years is calculated considering the exposure lasts from the age of 20–70 years and the foreseen days of exposure per year (e.g., 240 days per years if a full working life with a typical length of the working day in this sector is considered). The values obtained should be compared with the HGCZ values (Table 1) to assess the exposure over the years. To model the cumulative effect of WBV with the parameter R^A, the data obtained for each cluster must be fitted using polynomial fitting function. The parameterization of data allows the health risk probability to which the worker is exposed in any year to be determined, from a data set measured under specific velocity and surface conditions.

Step 5. Color map of risk factor evolution. Specifically the proposed method relies on the model the cumulative effect of WBV with the parameter R^A. Based on the data obtained, a bidimensional surface-type fitting must be carried out using a polynomial surface model, where the *x*-axis is the age of the worker and the *y*-axis represents the average speed at which the activity is performed.

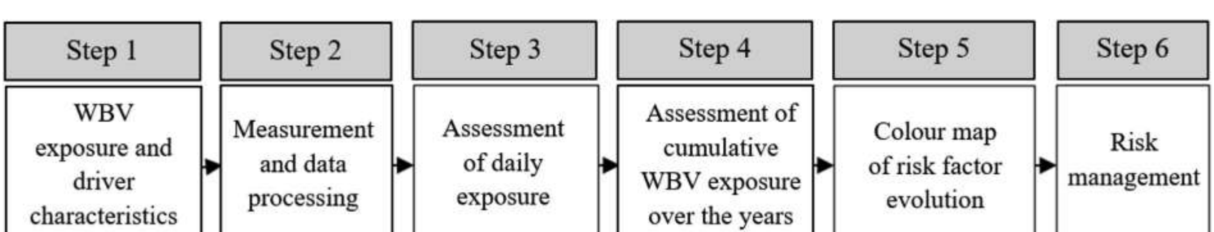

Fig. 1. Diagram of the proposed methodology.

Step 6. Risk management. The color map obtained in Step 5 can be used to manage WBV exposure. The safety manager can use it to assess the long-term cumulative WBV exposure. Based on the characteristics and speed at which the activity is performed, the safety manager can assess if the worker will reach a high probability of an adverse health effect (limit defined in ISO2631-5:2018, Table 1) throughout his working life. In addition, the color map can be used by drivers as an information and training tool. It provides information that can be used to raise awareness of the importance of the performance characteristics of the activity and the impact of long-term WBV exposure.

This methodology can be applied to any activities carried out under less severe WBV exposure conditions such as those listed in Article 4(b) of ISO 2631-5:2018. HEVs are used in these types of activities, such as driving with tractors, forestry machines, and mobile earth-moving machinery over rough surfaces (off-road, potholes, frequent crossing of railroad tracks, etc.). Given that the ride condition (i.e., surface condition and forward speed), type of operation, and the use of machinery significantly affect the S_d^A value, the proposed methodology should be applied on a specific case-by-case basis. In addition, as the type of crop may influence the operation performance and each country has set specific limits to WBV exposure, all these aspects have to be also considered when applying this methodology and the resulting risk management. The following section shows how to apply the proposed methodology to a given case, from the data collection step to the risk factor evolution and management.

4. Implementation of the proposed methodology to a HEV driver case

To illustrate the application of the proposed methodology, in this section, the preceding proposed process steps are applied to a real study case to generate the color map for risk assessment as the final goal, according to the procedure proposed in Section 3.

4.1. Whole-body vibration and driver's characteristics. Step 1

4.1.1. Experimental set-up

An experimental data measuring schedule was designed to assess the exposure to WBV in HEV drivers as a case study to test the implementation of the ISO standards and the proposed methodology. The magnitude to be measured was the acceleration at the interface between the seat pad and the ischial tuberosities. The experimental design included a monitoring of the exposure to WBV in a standardized test route comprising a variety of representative real surface conditions for HEV displacement. Previous studies analyzing the transmission of vibrations through the seat in agricultural tractors have considered different types of surfaces, such as tarmac and rough track (Adam & Jalil, 2017; Giordano, Facchinetti, & Pessina, 2015). In this case, the path and length of the routes were established to include the possibility of obtaining vibration data sets corresponding to different speeds representative of typical surface conditions found in HEV drivers (i.e.. the route included 1 km of off-road, 4 km of unpaved road, and 5 km of tarmac road). The test locations were specifically chosen for two main reasons: (1) low traffic disturbance (to achieve a stable environment during the test and to minimize interference due to external interruptions) and (2) diversity and representativeness of the sample of road surfaces.

The same route was also used to evaluate the vibrations transmitted by the vehicle to the driver at a wide range of velocities. In this case, speed was monitored by a Global Positioning System (GPS) attached to the vehicle. The lowest speed value (5 km/h) was chosen for reproducing realistic travel conditions with the usual lower speed. The highest speed value (25 km/h) was chosen because it is the maximum speed limit (HEV speed regulation). The vehicle used throughout the entire study was a tractor classified as Class II Category A according to the Directive 78/764/EEC (1978). As there are large variations in the magnitude of vibration depending on the type of vehicle (Paddan & Griffin, 2002), all measurements were performed with the same tractor to eliminate this variation, as the aim of this study case was to test the use of the standards to propose a methodology for exposure assessment.

The duration of the measurements was selected to provide adequate data to be representative of the exposure in different conditions (surface and displacement velocity). The displacement through unpaved roads and tarmac roads are recurring tasks; therefore, performing several subsequent measurements observing a minimum measurement time at an average speed (with a maximum speed dispersion of 5 km/h) is enough to ensure that the result is representative of driver exposure. However, off-road travelling is a nonrepetitive task; hence, the terrain was studied and characterized in a first approach, and successive measurements of sufficient duration were performed.

With regard to the driver, a healthy male adult was chosen to participate in this field study. The reasons for this option are mainly as follows: the participant had more than 20 years of driving experience (HEV including trucks and agricultural tractors) without current pain and history of MDs; he was 48 years old and his height and weight were 1.85 m and 120 kg, respectively, with a body mass index of 35.06 kg/m^2. This high body mass index (BMI > 35 kg/m^2) indicates that this person suffers from obesity, and because body weight is related to spinal loading, this factor increases the risk of low back pain and lumbar disc degeneration (Liuke et al., 2005). Therefore, the subject belongs to a high-risk group for the development of MDs. Because the driver belongs to a high-risk group, a comparison of assessment methods in humans who may be at higher risk is of particular interest. It is worth noting that a high BMI within the highest body mass percentile range (BMI > 26.1 kg/m^2 and 95th percentile, i.e., a body mass larger than 109 kg) are those values defined in ISO 2631-5:2018 Annex A.3 that maximize the spinal load; hence, it is an interesting study case for its specific characteristics. Regarding posture groups classified in Annex A of ISO 2631-5, the driver's posture was the posture group number 3. Posture group 3 and the anthropometric characteristics of the driver were used as input of the model. As with the vehicle selection, all the measurements were performed by the same subject to eliminate the uncertainty associated with variables linked to the anthropometric characteristics of the operator.

4.1.2. Measurement equipment

A tri-axial accelerometer (SV38, SVANTEK) was used to measure the acceleration transmitted to the seat pad. The instrument enables the sampling of the experimental acceleration with a frequency of 6000 Hz in each direction: fore-to-aft (x axis), left-to-right (y axis) and buttocks-to-head (z axis). Raw unweighted acceleration signal was recorded and stored in a data logger (SV106, SVANTEK) connected to the accelerometer. According to the ISO2631-5:2018 standard, the sign of the acceleration signal was also recorded. The equipment meets the ISO 8041, ISO 10323-1, and ISO 2631 requirements for measurements. The time and position of the vehicle were also simultaneously recorded via a GPS logger.

4.2. Measurement and data processing. Step 2

The procedure adopted for the field testing consisted of three steps. First, the sensors were installed: the accelerometer was placed on the seat surface and its position was adjusted to ensure the correct positioning of the axes. It was fixed with an adhesive tape to avoid relative displacement between the seat surface and the sensor. In addition, the GPS was placed on the surface of the vehicle dashboard.

Second, the measurement with both sensors started simultaneously and the test started. During the measurements, the subject remained seated and did not lose contact with the seat surface (the subject was instructed and supervised not to get up from the seat just to ensure that the exposure did not include bad acceleration data measured during loss of contact). Moreover, the driver was monitored performing the activity under normal working conditions. If there was a significant anomaly in the recorded test data, the experiment was carried out again. The test was performed several times for each test section (at least three times) to reduce random errors. From the data obtained in each measurement, those with a length of more than 90 s and a maximum speed deviation of ±2.5 km/h were selected. In this data selection, we followed the recommendations of ISO 2631-1, which states that a minimum measurement duration of 108 s for a lower frequency limit of 1 Hz is required to assure an error less than 3 dB at a 90% confidence level. This data preprocessing step was intended to eliminate the vibration measurements in nonstable velocity periods and to ensure the minimum measurement duration to provide representative results of the exposure in tested conditions.

Finally, based on the acceleration data and other recorded data from the experiment, the daily compressive dose S_d^A is calculated according to Section 2. The acceleration measured at seat surface was used for the seat and backrest in the model. To compare different exposures, the same set of conditions is used to normalize the measured exposure to a typical/realistic daily exposure. In our test, the duration exposure conditions were chosen in such a way that they maximize the spinal load, as the values defined in Annex A.3 of ISO2631-5: the daily exposure duration were normalized to 4 h to compare the results obtained in the different exposure conditions. An exposure of 240 days per year for the ages from 20 to 70 years was considered as the lifetime exposure history.

The daily compressive dose (most unfavorable vertebra) and the average velocity were used as input variables (Fig. 2), and the objective was to obtain a grouping of the daily compressive dose to test the appearance of differences between different groups based on the mean velocity.

Therefore, a clustering process was carried out using the k-means++ algorithm in terms of the average velocity for different types or roads. As described in Section 2, prior to the application of the algorithm, it is necessary to establish the number of clusters to carry out this process. For the selection of the optimal number of clusters, the Elbow and the GAP methods were used. Fig. 3 shows the results of the analysis using both methods. In this study, the optimum number of clusters obtained is $k = 4$, for both methods, according to the selection criteria given in Section 2.

Once the optimum number of clusters was set up, the clustering process was subsequently applied. Fig. 4 shows the obtained results. The data set assigned to cluster 1 was recorded on off-road; cluster 2 contains data recorded on unpaved road, and cluster 4 contains data recorded on tarmac road. Unlike the other clusters, cluster 3 contains data recorded coming from both unpaved and tarmac road. The mean travel speed for each cluster was also calculated ($\overline{v_{c1}} = 6.7 km/h$; $\overline{v_{c2}} = 15.4$ km/h; $\overline{v_{c3}} = 21.1$ km/h; $\overline{v_{c4}} = 23.7$ km/h).

To continue with this analysis, we can see that the driver clearly adapts his speed to the type of surface. Therefore, on "off-road," the maximum speed the driver reaches is lower than that on regular surfaces. On the other hand, it is observed that (1) on the same type of surface, the higher the speed is, the parameter S_d^A increases and (2) the greater the surface irregularity is, parameter S_d^A increases at the same speed.

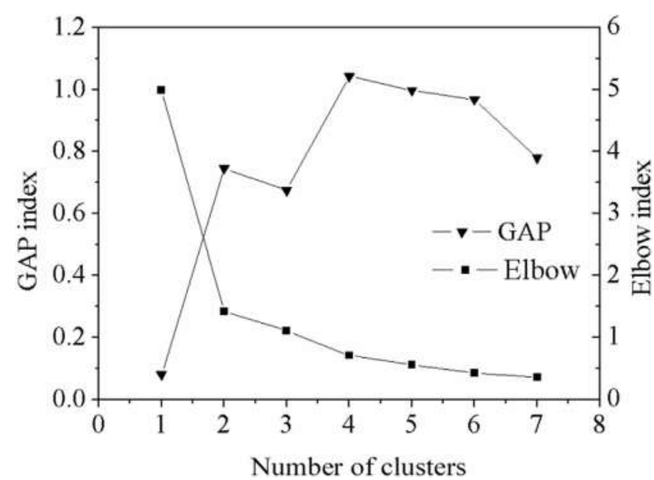

Fig. 3. Selection of the optimum number of clusters.

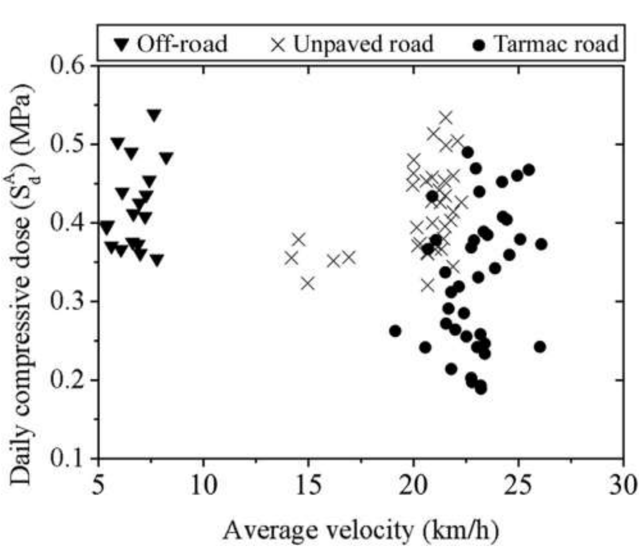

Fig. 2. S_d^A versus average velocity.

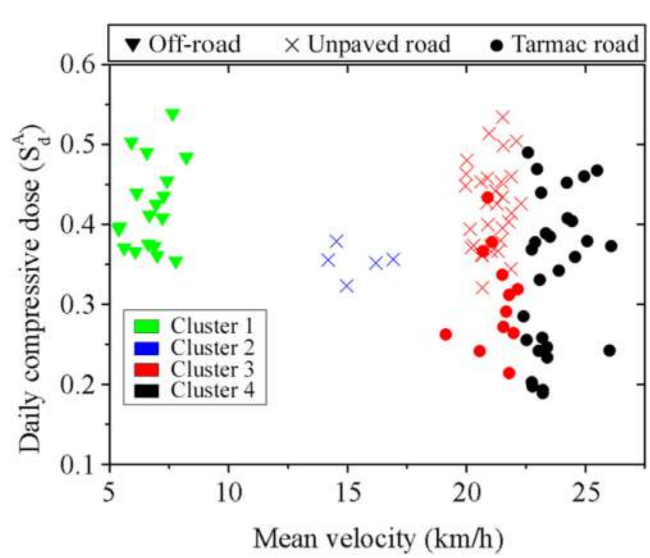

Fig. 4. K-means clustering of S_d^A in terms of mean speed.

4.3. Assessment of daily exposure. Step 3

On the basis of the acceleration data taken from the field experiment, the methods defined in ISO 2631-1 based on A(8) and VDV parameters were applied, as explained in Section 2. With these parameters, we performed an analysis considering the different clusters obtained in the previous section. First, the A(8) parameter and the highest S_d^A value of the vertebral level, calculated using the samples belonging to each cluster, were compared (Fig. 5), and their results were compared with those coming from the HGCZ boundaries associated with probabilities of adverse health effects. The figure shows a linear relationship between the two parameters, and the results show similar assessments in clusters 2 and 4, regarding the values obtained for both parameters. Notably, two data are above the A(8) HGCZ boundary in cluster 1, and only one data is above the S_d^A HGCZ boundary in the same cluster. In addition, there are three data in cluster 3 that exceed the S_d^A HGCZ boundary and two data above the A(8) HGCZ boundary.

Secondly, the VDV parameter and the highest S_d^A value of the vertebral level were calculated. Fig. 6 shows the relationship between both parameters for each of the predefined clusters. The general observation is that both parameters are correlated. For the VDV, this methodology is more restrictive than that based on S_d^A. Thus, it can be seen that the data set of cluster 1 as well as some data of clusters 3 and 4 exceed the VDV HGCZ boundary. It should be noted that only two samples of cluster 1 and three samples of cluster 3 exceeded the S_d^A HGCZ boundary with respect to the S_d^A.

Therefore, health risks predicted by the VDV assessment method are higher than those predicted by A(8)- and S_d^A-based methods. In our experiment, in which the worker exposure contains multiple shocks, this result can be explained as VDV values would be more restrictive than A(8) because the VDV method is more sensitive to shocks. However, there are two methods (VDV and S_d^A), both assessing WBV exposure but providing different assessments when data contain shocks. This can lead to a remarkable confusion as some operations could be considered safe when they are not, and vice versa, depending on the chosen assessment method.

In fact, similar results, but with previous standards, were provided by other research studies on the previous standard ISO 2631-5:2004 and ISO 2631-1 for load-haul-dump vehicles (Eger et al., 2008), railroad locomotives (Cooperrider & Gordon, 2006; Johanning et al., 2006), and front-end loader (Blood et al., 2012). Eger et al. (2008) already suggested that research should be conducted to discuss whether the limits for low and high probabilities of adverse health effects suggested in ISO 2631-5:2004 would require some revision. However, although the method of calculation of a compressive dose is different in ISO 2631-5:2018, the limits of low and high probability of an adverse health effect are the same as those published in ISO2631-5:2004; thus, this result supports this argument in a different context.

4.4. Assessment of cumulative WBV exposure over the years. Step 4

For the data set of each cluster, the average S_d^A value was calculated for the vertebral levels T12/L1 to L5/S1 (Table 2), according to the surface on which they were measured. As the k-means++ method splits the data into nonoverlapping groups, and the Elbow and Gap criteria were applied to select the number of clusters, the

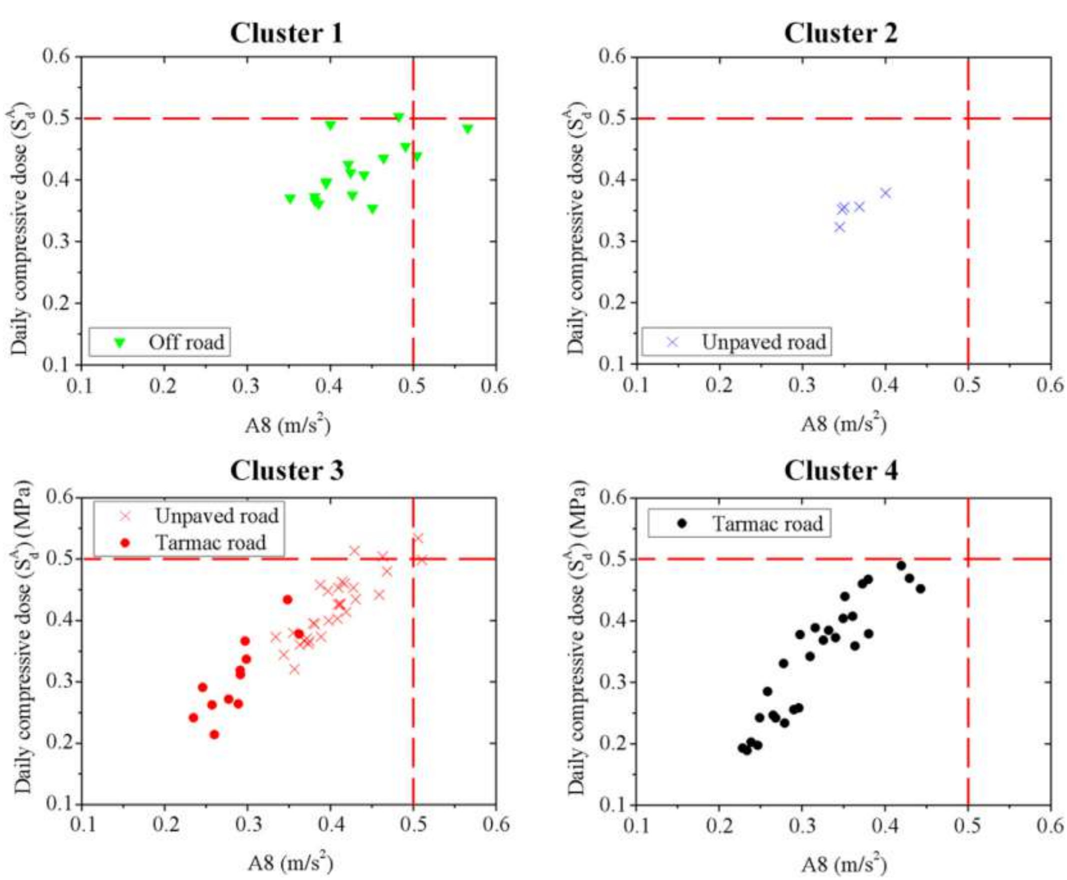

Fig. 5. Relationship between A(8) and S_d^A in the four generated clusters.

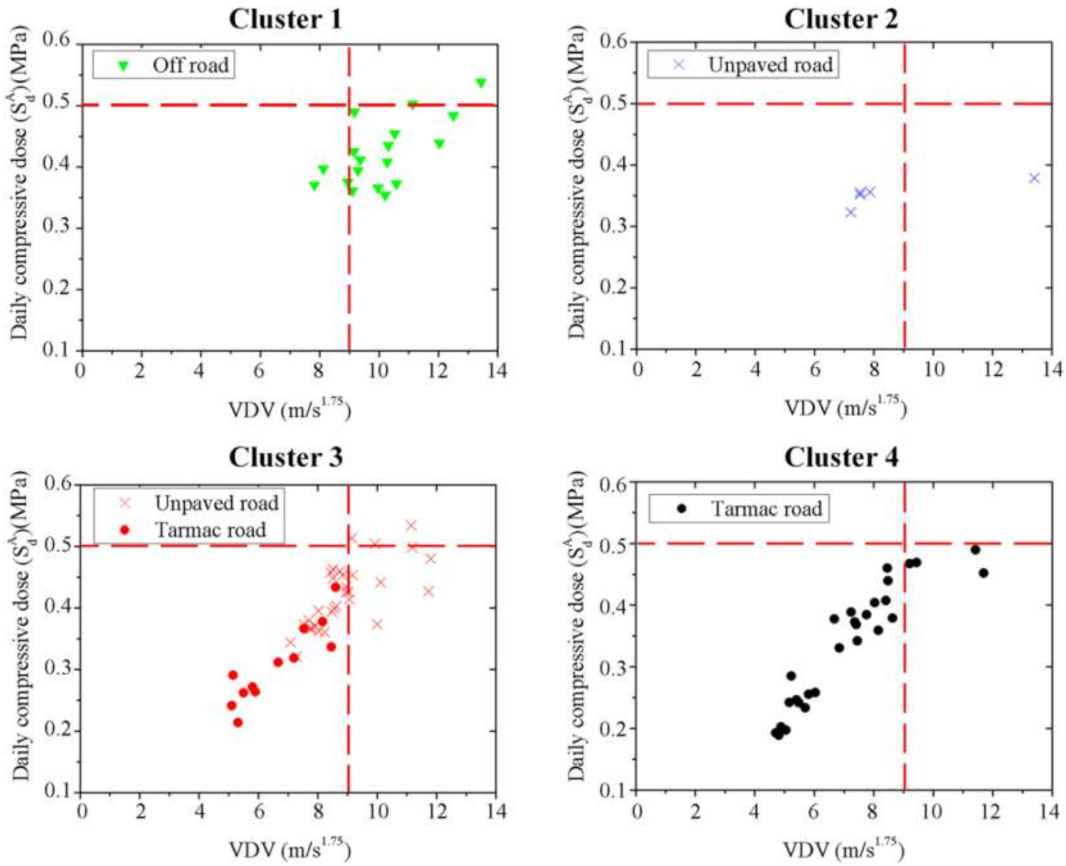

Fig. 6. Relationship between VDV and S_d^A in the four generated clusters.

data obtained are statistically different by the attributes used in the clustering procedure. If we compare the maximum values S_d^A with the HGCZ value, all values indicate a low probability of an adverse health effect to occur after vibration exposition. The maximum S_d^A value of each case was used to calculate the R^A parameter as defined in Section 2, considering the exposure lasts from the age of 20–70 years for 240 days per year (full working life with a typical length of the working day in this sector). The evolution of the R^A values over exposure time is shown in Fig. 7.

By analyzing the evolution of the R^A parameter over exposure time, the WBV exposure daily pattern per year allows us to predict when the subject will exceed the boundary values associated with low and high probabilities of adverse health effects. In this case, the curve C-I (off-road – cluster 1) reaches the limit when the driver is 61 years old, and C-II (unpaved road – cluster 3) reaches the limit when the driver is 62 years old. The boundary exceeded in none of the other cases.

Further analysis and comparison on these curves, and specifically those corresponding to the same type of surface, also allow us to note how the slope of the curve increases as the speed increases. This implies a higher probability of occurrence of an adverse health effect. Therefore, speed is an important factor to consider when trying to reduce the severity of the exposure. In addition, as the irregularity increases, so does the risk. The nature of the terrain and the characteristics of the activity have a great impact on the magnitude of the vibrations transmitted, and they are both very relevant factors. Although the off-road terrain is compressive in a majority of cases and accordingly the level of transmitted vibration would become lower, in our study, the high surface irregularity and the unevenness of the off-road terrain result in an increased severity of the transmitted vibration in com-

Table 2
S_d^A values for the vertebral levels T12/L1 to L5/S1 for the defined clusters.

Surface	Cluster		T12/L1	L1/L2	L2/L3	L3/L4	L4/L5	L5/S1	Max
Off-road	1	S_d^A	0.385	0.393	0.406	0.421	0.414	0.382	0.421
		σ	0.048	0.049	0.052	0.055	0.054	0.045	
Unpaved road	2	S_d^A	0.353	0.330	0.320	0.320	0.315	0.311	0.353
		σ	0.020	0.020	0.022	0.022	0.022	0.018	
	3	S_d^A	0.419	0.408	0.401	0.406	0.400	0.381	0.419
		σ	0.054	0.050	0.049	0.049	0.048	0.046	
Tarmac road	3	S_d^A	0.301	0.291	0.291	0.296	0.292	0.283	0.301
		σ	0.063	0.059	0.061	0.064	0.063	0.060	
	4	S_d^A	0.335	0.318	0.310	0.313	0.305	0.296	0.335
		σ	0.099	0.091	0.085	0.085	0.082	0.081	

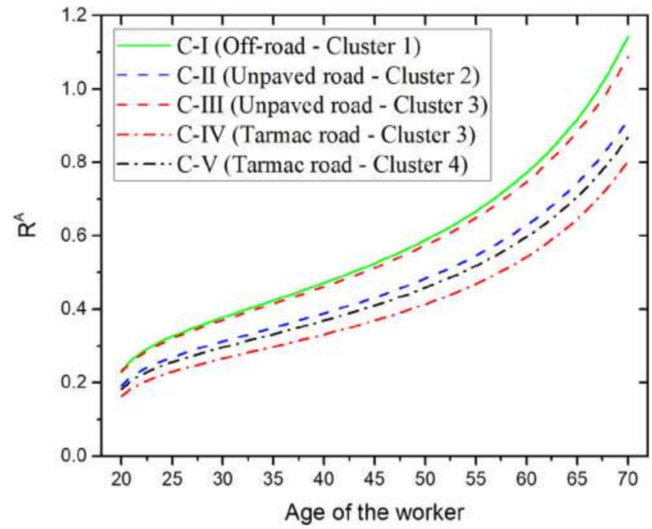

Fig. 7. Evolution of R^A values over the exposure time.

parison to other surfaces, even though the forward speed is lower than that in other surfaces. This fact is noted by comparing the curves CI (off-road), C-II/C-III (unpaved road), and C-IV/C-V (tarmac road).

Because this result deserves attention and will prove to be useful to establish a methodology to predict the probability of an adverse health effect, the data obtained were fitted using polynomial fitting tools in MATLAB software (Table 3) in order to model the cumulative effect of WBV with the parameter R^A. Eq. (9) shows the equation of the general model polynomial fitting function applied to our data. The degree of the obtained polynomial fitting becomes three; this is due to the fact that greater degrees do not improve the goodness of the fit, causing an overfitting or badly conditioned problem.

$$R^A(t) = p1 \cdot t^3 + p2 \cdot t^2 + p3 \cdot t + p4 \quad (9)$$

t is the exposure time. The above polynomial functions can be used to predict the risk factor in terms of the exposure time of worker, depending on the type of road.

4.5. Color map of risk factor evolution. Step 5

Specifically, the proposed method relies on the analysis carried out to obtain Figs. 8 and 9. The parameterization of data in Table 3 allows the health risk probability to which the worker is exposed in any year to be determined, from a data set measured under specific velocity and surface conditions. Based on the experimental data for each type of road surface and exposure time, a bidimensional surface-type fitting was carried out using a polynomial surface model, where the *x*-axis is the number of years of exposure and the *y*-axis represents the average speed at which the activity is performed; thus, the bidimensional polynomial models for surfaces are given by Eq. (10). In this equation, the polynomial surface fitting coefficients and the goodness of fit statistics are shown in

Table 3
Coefficients of the general model polynomial fitting model.

Curve		*p*1	*p*2	*p*3	*p*4	Goodness of fit R-square
C-I	Off-road Cluster 1	1.034e−05	−0.001123	0.04978	−0.3877	0.99
C-II	Unpaved road Cluster 2	8.094e−06	−0.00884	0.03966	−0.3043	0.99
C-III	Unpaved road Cluster3	9.607e−06	−0.001049	0.04708	−0.3612	0.99
C-IV	Tarmac road Cluster 3	6.901e−06	−0.0007538	0.03382	−0.2565	0.99
C-V	Tarmac road Cluster 4	7.681e−06	−0.000839	0.03764	−0.2888	0.99

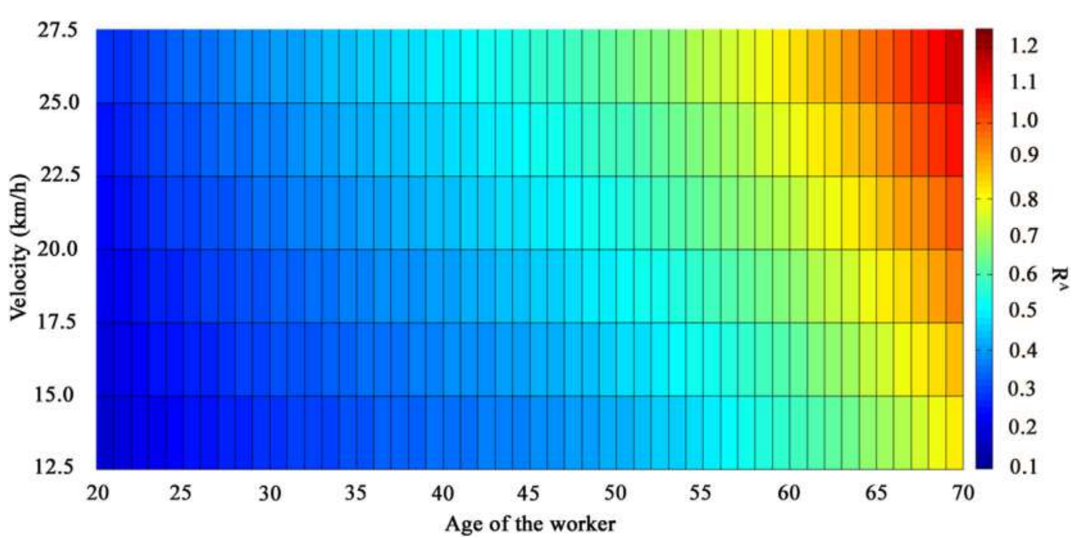

Fig. 8. Unpaved road.

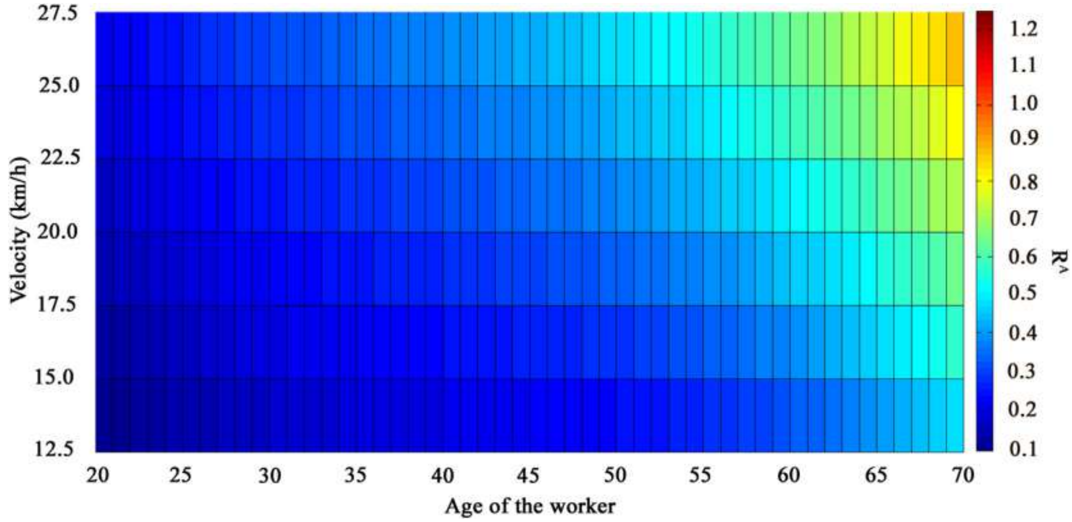

Fig. 9. Tarmac road.

Table 4
Coefficients (with 95% confidence bounds).

	p_{00}	p_{10}	p_{01}	p_{20}	p_{11}	p_{30}	p_{21}	Goodness of fit R-square
Unpaved road	−0.4566	0.04434	0.006797	−0.001055	−5.324e−05	8.85e−06	4.834e−06	0.99
Tarmac road	−0.4578	0.03717	0.008178	−0.000927	−6.406e−05	7.291e−06	5.816e−06	0.99

Table 4. The degree of the obtained polynomial fitting becomes three in the x-axis and one in the y-axis.

$$R^A = f(x,\ y) = p_{00} + p_{10} \cdot x + p_{01} \cdot y + p_{20} \cdot x^2 + p_{11} \cdot x \cdot y + p_{30} \cdot x^3 + p_{21} \cdot x^2 \cdot y \quad (10)$$

A graphic interpretation is useful to extract conclusions and could be done by representing the z-axis by a color map that shows the R^A value. This analysis deepens into the evolution of the cumulative effect of WBV exposure for displacements in a wide range of velocities on different surface types (Fig. 8 and Fig. 9).

It should be noted that this process was carried out for displacements performed on unpaved road and tarmac road. As off-road operations occur in a reduced speed range because they demand continuous concentration and require conscious decision-making to choose the appropriate acceleration, trajectory, etc., off-road surface has not been included in the analysis.

It should be noted that Fig. 8 and Fig. 9 are the result of applying the proposed methodology to this study case. Therefore, the minimum speed was 12.5 km/h because it was the minimum forward speed when the real task was performed. Consequently, the results obtained in the illustrative case should not simply be generalized to other different activities. In any case, the risk assessment professionals should apply the proposed methodology following all the proposed steps when those activities require WBV risk management.

4.6. Risk management. Step 6

The proposed methodology relies on using Figs. 8 and 9 to estimate the health risk probability (based on the risk factor R^A) to which the worker is exposed in terms of the exposure years and a different range of speeds, with no need to calculate the dose, which is sometimes a difficult task. Using the color maps obtained for each vehicle, the safety manager could perform a quick evaluation of the worker's activities according to the average speed at which they are performed, with an entire working life perspective that covers the whole life of the subject. So, the characterization of the activities of the worker, together with the use of R^A curves and surfaces, allows for performing a quick analysis that ensures the safety of the worker.

Finally, the objective of the proposed procedure is aligned with that defined in the Framework Directive 89/391/EEC, which states that the employer must implement measures to ensure an improvement in the level of protection of workers as well as the establishment of the necessary organization and means. The application of this methodology provides vital information to ensure that WBV risk management is carried out correctly. Using these methodological basis, it will be very easy to perform an appropriate organization of the work and the design of the operations in an optimal way from the point of view of the worker health and safety in long term. In addition, safety managers have a simple tool that allows them to define preventive organizational measures that should result in a reduction of WBV risk exposure. In this way, it is ensured that workers could carry out their operations by limiting themselves to the safe region given by this method, emphasizing that the whole working life of the subject has been considered in the development of this procedure.

5. Conclusions and practical applications

Musculoskeletal disorders have a high prevalence among occupational populations as well as a high economic and social impact. Whole-body vibration (WBV) exposure is related to the emergence of musculoskeletal disorders and degeneration of the lumbar spine; therefore, international standards have focused on its assessment. In this sense, ISO 2631-1 and ISO 2631-5 describe models for assessing exposure to WBV. In this research, WBV exposure was analyzed using both models to assess the WBV exposure associated with a HEV operation on a variety of surfaces and speeds. Based on the obtained results and the proposed modeling of the risk factor of occurrence of adverse health effects as established in the standard ISO 2631-5:2018, a methodology was developed to perform a quick evaluation of risks due to the cumulative effect of WBV exposure associated with HEV operation as a function of HEV speed.

The research performed in this paper allows us to draw two main conclusions on the basis of the results and methods discussed in the previous sections:

The first main conclusion is that the results obtained in the evaluation with the A(8), VDV y S_d^A methods provide different assessments leading to different possible consideration of safe operations when the risks associated with them may actually be high. Although the ISO2631-5:2018 methods have been modified from ISO2631-5:2004, HGCZ boundaries should be revised for the sake of consistency.

The second main conclusion is that a method was proposed to assess the health risk probability to which the HEV worker is exposed in terms of the exposure years and a different range of speeds. The methodology proposed in this study supports the design of activities performed with HEV, ensuring that the probability of an adverse health effect is low in the entire working life of the driver. In addition, this methodology reduces the computational time that would require recalculating the S_d^A and R^A values associated with other speed values, as they were calculated from the parameterized R^A curves.

Furthermore, this methodology can contribute to improving the quality of life of professional drivers during and after their working life as this method can be applied from the start of the first job in which the worker is exposed to WBV. In the context of the increase in life expectancy and raising of retirement ages that make suffering from WBV-related diseases more likely, it is very important to consider the entire working life as this method does. Finally, the designed methodology contributes to the development of the EU Strategic Framework on Health and Safety at Work 2014-2020 (Brussels, 6.6.2014COM(2014)) in two of its three major health and safety at work challenges (i.e., it allows the improvement of the prevention of work-related diseases, considering the aging of the EU workforce).

Finally, the proposed methodology is designed and configured to be a practical tool to support safety and health professionals in their objective of assessing the global exposition to WBV throughout their working life. It should be noted that because other relevant factors contribute to the long-term occupational health hazards, a comprehensive health surveillance to prevent possible negative effects on workers must always be conducted by professionals, supplementing the procedure described in this article.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

María Luisa de la Hoz Torres and Antonio Aguilar Aguilera wish to thank the support of the Ministerio de Ciencia, Innovación y Universidades of Spain under an FPU grant. This work has been supported by "Junta de Andalucía" (Spain) and the European Regional Development Fund (ERDF) under project B-TEP-362-UGR18.

References

Adam, S. A., & Jalil, N. A. A. (2017). Vertical suspension seat transmissibility and SEAT values for seated person exposed to whole-body vibration in agricultural tractor preliminary study. *Procedia engineering, 170*, 435–442. https://doi.org/10.1016/j.proeng.2017.03.070.

Arthur, D., & Vassilvitskii, S., (2007). K-means++: The Advantages of Careful Seeding. In SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 1027–1035.

Bakusic, J., Lenderink, A., Lambreghts, C., Vandenbroeck, S., Verbeek, J., Curti, S., Mattioli, S. & Godderis, L., (2018). Alert and sentinel approaches for the identification of work-related diseases in the EU. European Risk Observatory Report. In Publications Office of the European Union. (Ed.). European Agency for Safety and Health at Work, Luxembourg.

Blood, R. P., Rynell, P. W., & Johnson, P. W. (2012). Whole-body vibration in heavy equipment operators of a front-end loader: Role of task exposure and tire configuration with and without traction chains. *Journal of Safety Research, 43*(5–6), 357–364. https://doi.org/10.1016/j.jsr.2012.10.006.

Bovenzi, M., & Betta, A. (1994). Low-back disorders in agricultural tractor drivers exposed to whole-body vibration and postural stress. *Applied Ergonomics, 25*(4), 231–241. https://doi.org/10.1016/0003-6870(94)90004-3.

Burström, L., Nilsson, T., & Wahlström, J. (2015). Whole-body vibration and the risk of low back pain and sciatica: A systematic review and meta-analysis. *International Archives of Occupational and Environmental Health, Health, 88*(4), 403–418. https://doi.org/10.1007/s00420-014-0971-4.

Cooperrider, N., & Gordon, J. (2006). Shock and impact on North American locomotives evaluated with ISO 2631 parts 1 and 5. *Proceedings of the First American Conference on Human Vibration*.

De la Hoz Torres, M., López-Alonso, M., Padillo, D.R. & Martínez-Aires, M., (2017). Analysis of whole-body vibrations transmitted by earth moving machinery, Occupational Safety and Hygiene V: Selected papers from the International Symposium on Occupational Safety and Hygiene (SHO 2017), April 10-11, 2017, Guimarães, Portugal. CRC Press, p. 453. 10.1201/9781315164809.

De la Hoz-Torres, M. L., Aguilar-Aguilera, A. J., Martnez-Aires, M. D., & Ruiz, D. P. (2019). Assessment of whole-body vibration exposure using ISO2631-1: 2008 and ISO2631-5: 2018 standards. *Paper presented at the INTER-NOISE and NOISE-CON Congress and Conference Proceedings*.

Costa, N., Arezes, P., & Melo, R. (2014). Effects of occupational vibration exposure on cognitive/motor performance. *International Journal of Industrial Ergonomics, 44* (5), 654–661. https://doi.org/10.1016/j.ergon.2014.07.005.

Directive 78/764/EEC of 25 July 1978 on the approximation of the laws of the Member States relating to the driver's seat on wheeled agricultural or forestry tractors. Council of European Union. Retrieved from https://eur-lex.europa.eu/homepage.html.

Directive of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (89/391/EEC). Retrieved from https://eur-lex.europa.eu/homepage.html.

Directive 2002/44/EC of the European Parliament and of the Council of 25 June 2002 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (vibration). Retrieved from https://eur-lex.europa.eu/homepage.html.

Eger, T., Stevenson, J., Boileau, P.-É., & Salmoni, A. (2008). Predictions of health risks associated with the operation of load-haul-dump mining vehicles: Part 1—Analysis of whole-body vibration exposure using ISO 2631–1 and ISO-2631-5 standards. *International Journal of Industrial Ergonomics, 38*(9-10), 726–738. https://doi.org/10.1016/j.ergon.2007.08.012.

Giordano, D. M., Facchinetti, D., & Pessina, D. (2015). Comfort efficiency of the front axle suspension in off-road operations of a medium-powered agricultural tractor. *Contemporary Engineering Sciences, 8*, 1311–1325. https://doi.org/10.12988/ces.2015.56186.

Health and Safety Executive, (2018). Health and safety at work. Summary statistics for Great Britain 2017. Health and safety Executive, Great Britain.

International Organisation for Standardisation (2004). *ISO 2631-5 Mechanical Vibration and Shock – Evaluation of Human Exposure to Whole-body Vibration – Part 5: Method for Evaluation of Vibration Containing Multiple Shocks*. Geneva, Switzerland: ISO.

International Organisation for Standardisation (2008). *ISO 2631-1 Mechanical Vibration and Shock – Evaluation of Human Exposure to Whole-body Vibration – Part 1: General Requirements*. Geneva, Switzerland: ISO.

International Organisation for Standardisation (2018). *ISO 2631-5 Mechanical Vibration and Shock – Evaluation of Human Exposure to Whole-body Vibration – Part 5: Method for Evaluation of Vibration Containing Multiple Shocks*. Geneva, Switzerland: ISO.

Griffin, M., Howarth, H., Pitts, P., Fischer, S., Kaulbars, U., Donati, P., & Bereton, P. (2008). Non-binding guide to good practice with a view to implementation of directive 2002/44/EC on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (vibrations). *European Commission*. https://doi.org/10.2767/41117.

Johanning, E., Fischer, S., Christ, E., Gores, B., & Luhrman, R. (2006). Railroad locomotive whole-body vibration study: Vibration, shocks and seat ergonomics. *Proceedings of the first american conference on human vibration*.

Johnson, P. W., Dennerlein, J., Ramirez, L. M., Arias, C., & Rodríguez, A. C. (2015). Assessment of continuous and impulsive whole body vibration exposures in heavy equipment mining vehicles. In *Proceedings 19th Triennial Congress of the IEA* (pp. 14).

Kia, K., Fitch, S. M., Newsom, S. A., & Kim, J. H. (2020). Effect of whole-body vibration exposures on physiological stresses: Mining heavy equipment applications. *Applied Ergonomics, 85*, 103065. https://doi.org/10.1016/j.apergo.2020.103065.
Kim, J. H., Dennerlein, J. T., & Johnson, P. W. (2018). The effect of a multi-axis suspension on whole body vibration exposures and physical stress in the neck and low back in agricultural tractor applications. *Applied Ergonomics, 68*, 80–89. https://doi.org/10.1016/j.apergo.2017.10.021.
Kittusamy, N. K., & Buchholz, B. (2004). Whole-body vibration and postural stress among operators of construction equipment: A literature review. *Journal of Safety Research, 35*(3), 255–261. https://doi.org/10.1016/j.jsr.2004.03.014.
Kumar, A., Varghese, M., Mohan, D., Mahajan, P., Gulati, P., & Kale, S. (1999). Effect of whole-body vibration on the low back: A study of tractor-driving farmers in north India. *Spine, 24*(23), 2506. https://doi.org/10.1097/00007632-199912010-00013.
Liuke, M., Solovieva, S., Lamminen, A., Luoma, K., Leino-Arjas, P., Luukkonen, R., & Riihimäki, H. (2005). Disc degeneration of the lumbar spine in relation to overweight. *International Journal of Obesity, 29*(8), 903–908. https://doi.org/10.1038/sj.ijo.0802974.
Livshits, G., Popham, M., Malkin, I., Sambrook, P. N., MacGregor, A. J., Spector, T., & Williams, F. M. (2011). Lumbar disc degeneration and genetic factors are the main risk factors for low back pain in women: The UK Twin Spine Study. *Annals of the Rheumatic Diseases, 70*(10), 1740–1745. https://doi.org/10.1136/ard.2010.137836.
Meucci, R. D., Fassa, A. G., & Faria, N. M. (2015). Prevalence of chronic low back pain: Systematic review. *Revista de Saude Publica, 49*, 1. https://doi.org/10.1590/S0034-8910.2015049005874.
Milosavljevic, S., Bagheri, N., Vasiljev, R. M., Mcbride, D. I., & Rehn, B. (2011). Does daily exposure to whole-body vibration and mechanical shock relate to the prevalence of low back and neck pain in a rural workforce? *Annals of Work Exposures and Health, 56*(1), 10–17. https://doi.org/10.1093/annhyg/mer068.
Milosavljevic, S., Bergman, F., Rehn, B., & Carman, A. B. (2010). All-terrain vehicle use in agriculture: Exposure to whole body vibration and mechanical shock. *Applied Ergonomics, 41*(4), 530–535. https://doi.org/10.1016/j.apergo.2009.11.002.
Miyamoto, M., Shirai, Y., Nakayama, Y., Gembun, Y., & Kaneda, K. (2000). An epidemiologic study of occupational low back pain in truck drivers. *Journal of Nippon Medical School, 67*(3), 186–190. https://doi.org/10.1272/jnms.67.186.
Nielsen, K., Jørgensen, M.B. & Malgorzata Milczarek, L.M., (2018). Healthy workers, thriving companies - a practical guide to wellbeing at work, in: work, European Agency for Safety and Health at Work. (Ed.), Luxembourg.
Paddan, G. S., & Griffin, M. J. (2002). Evaluation of whole-body vibration in vehicles. *Journal of Sound and Vibration, 253*(1), 195–213. https://doi.org/10.1006/jsvi.2001.4256.
Paschold, H. W., & Sergeev, A. V. (2009). Whole-body vibration knowledge survey of US occupational safety and health professionals. *Journal of Safety Research, 40*(3), 171–176. https://doi.org/10.1016/j.jsr.2009.02.008.
Punnett, L., & Wegman, D. H. (2004). Work-related musculoskeletal disorders: The epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology, 14*(1), 13–23. https://doi.org/10.1016/j.jelekin.2003.09.015.
Raffler, N., Rissler, J., Ellegast, R., Schikowsky, C., Kraus, T., & Ochsmann, E. (2017). Combined exposures of whole-body vibration and awkward posture: A cross sectional investigation among occupational drivers by means of simultaneous field measurements. *Ergonomics, 60*(11), 1564–1575. https://doi.org/10.1080/00140139.2017.1314554.
Rehn, B., Nilsson, T., Lundström, R., Hagberg, M., & Burström, L. (2009). Neck pain combined with arm pain among professional drivers of forest machines and the association with whole-body vibration exposure. *Ergonomics, 52*(10), 1240–1247. https://doi.org/10.1080/00140130902939889.
Shiri, R., Karppinen, J., Leino-Arjas, P., Solovieva, S., & Viikari-Juntura, E. (2009). The association between obesity and low back pain: A meta-analysis. *American Journal of Epidemiology, 171*(2), 135–154. https://doi.org/10.1093/aje/kwp356.
Singh, A., Nawayseh, N., Singh, L., Singh, S., & Singh, H. (2019). Investigation of Compressive Stress on Lumbar Spine due to Whole Body Vibration Exposure in Rotary Tillage Operation. International Journal of Automotive and Mechanical Engineering, 16(2), 6684-6696. 10.15282/ijame.16.2.2019.16.0503.
Smets, M. P., Eger, T. R., & Grenier, S. G. (2010). Whole-body vibration experienced by haulage truck operators in surface mining operations: A comparison of various analysis methods utilized in the prediction of health risks. *Applied Ergonomic, 41*(6), 763–770. https://doi.org/10.1016/j.apergo.2010.01.002.
Spielholz, P., Cullen, J., Smith, C., Howard, N., Silverstein, B., & Bonauto, D. (2008). Assessment of perceived injury risks and priorities among truck drivers and trucking companies in Washington State. *Journal of Safety Research, 39*(6), 569–576. https://doi.org/10.1016/j.jsr.2008.09.005.
Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63*(2), 411–423. https://doi.org/10.1111/1467-9868.00293.
Wilder, D. G., Aleksiev, A. R., Magnusson, M. L., Pope, M. H., Spratt, K. F., & Goel, V. K. (1996). Muscular response to sudden load: A tool to evaluate fatigue and rehabilitation. *Spine, 21*(22), 2628–2639. https://doi.org/10.1097/00007632-199611150-00013.
Woolf, A. D., & Pfleger, B. (2003). Burden of major musculoskeletal conditions Retrieved from. *Bulletin of the world health organization, 81*, 646–656 https://apps.who.int/iris/handle/10665/269026.
Yazdani, A., Hilbrecht, M., Imbeau, D., Bigelow, P., Neumann, W. P., Pagell, M., & Wells, R. (2018). Integration of musculoskeletal disorders prevention into management systems: A qualitative study of key informants' perspectives. *Safety Science, 104*, 110–118. https://doi.org/10.1016/j.ssci.2018.01.004.

María L. de la Hoz-Torres is a PhD candidate at the University of Granada and research assistant in the research group TEP-968 Technologies for a circular economy. She completed her Ms. in Safety Management and Acoustics Engineering in 2016 and 2017 respectively. She has a full time research grant. Her research focusses in planning for a healthy work environment through risk-based design methods.

Antonio J. Aguilar is a PhD student in Civil Engineering at the University of Granada and research assistant in the research group TEP-968 Technologies for a circular economy. He has a full time research grant. His research focusses in development of tools for analysis and improvement of indoor environmental quality in workplaces.

M. D. Martínez-Aires received the Bs in Organization Industrial Engineering in 2004 and Technical Architecture in 1993. The PhD degree in Civil Engineering in 2009 from the University of Granada, Spain. Currently, she is an Associate Professor in the Department of Building Construction of the University of Granada. She has been co-director of Management and Integral Safety in Buildings Master's Degree (2007-2017). She has authored many articles published in international journals, national and international conference proceedings and book chapters, always in the field of Occupational Risk Prevention. She is a featured reviewer in journals from Elsevier, John Wiley & Sons and MPDI. Her research interests include Health and Safety, Prevention through Design (PtD), Nanomaterials, Ergonomics and BIM.

Diego P. Ruiz received the Bs in Physics with a major in Electronics in 1991 and the Ph.D. in 1995 from the University of Granada, Spain. Currently, he is Associate Professor in the Applied Physics Department of the University of Granada. He is director of Acoustics Engineering Master's Degree and professor in a Master's course in occupational hazard prevention. He has authored more than 140 articles in international journals and more than 150 publications in national and international conferences. He has been advisor of 11 doctoral thesis, more than 78 or final degree/masters' research projects, and has been the main researcher of 6 research projects and member of 15 research projects. His research has been linked to the modeling of experimental data, signal processing and data analysis, including the development of techniques and algorithms, in fields related to electromagnetism, acoustics and thermal and acoustic insulation in buildings, Health and Safety, risk prevention and its modelling, and BIM. The lines of research have been funding by research public agencies and companies from the public and private sector.

Journal of Safety Research 78 (2021) 242–250

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Analysis of effects of driver's evasive action time on rear-end collision risk using a driving simulator

Dhwani Shah, Chris Lee *

Department of Civil and Environmental Engineering, University of Windsor, ON, N9B 3P4, Windsor, Canada

ARTICLE INFO

Article history:
Received 14 December 2020
Received in revised form 7 March 2021
Accepted 1 June 2021
Available online 15 June 2021

Keywords:
Rear-end collision
Surrogate safety measure
Evasive action
Truck
Driving simulator

ABSTRACT

Introduction: Driver's evasive action is closely associated with collision risk in a critical traffic event. To quantify collision risk, surrogate safety measures (SSMs) have been estimated using vehicle trajectories. However, vehicle trajectories cannot clearly capture presence and time of driver's evasive action. Thus, this study determines the driver's evasive action based on his/her use of accelerator and brake pedals, and analyzes the effects of the driver's evasive action time (i.e., duration of evasive action) on rear-end collision risk. *Method:* Fifty drivers' car-following behavior on a freeway was observed using a driving simulator. An SSM called "Deceleration Rate to Avoid Crash (DRAC)" and the evasive action time were determined for each driver using the data from the driving simulator. Each driver tested two traffic scenarios – Cars and Trucks scenarios where conflicting vehicles were cars and trucks, respectively. The factors related to DRAC were identified and their effects on DRAC were analyzed using the Generalized Linear Models and random effects models. *Results:* DRAC decreased with the evasive action time and DRAC was closely related to drivers' gender and driving experience at the road sections where evasive action to avoid collision was required. DRAC was also significantly different between Cars and Trucks scenarios. The effect of the evasive action time on DRAC varied among different drivers, particularly in the Trucks scenario. *Conclusions:* Longer evasive action time can significantly reduce crash risk. Driver characteristics are more closely related to effective evasive action in complex driving conditions. *Practical Applications:* Based on the findings of this study, driver warning information can be developed to alert drivers to take specific evasive action that reduces collision risk in a critical traffic event. The information is likely to reduce the variability of the driver's evasive action and the speed variations among different drivers.

1. Introduction

Vehicle crashes and passengers' injury severity have been analyzed using historical crash data. However, due to rare occurrence of crashes, it takes long time to collect enough crash data to identify safety problems. Thus, surrogate safety measures (SSMs) estimated from individual vehicle records (e.g., speed, headway) have been used to observe conflicts among vehicles. Since these individual vehicle records can be collected in relatively shorter time period (e.g., several hours or days instead of multiple years), safety problems can be identified and countermeasures to prevent crashes can be implemented faster.

SSMs determine the probability of collision between two vehicles using vehicle trajectories that keep track of individual vehicle's position over time. From the trajectories, vehicle's instantaneous speed and acceleration/deceleration, and their spacing with the lead vehicle can be calculated. This paper will focus on the probability of rear-end collision between the lead and following vehicles.

Conventional temporal proximity-based SSMs represent the time remaining to a collision if the following vehicle driver does not take evasive action (Laureshyn et al., 2017; Mahmud et al., 2017). In this study, evasive action is defined as the driver's action or maneuver to decelerate to avoid a collision with the lead vehicle. Examples of this type of SSM are Time-to-Collision (TTC) (Hayward, 1972) and Post Encroachment Time (PET) (Allen et al., 1978). Temporal proximity-based SSMs essentially reflect the time available for evasive action. Thus, a lower value of these SSMs indicates that shorter time is available for evasive action and, hence, it is more difficult to avoid a collision and collision risk is higher.

However, these SSMs do not account for the factors related to evasive action such as driver characteristics (e.g., perception and reaction time which determines the start time of evasive action) and vehicle performance characteristics (e.g., deceleration rate

* Corresponding author.
E-mail addresses: shah181@uwindsor.ca (D. Shah), cclee@uwindsor.ca (C. Lee).

https://doi.org/10.1016/j.jsr.2021.06.001
0022-4375/

during braking, maximum deceleration rate). Thus, these SSMs do not capture how the driver's evasive action affects collision risk.

In this regard, the relationship between the driver's evasive action and collision risk is illustrated in Fig. 1. In undisturbed traffic, drivers do not take evasive action but it will not lead to a collision (i.e., no collision risk) assuming that they do not make errors. However, in a critical event, drivers take evasive action and it can potentially lead to a collision as one of possible outcomes (Laureshyn, 2017). Li et al. (2019) also found that driver's evasive action (swerving to avoid conflicts with vehicles or pedestrians) was significantly related with collision risk. This shows that it is important to analyze the effects of evasive action on collision risk using SSMs.

Unlike temporal proximity-based SSMs, deceleration-based SSMs account for the driver's evasive action more explicitly (Mahmud et al., 2017). For instance, Time-to-Accident (TA) represents the time remaining to a collision from the time when the driver's evasive action starts (Hyden, 1987). Deceleration Rate to Avoid Crashes (DRAC) (Cooper & Ferguson, 1976) represents the deceleration required for avoiding (or not being able to avoid) crashes if the driver takes evasive action.

These deceleration-based SSMs also reflect the difference in collision risk among different drivers and vehicles. These SSMs incorporate differences in driver's perception and reaction time (PRT) and vehicle's deceleration rate. In particular, PRT is used to estimate the start time of the driver's evasive action after the driver perceives the lead vehicle's motion. For instance, Wu et al. (2018) used the fixed value of PRT (=1.5 s) for calculating both lead and following vehicles' stopping sight distance to determine the rear-end collision risk index (RCRI). Alternatively, Triggs and Harris (1982) assumed that PRT follows a lognormal distribution based on their observed distribution of different drivers' PRT. Wang and Stamatiadis (2013) and Kuang et al. (2015) developed SSMs based on the lognormal distribution of PRT. However, the fixed value or the distribution of PRT may not reflect each driver's actual PRT since PRT does not only vary among different drivers, but also different spacings with the lead vehicle for the same driver (Wang et al., 2016). Thus, each driver's start time of evasive action cannot be accurately determined in this approach.

In addition, it is difficult to determine whether the driver actually took evasive action or not based on vehicle trajectories only. For instance, the speed of vehicle can naturally change even when the driver did not take any action due to variations in road geometric and environmental conditions such as gradient and friction of road surface. Also, the speed may not noticeably change if the driver gradually decelerates by releasing the accelerator pedal or applying brake slowly.

To determine the presence and time of driver's evasive action, vehicle dynamics data can be used. Examples of vehicle dynamics variables are the use of accelerator and brake pedals. However, unlike vehicle trajectories, vehicle dynamics cannot be observed using roadside cameras. Instead, vehicle dynamics data can be collected using a driving simulator. Although some past studies determined SSMs using the data from a driving simulator, they did not use vehicle dynamics data (Levulis et al., 2015). Recently, Gold et al. (2018) evaluated the effect of brake application using driving simulator data. Although the study found that the driver's brake application influenced collision risk, it mainly focused on prediction of probability of brake application in a critical event.

However, there is a limitation in determining SSMs using driving simulator data. For instance, due to the driver's discomfort and motion sickness in virtual traffic environments, the driver's simulated behavior may not be naturalistic and realistic. However, Brooks et al. (2010) found that simulator sickness was not a significant problem for young drivers. Also, a driving simulator can control confounding effects of external factors (e.g., congestion, weather, vehicle composition) and identify the isolated effect of each factor on collision risk unlike field studies. For these reasons, simulator-based SSMs are more advantageous than SSMs determined using roadside cameras.

Moreover, relatively few studies have considered differential effects of the types of lead and following vehicles (e.g., car and truck) on rear-end collision risk. For instance, Li et al. (2017) found that TTC was shorter and rear-end collision risk was higher for car drivers when they followed cars as opposed to trucks at a freeway weaving section. They explained that this is because large trucks generally obstruct car drivers' sight and car drivers maintain longer gap with the lead truck than the lead car. Similarly, Zhao and Lee (2017) found that PET was shorter for car-following-car than car-following-truck due to shorter spacing with the lead car than the lead truck. However, in spite of lower collision risk, the collision involving trucks leads to more severe injury (Zheng et al., 2018).

Considering high severity of truck-involved crashes, it is important to comprehensively understand how drivers behave differently when conflicting vehicles are trucks compared to cars and how this is related to collision risk. Thus, the objective of this study is to analyze the effects of the driver's evasive action and the type of lead vehicle (car or truck) on car driver's rear-end collision risk using both vehicle dynamics and trajectories from a driving simulator.

2. Description of data

To estimate rear-end collision risk, the driver's car-following behavior in different traffic conditions was observed using a driving simulator in June 2016. The driving simulator includes a driver seat, a steering wheel and accelerator/brake pedals in the cab, and three LCD monitors. A total of 50 licensed drivers (34 males and 16

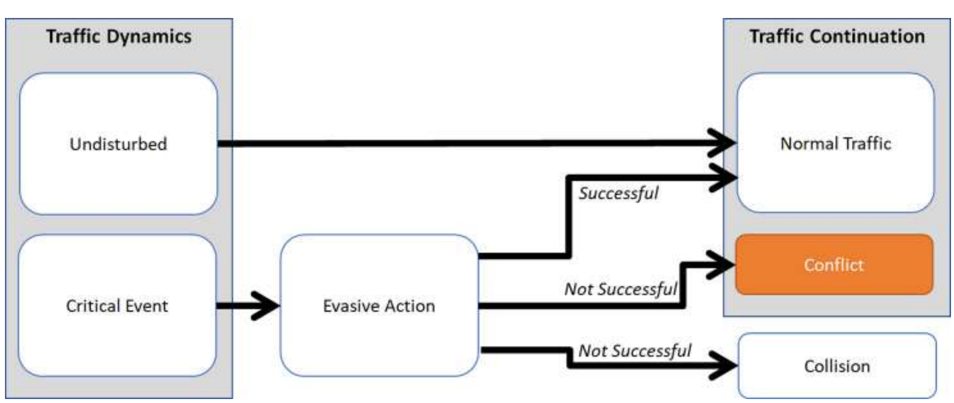

Fig. 1. Relationship between driver's evasive action and collision risk. (Source: Laureshyn et al., 2017).

females) participated in the driving simulator experiment. Among them, 37 drivers were 25 years old or younger, 7 drivers were 26–35, and 6 drivers were above the age of 36 years. The study was approved by the University Research Ethics Board.

Each participant drove a section of two-lane one-way freeway in the following two hypothetical traffic scenarios as a car driver: (1) Cars scenario: all conflicting vehicles are cars (except two stopped trucks in the middle of the freeway), and (2) Trucks scenario: all conflicting vehicles are trucks. The posted speed limit on the freeway was set to 100 km/h. Average speeds of these conflicting cars and trucks were 100 km/h and 90–95 km/h, respectively. In each scenario, the participants encountered different traffic events while driving on the highway as shown in Fig. 2.

In the beginning, the drivers were placed on an entrance ramp. Then the drivers started accelerating to merge into the mainline freeway. Afterwards, the drivers encountered various driving conditions such as approaching two stopped trucks (one stopped truck in each lane) and a merging vehicle from an on-ramp. The freeway was classified into seven different sections based on driving conditions as shown in Fig. 2. The driving conditions in each road section and the length of each section are described as follows:

Section 1: As the drivers merge into the mainline freeway, they see that other cars are approaching them on the freeway. They are required to adjust their speed tactfully to complete the merging process. (Length: 2.51 km)
Section 2: After the merge, drivers can drive normally. They can see trucks ahead but they may not perceive that the trucks are stopped. (Length: 1.15 km)
Section 3: As the drivers approach the two stopped trucks, they are required to reduce speed to completely stop behind one of the trucks. (Length: 3.06 km)
Section 4: As the stopped trucks start moving, the drivers can drive normally without having to change lanes. (Length: 8.72 km)
Section 5: The drivers can see that a vehicle is merging into the mainline freeway from an on-ramp. They can yield to the merging vehicle or accelerate to avoid conflict with the merging vehicle. (Length: 2.24 km)
Section 6: After passing the merging vehicle and before seeing the message that they must exit the freeway, the drivers can drive normally without having to change lanes. (Length: 6.32 km)
Section 7: The drivers see the message that they must exit the freeway. They are required to exit via an off-ramp and stop on the ramp. (Length: 3.62 km)

The participants drove all seven sections in the Cars and Trucks scenarios separately. Traffic kinematics and vehicle dynamics data as shown in Table 1 were extracted at every 1/60th second from the driving simulator. The accelerator pedal position is expressed as a value between 0 and 1. The value gradually increases from 0 to 1 as the driver steps on the accelerator pedal. Similarly, the value of brake pedal force gradually increases as the driver steps on the brake pedal. After the participants completed testing the scenario, they were asked to provide their age and driving experience.

The speed was used to calculate acceleration and deceleration rates. The rates were calculated as the difference between the final speed at the current time frame i and the initial speed at the time frame (i – 60) divided by the number of time frames in one second. Thus, the acceleration and deceleration represent the rate of change in speed per second. Table 2 compares average traffic kinematic variables of all drivers among the seven road sections. The table shows that the mean speed, mean spacing, and maximum deceleration rate significantly varied across different sections due to differences in driving conditions. Also, the differences between the Cars and Trucks scenarios varied across different sections.

It was found that the mean speed was significantly lower in Section 3 than the other sections. This is expected because the drivers were required to reduce speed and stop behind the trucks in Section 3.

It was also found that both mean speed and mean spacing were lower for the Cars scenario than the Trucks scenario for Sections 3 and 7 unlike the other sections. On the other hand, the mean speed difference between the lead and subject vehicles was significantly higher for the Trucks scenario than the Cars scenario in Section 3. Table 2 also shows that the average of maximum deceleration rate was generally higher (in absolute values) for the Trucks scenario than the Cars scenario. This reflects that the drivers tend to apply harder deceleration to avoid conflicts with the lead truck than the lead car. As expected, the absolute values of maximum decel-

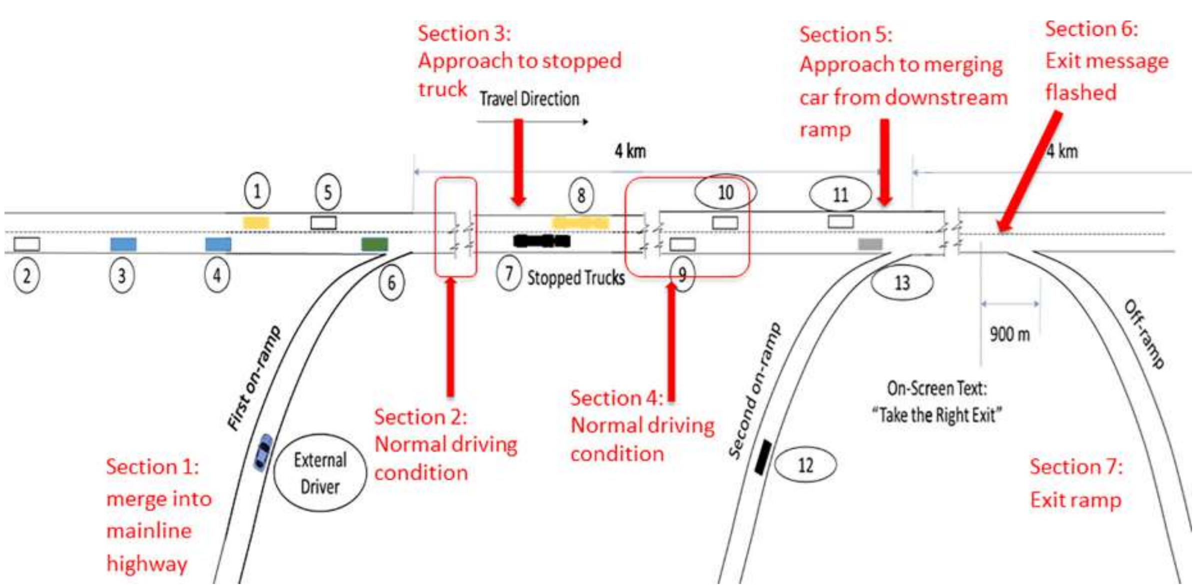

Fig. 2. Classification of road sections in driving simulator experiment. (*Note:* Numbers denote different conflicting vehicles.)

Table 1
List of variables.

Type	Variables	Description	Descriptive statistics
Traffic kinematics	Speed (km/h)	Speed of the subject vehicle	Mean: 80.7 km/h Maximum: 163.9 km/h Minimum: 0 km/h
	Spacing (m)	Front-to-rear spacing with the lead vehicle	Mean: 123.55 m Maximum: 2622.15 m Minimum: 0 m
Vehicle dynamics	Accelerator pedal position	1 = fully pressed 0 = fully released	Mean: 0.38 Maximum: 1 Minimum: 0
	Brake pedal force (lb)		Mean: 1.82 lb Maximum: 180 lb Minimum: 0 lb
Driver character-istics	Gender	0 = Male, 1 = Female	34 male and 16 female drivers
	Age (years)	Age 1 = 25 or younger Age 2 = 26–35 Age 3 = 36 or older	37 drivers 7 drivers 6 drivers
	Driving experience (years)	Number of years since the first license	Mean: 7.5 years Maximum: 50 years Minimum: 1 year

Table 2
Average traffic kinematic variables by scenario and road section.

Variable	Scenario	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Mean speed (km/h)	Car	101.3	117.1	41.9	96.6	110.5	103.7	98.1
	Truck	93.8	115.0	48.2	98.6	84.5	103.7	104.2
Mean spacing (m)	Car	394.5	285.1	96.0	126.6	157.6	131.4	115.2
	Truck	268.6	286.9	107.4	119.3	82.3	116.0	154.6
Mean speed difference (km/h)	Car	15.6	16.5	28.1	11.8	20.4	10.6	7.5
	Truck	12.1	23.1	41.1	17.0	14.7	11.3	9.3
Max. deceleration (m/s^2)	Car	−1.05	−0.59	−5.97	−1.35	−1.17	−1.32	−1.08
	Truck	−1.19	−0.67	−6.95	−1.70	−1.90	−1.03	−1.12
Mean acceleration (m/s^2)	Car	0.69	0.51	0.65	0.52	0.40	0.36	0.39
	Truck	0.77	0.59	0.85	0.51	0.60	0.40	0.33

eration rate were significantly higher in Section 3 than the other sections for both scenarios due to the stopped trucks. Similarly, the mean acceleration rate was also generally higher for the Trucks scenario than the Cars scenario. This shows that the drivers tend to apply higher acceleration to change lanes to overtake trucks or compensate for the delay caused by trucks.

3. Methods

3.1. Determination of driver's evasive action

In this study, the driver's evasive action was determined based on two vehicle dynamics variables – the accelerator pedal position and the brake pedal force. As the driver controls speed by releasing the accelerator pedal or pressing the brake pedal, the changes in values of accelerator pedal position and brake pedal force indicate the driver's evasive action. For instance, Fig. 3 shows that the profiles of speed, accelerator pedal position, and brake pedal force as one driver approached the stopped trucks in Section 3. The figure shows that the driver occasionally released accelerator pedal or applied brake. When the driver was very close to the stopped trucks, the brake pedal force significantly increased and speed abruptly dropped. This clearly indicates that the driver took evasive action to avoid a collision with the stopped trucks.

However, while the driver was approaching the stopped trucks in a longer distance and gradually decelerating, the speed did not noticeably drop despite the driver's action (i.e., released accelerator pedal, pressed brake pedal with low force). In these conditions, it is difficult to determine when the driver took evasive action solely from the speed profile. This example demonstrates that the driver's use of accelerator and brake pedals is important to determine the time of driver's evasive action.

In fact, the drivers decelerate not only because they want to avoid a collision in emergency situations. It is possible that the drivers decelerate because they do not feel safe or they want to drive more cautiously in non-emergency situations. From a broader perspective, such action can also be considered as the driver's evasive action. In this study, the time period during which the drivers either released the accelerator pedal or applied the brake was defined as the "evasive action time." The method of determining the evasive action is illustrated in Fig. 4.

As there was no true disincentive for the drivers to avoid crashes in the driving simulator experiment, it's possible that the participants drove carelessly, unlike their actual driving. Thus, the validity of the simulator data was evaluated based on the crash rate (=number of crashers per 100 km) as follows. In this study, a total of 7 crashes occurred – 6 crashes in the Trucks scenario and 1 crash in the Cars scenario. Thus, the crash rate in this study is 0.253 crashes/100 km (=7 crashes/2762 km), which is not unrealistically high. This low crash rate indicates that most participants normally and carefully drove similar to their actual driving.

3.2. Estimation of rear-end collision risk

Rear-end collision risk for different drivers and scenarios was analyzed based on the Deceleration Rate to Avoid Crashes (DRAC). DRAC is defined as the minimum deceleration rate required for the

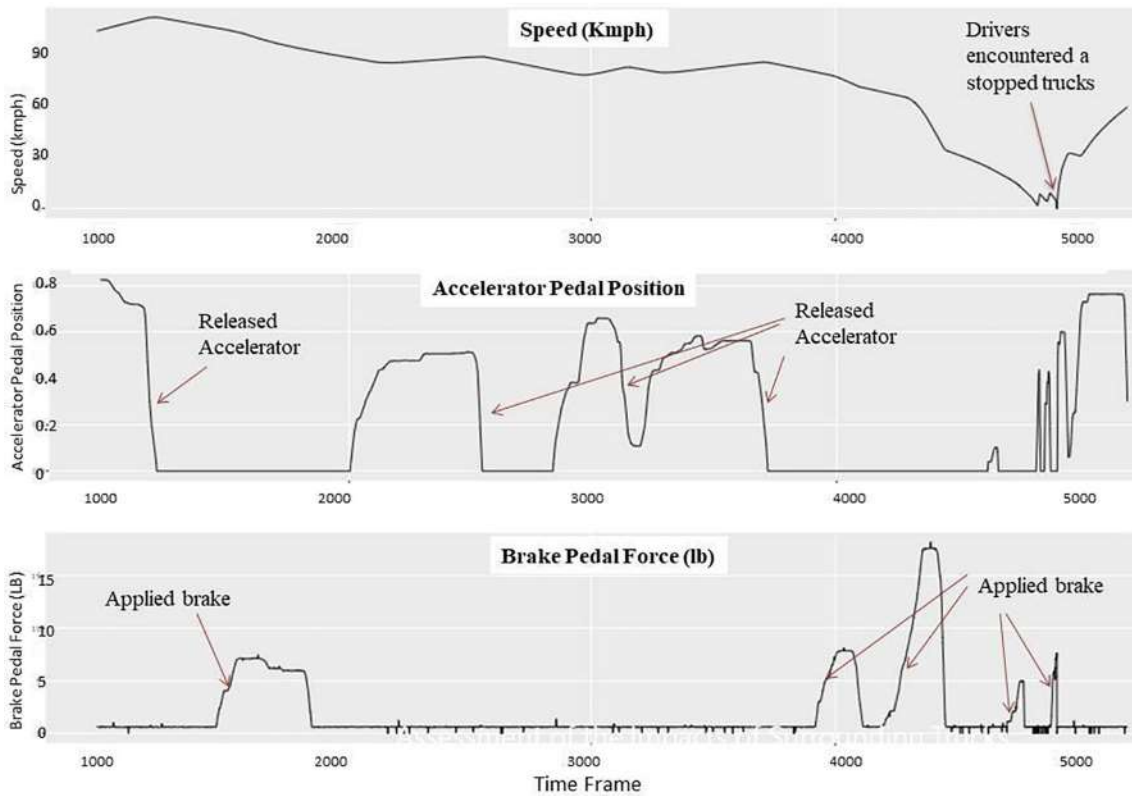

Fig. 3. Profiles of speed, accelerator pedal position and brake pedal force in Section 3. *Note:* The horizontal axis represents the number of time frames (each time frame = 1/60 s).

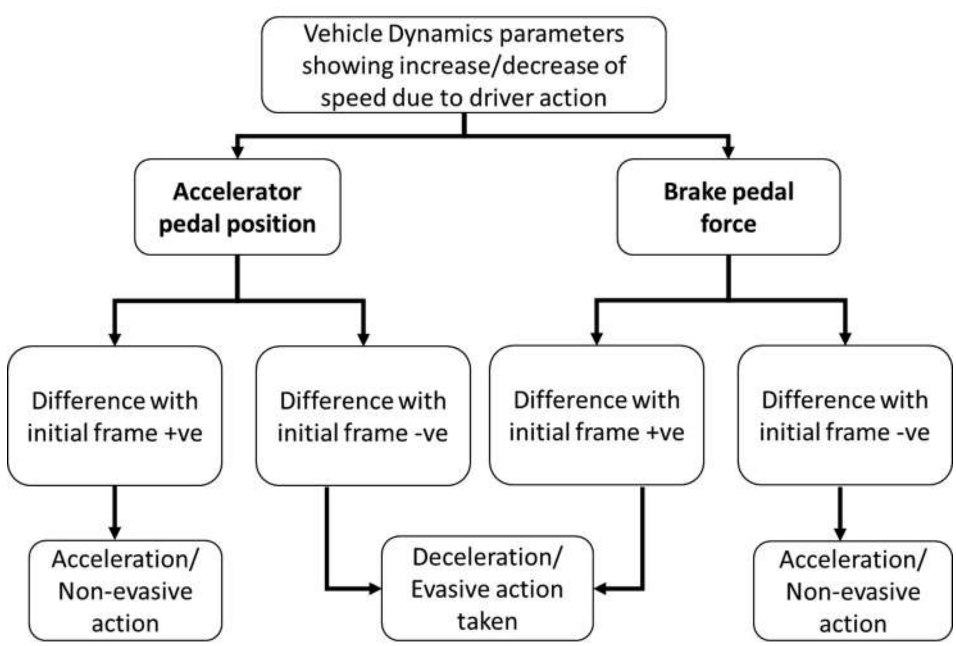

Fig. 4. Determination of evasive action using vehicle dynamics data.

following vehicle to stop behind the lead vehicle. DRAC is calculated using the following equation:

$$DRAC(t) = \frac{(V_F(t) - V_L(t))^2}{2S(t)}, \; V_F(t) > V_L(t) \quad (1)$$

where $V_F(t)$ = the following vehicle's speed at time t, $V_L(t)$ = the lead vehicle's speed at time t, and $S(t)$ = front-to-rear spacing between the following and lead vehicles at time t. DRAC can be determined only when the following vehicle's speed is higher than the lead vehicle's speed. Higher value of DRAC indicates higher collision risk since it is more difficult to apply higher deceleration to avoid collision.

3.3. Collision risk prediction models

To identify the variables and capture their effect on collision risk, collision risk prediction models were developed using two types of the models – (1) Generalized Linear Model (GLM) and (2) the random effects model. These models can also be used to predict collision risk based on the explanatory variables.

In this study, the GLM relates the mean DRAC (i.e., average of all values of DRAC in different time frames) to various explanatory variables to identify the isolated effects of the variables on collision risk. The GLM is expressed as shown in the following equation:

$$DRAC = \exp(\alpha + \beta_i x_i) \quad (2)$$

where $DRAC$ = mean DRAC, α = constant, β_i = coefficient, and x_i = explanatory variables. Since the speed difference between lead and following vehicles and the spacing between them were used to calculate DRAC, the speed difference and spacing were excluded from the models.

Unlike the GLM, which assumes a fixed value of the intercept (α), the random effects model assumes that the intercept varies among different observations. The model assumes that the intercept is normally distributed with a mean intercept α and standard deviation τ. In this study, it was assumed that the variability of intercept is due to random effect of the evasive action time on mean DRAC among different drivers.

4. Results and discussion

4.1. Comparison of evasive action time and DRAC

DRAC was separately calculated during the evasive action time and non-evasive action time (i.e., the time period during which the drivers neither applied brake nor released accelerator pedal). Table 3 shows mean values of DRAC for all drivers during the evasive action time, the non-evasive action time, and the total time. The table also compares mean DRAC among different sections. It was found that mean DRAC was highest in Section 3 followed by Section 5. This shows that rear-end collision risk was higher when the drivers approached the stopped trucks or the merging lane. Also, mean values of DRAC were consistently higher for the Trucks scenario than the Cars scenario for all sections. In particular, the difference in mean DRAC between the Cars and Trucks scenarios was statistically significant for Sections 3 and 4 at a 95% confidence interval ($p < 0.05$).

It was also found that mean DRAC was consistently higher during the evasive action time than the non-evasive action time for all sections. In particular, mean DRAC was significantly different between the evasive and non-evasive action times for Sections 3–7 for the Cars scenario and Sections 3, 6 and 7 for the Trucks scenario at a 95% confidence interval ($p < 0.05$).

However, the comparison of mean DRAC only shows which section has a higher collision risk than the others, but it cannot identify the effects of evasive action time and the other factors on mean DRAC. Thus, the results of GLM and random effect model were discussed in the next section.

4.2. GLM for all road sections

The estimated parameters of GLM for all road sections in the Cars and Trucks scenarios are shown in Table 4. All variables were statistically significant at a 95% confidence interval. It was found that there were some differences in the results between the Cars and Trucks scenarios. First, female driver has a negative effect on mean DRAC in the Cars scenario. This indicates that female drivers were generally more cautious while driving and their collision risk was lower than male drivers in when they followed cars. However, the driver's gender was not significant in the Trucks scenario. This is potentially because both male and female drivers were cautious when they followed trucks.

It was also found that driving experience has a negative effect on mean DRAC in the Trucks scenario. This indicates that more experienced drivers had lower collision risk than less experienced drivers when they followed trucks. This may be because more driving experience and better driving skill are required to avoid a collision with trucks than cars.

As expected, it was found that shorter evasive action time increased collision risk in the Cars scenario. However, the effect of the evasive action time was not significant in the Trucks sce-

Table 4
Estimated parameters of GLMs for all road sections Cars scenario.

(a) Cars scenario

Parameter	Estimates	p-value
Intercept	−2.10	0.0018
Female	−0.15	0.0497
Average speed (km/h)	0.06	<0.0001
Evasive action time	−0.03	0.0298
Section 1	−7.75	<0.0001
Section 2	−7.49	<0.0001
Section 3*	–	–
Section 4	−4.59	<0.0001
Section 5	−6.14	<0.0001
Section 6	−5.89	<0.0001
Section 7	−6.49	<0.0001
R-square	0.77	

(b) Trucks scenario

Parameter	Estimates	p-value
Intercept	−3.40	<0.0001
Average speed (km/h)	0.06	<0.0001
Driving experience (years)	−0.02	0.0003
Section 1	−5.56	<0.0001
Section 2	−6.01	<0.0001
Section 3*	–	–
Section 4	−4.65	<0.0001
Section 5	−3.41	<0.0001
Section 6	−5.74	<0.0001
Section 7	−5.76	<0.0001
R-square	0.69	

*Section 3 is the base case and the other sections are compared to Section 3.

Table 3
Mean DRAC by road section in different time periods.

	Scenario	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Evasive action time	Car	0.04	0.09	0.47	0.07	0.19	0.06	0.04
	Truck	0.08	0.12	0.76	0.14	0.42	0.09	0.06
Non-evasive action time	Car	0.04	0.08	0.34	0.07	0.17	0.06	0.03
	Truck	0.07	0.12	0.58	0.14	0.24	0.09	0.05
Total time	Car	0.04	0.09	0.42	0.07	0.18	0.06	0.04
	Truck	0.07	0.12	0.73	0.14	0.36	0.09	0.06

nario. This is potentially because the drivers perceived that it is easier for them to avoid a collision with the lead car relative to the lead truck and their evasive action time is more likely to vary among different drivers. On the other hand, it is more difficult to avoid a collision with the lead truck relative to the lead car due to the limited sight and larger difference in speed. Thus, most drivers must have taken precautions when they followed trucks and, hence, the variation in the evasive action time is likely to be lower in the Trucks scenario.

Similar to the comparison of mean DRAC, the model result shows that DRAC was highest for Section 3 among the seven road sections. This shows that approaching the stopped trucks was the most critical condition that increases chances of collision.

4.3. GLM for individual road sections

To compare effects of variables on mean DRAC among different road sections, GLMs were also developed for each road section separately. The estimated parameters of the models are shown in Table 5. The models only include the variables that are statistically significant at a 95% confidence interval. The table shows that significant variables associated with collision risk were different for different road sections, similar to Meng and Weng (2011).

Driver's gender was significant for Section 3 in the Cars scenario and Sections 5 and 7 in the Trucks scenario. This shows that female drivers generally had lower collision risk than male drivers when they approached the stopped trucks or the merging lane. Driving experience also had similar effects as driver's gender. More experienced drivers had lower collision risk than less experienced drivers in the same road section. However, in spite of significance of age in the models, the effect of age on collision risk could not be analyzed due to relatively small number of drivers over 35 and wide range of age in this age group compared to the drivers under 35. Thus, these results reflect that female and more experienced drivers are more likely to be cautious or skillful in controlling their speed in more complex traffic conditions.

It is worth noting that the effects of gender and driving experience were significant for Section 3 in the Cars scenario, whereas they were significant for Section 5 in the Truck scenario. This indicates that female and more experienced drivers perceived that it was more critical to avoid a merging truck near the on-ramp than a stopped truck and took evasive action earlier before reaching the merge area. Thus, drivers' evasive action not only depends on traffic conditions, but also the type of conflicting vehicles (cars or trucks).

It is also worth noting that the evasive action time was only significant for Sections 2, 5, and 7 in the Cars scenario and Sections 3 and 5 in the Trucks scenario. When the drivers see the trucks ahead in Section 2 (although they cannot see the trucks were stopped) and Section 3, they are more likely to reduce speed to check if they can avoid approaching trucks and stop behind the stopped trucks. Similarly, when the drivers see the merging vehicle from the on-ramp in Section 5, they are more likely to reduce speed to yield to the merging vehicle or change to inner lanes to avoid conflicts with the merging vehicle. Lastly, when the drivers are required to exit from the freeway in Section 7, they are more likely to reduce speed in the right lane while approaching the exit ramp. This result shows that longer evasive action time can help reduce collision risk more effectively in the road sections where drivers are required to take action to avoid crash or complete the task.

Table 5
Estimated parameters of GLMs for each road section.

(a) Cars scenario

Parameter	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Constant	−17.80 (0.02)	−8.78 (0.09)	−7.67 (0.03)	−10.77 (<0.001)	−3.87 (0.09)	−10.12 (<0.001)	−4.76 (0.06)
Female	–*	–	−0.33 (0.02)	–	–	–	–
Driving experience (years)	–	–	−0.02 (0.03)	–	–	–	–
Age 1 (25 or younger)	–	–	−0.84 (0.01)	–	–	–	–
Age 2 (26–35)	–	–	−0.78 (0.01)	–	–	–	–
Avg. speed (km/h)	0.12 (0.04)	0.10 (<0.001)	0.18 (0.03)	0.08 (<0.001)	0.03 (0.003)	0.07 (<0.001)	0.04 (0.001)
Evasive action time	–	−1.28 (0.02)	–	–	−0.19 (0.02)	–	−0.15 (0.001)
R-square	0.66	0.60	0.33	0.53	0.76	0.36	0.47

(b) Trucks scenario

Parameter	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Constant	–	−15.27 (<0.001)	−2.75 (0.25)	−8.65 (<0.001)	5.38 (<0.001)	−9.83 (<0.001)	−16.55 (<0.001)
Female	–	–	–	–	−0.56 (0.04)	–	1.69 (0.004)
Driving experience (years)	–	–	–	–	−0.10 (<0.001)	–	–
Age 1 (25 or younger)	–	–	–	–	−1.57 (0.008)	–	–
Age 2 (26–35)	–	–	–	–	−1.85 (0.003)	–	–
Avg. speed (km/h)	–	0.10 (<0.001)	0.12 (0.001)	0.06 (<0.001)	–	0.06 (<0.001)	0.12 (<0.001)
Evasive action time	–	–	−0.12 (0.003)	–	−0.32 (<0.001)	–	–
R-square	–	0.70	0.46	0.79	0.63	0.61	0.41

Note: The values in parentheses are p-values.
*Not statistically significant at a 95% confidence interval.

4.4. Random effects model

The estimated parameters of random effects models for all road sections are shown in Table 6. Similar to GLMs, the results of random effects models show that the effects of driver's gender and driving experience on mean DRAC were significant for the Cars and Trucks scenarios, respectively. Also, the effect of average speed was significant and collision risk was highest in Section 3 for both Cars and Trucks scenarios.

To investigate random effect of the evasive action time, the evasive action time was included in the models as a random parameter. The significance of random effect of the evasive action time is indicated by higher R-square value for the models with the random parameter and the model without the random parameter (0.96 vs. 0.64 for the Cars scenario and 0.96 vs 0.35 for the Trucks scenario).

In particular, the random effect of the evasive action time was more significant in the Trucks scenario than the Cars scenario since an increase in R-square value was larger (0.61 vs. 0.32). This implies that the pattern of evasive action was more variable among different drivers for the same evasive action time and, consequently, their effect on collision risk was also more variable when the drivers followed lead trucks than lead cars. This is potentially because due to a larger size of lead trucks, the drivers had more difficulty with viewing the road ahead and taking appropriate evasive action. As a result, drivers are more likely to take different evasive actions.

Table 6
Estimated parameters of the random effect models for all road sections.

(a) Cars scenario		
Fixed effect parameter	Estimates	p-value
Intercept	0.19	0.0018
Female	−0.04	0.0497
Average speed (km/h)	0.01	<0.0001
Section 1	−0.69	<0.001
Section 2	−0.73	<0.001
Section 3*	–	–
Section 4	−0.62	<0.001
Section 5	−0.59	<0.001
Section 6	−0.66	<0.001
Section 7	−0.67	<0.001
Random effect parameter	Variance	Standard deviation
Evasive action time	0.01	0.09
R-square (without random effect parameter)	0.64	
R-square (with random effect parameter)	0.96	
Increase in R-square	0.32	
(b) Trucks scenario		
Fixed effect parameter	Estimates	p-value
Intercept	0.19	0.0018
Female	−0.04	0.0497
Average speed (km/h)	0.01	<0.0001
Section 1	−0.69	<0.001
Section 2	−0.73	<0.001
Section 3*	–	–
Section 4	−0.62	<0.001
Section 5	−0.59	<0.001
Section 6	−0.66	<0.001
Section 7	−0.67	<0.001
Random effect parameter	Variance	Standard deviation
Evasive action time	0.01	0.09
R-square (without random effect parameter)	0.64	
R-square (with random effect parameter)	0.96	
Increase in R-square	0.32	

*Section 3 is the base case and the other sections are compared to Section 3.

Relatively less significant random effect of the evasive action time in the Cars scenario indicates that there was less variation in the effect of the evasive action time on collision risk among different drivers. Hence, the pattern of evasive action was more uniform when the drivers followed cars than trucks.

5. Conclusions and recommendations

This study analyzes the effects of the driver's evasive action time on rear-end collision risk using a surrogate safety measure (SSM) – Deceleration Rate to Avoid Crash (DRAC). For the analysis, 50 drivers' behavior on a two-lane one-way freeway was observed in a driving simulator experiment. The drivers tested two traffic scenarios - Cars and Trucks scenarios where conflicting vehicles were cars and trucks, respectively.

The presence and time of drivers' evasive action were determined based on two vehicle dynamics variables - the accelerator pedal position and the brake pedal force. Mean DRAC of all drivers was separately calculated for each scenario and each of 7 road sections which were classified based on driving conditions. Then significant factors affecting mean DRAC and their relationship with collision risk were identified using generalized linear models (GLMs) and random effect models. There are a few noteworthy findings in this study as follows.

First, DRAC decreases as the evasive action time increases. This shows that if the driver takes evasive action for a longer time period, rear-end collision risk is reduced. In particular, longer evasive action time can significantly reduce collision risk in the conditions where drivers are required to take action to avoid a collision. Second, the effects of driver characteristics (gender and driving experience) on collision risk were particularly significant in the road sections where the drivers approached the stopped trucks and merging vehicles from an on-ramp. This implies that driver characteristics are more closely related to effective evasive action to reduce collision risk in complex driving conditions. Third, the effects of variables on rear-end collision risk were different between the Cars and Trucks scenarios. This indicates that the type of lead vehicle has differential effects on collision risk in a given driving condition. Lastly, the evasive action time had a more significant random effect in the Trucks scenario than the Cars scenario. This is potentially because a large size of trucks hinders the drivers' sight and delays their evasive action, which results in higher variation in evasive action among different drivers.

In summary, these findings help better understand how drivers' characteristics and their evasive action are associated with rear-end collision risk in various driving conditions. Based on this understanding, the evasive action that effectively reduces the frequency of high deceleration in a given driving condition can be identified. Then, warning messages can be provided via variable message signs to alert drivers to take more specific evasive action in a critical traffic event. For instance, a variable message sign can be installed upstream of a freeway merge area and it can display the warning information that guides drivers to reduce the speed to a specific safe speed in the next few seconds when high ramp traffic volume is detected. More specific warning messages are likely to reduce the variability of the driver's evasive action and the speed variations among different drivers compared to generic warning messages. Less speed variations will lower the chances of crash occurrence (Hamzeie et al., 2017).

However, there were some limitations in this study. For instance, SSMs could not be validated using the crash data due to a small number of crashes during the driving simulator experiment. Also, as the driver's evasive action was analyzed only in a limited number of driving conditions, more general effects of evasive action on collision risk could not be investigated.

In future work, it is recommended that the effects of the driver's evasive action on collision risk be analyzed more extensively in different vehicle types, geometric, traffic, and weather conditions using vehicle dynamics and trajectories. It is also recommended that more effective countermeasures be developed to reduce collision risk and their safety effects be evaluated using SSMs.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by Natural Sciences and Engineering Research Council of Canada [Grant number: RGPIN-2019-04430]. This research is also supported through research infrastructure funded by Canada Foundation for Innovation [Project number: 30508].

References

Allen, B. L., Shin, B. T., & Cooper, P. J. (1978). Analysis of traffic conflicts and collisions. *Transportation Research Record, 667*, 67–74.
Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., ... Wills, R. F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis & Prevention, 42*, 788–796.
Cooper, D. F., & Ferguson, N. (1976). Traffic studies at t-junctions – A conflict simulation model. *Traffic Engineering & Control, 17*(7), 306–309.
Gold, C., Happee, R., & Bengler, K. (2018). Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis & Prevention, 116*, 3–13.
Hayward, J. C. (1972). Near-miss determination through use of a scale of danger. *Highway Research Record, 384*, 24–34.
Hamzeie, R., Savolainen, P. T., & Gates, T. J. (2017). Driver speed selection and crash risk: Insights from the naturalistic driving study. *Journal of Safety Research, 63*, 187–194.
Hyden, C. (1987). *Development of a method for traffic safety evaluation: The Swedish traffic conflicts technique*. Lund Institute of Technology.
Kuang, Y., Qu, X., & Wang, S. (2015). A tree-structured crash surrogate measure for freeways. *Accident Analysis & Prevention, 77*, 137–148.
Laureshyn, A. (2017). *Surrogate measures in traffic safety analysis*. Sweden: Lund University.
Laureshyn, A., de Goede, M., Saunier, N., & Fyhri, A. (2017). Cross-comparison of three surrogate safety methods to diagnosecyclist safety problems at intersections in Norway. *Accident Analysis & Prevention, 105*, 11–20.
Levulis, S. J., DeLucia, P. R., & Jupe, J. (2015). Effects of oncoming vehicle size on overtaking judgments. *Accident Analysis & Prevention, 82*, 163–170.
Li, X., Rakotonirainy, A., & Yan, X. (2019). How do drivers avoid collisions? A driving simulator-based study. *Journal of Safety Research, 70*, 89–96.
Li, Y., Xu, C., Xing, L., & Wang, W. (2017). Evaluation of impacts of different car-following types on rear-end crash risk at freeway weaving sections using vehicle trajectory data. *Presented at the 96th Transportation Research Board Annual Meeting, Washington, D.C.*.
Mahmud, S. M. S., Ferreira, L., Hoque, M. S., & Tavassoli, A. (2017). Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. *International Association of Traffic and Safety Sciences (IATSS) Research, 41*, 153–163.
Meng, Q., & Weng, J. (2011). Evaluation of rear-end crash risk at work zone using work zone traffic data. *Accident Analysis and Prevention, 43*, 1291–1300.
Triggs, T., & Harris, W. (1982). *Reaction time of drivers to road stimuli. Technical report*. Monash University, Human Factor Group, Department of Psychology.
Wang, C., & Stamatiadis, N. (2013). Surrogate safety measure for simulation-based conflict study. *Transportation Research Record: Journal of the Transportation Research Board, 2386*, 72–80.
Wang, X., Zhu, M., Chen, M., & Tremont, P. (2016). Drivers' rear end collision avoidance behaviors under different levels of situational urgency. *Transportation Research Part C, 71*, 419–433.
Wu, Y., Abdel-Aty, M., Cai, Q., Lee, J., & Park, J. (2018). Developing an algorithm to assess the rear-end collision risk under fog conditions using real-time data. *Transportation Research Part C, 87*, 11–25.
Zhao, P., & Lee, C. (2017). Analysis and validation of surrogate safety measures by types of lead and following vehicles. *Transportation Research Record: Journal of the Transportation Research Board, 2659*, 137–147.
Zheng, Z., Lu, P., & Lantz, B. (2018). Commercial truck crash injury severity analysis using gradient boosting data mining model. *Journal of Safety Research, 65*, 115–124.

Dhwani Shah is a Ph.D. student at the University of Windsor, Canada. Her research focuses on analysis of driver behavior and collision risk using driving simulator data. Prior to her doctoral studies, Ms. Shah received Bachelor's and Master's degrees from CEPT University, India.

Chris Lee, Ph.D., P.Eng. is an Associate Professor of Civil and Environmental Engineering at the University of Windsor, Canada. His main research interests are in traffic safety, driver behavior, and traffic operation and control. Dr. Lee has conducted research on vehicle interaction and conflicts in car-truck mixed traffic, car-following and lane-changing behaviour, prediction of crash frequency and injury severity, safety evaluation of countermeasures, and advanced real-time speed control for freeways. Dr. Lee is an Associate Editor of two academic journals - Accident Analysis and Prevention and Canadian Journal of Civil Engineering.

Journal of Safety Research 78 (2021) 19–27

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Application of machine learning technique for optimizing roadside design to decrease barrier crash costs, a quantile regression model approach

Mahdi Rezapour *, Khaled Ksaibati

Wyoming Technology Transfer Center, 1000 E University Ave, Dept. 3295, Laramie, WY 82071, United States

ARTICLE INFO

Article history:
Received 13 December 2019
Received in revised form 16 June 2020
Accepted 2 June 2021
Available online 15 June 2021

Keywords:
Machine learning
Quantile regression model
Traffic barrier crash severity
Optimization
Benefit cost analysis

ABSTRACT

Introduction: In-transport vehicles often leave the travel lane and encroach onto natural objects on the roadsides. These types of crashes are called run-off the road crashes (ROR). Such crashes accounts for a significant proportion of fatalities and severe crashes. Roadside barrier installation would be warranted if they could reduce the severity of these types of crashes. However, roadside barriers still account for a significant proportion of severe crashes in Wyoming. The impact of the crash severity would be higher if barriers are poorly designed, which could result in override or underride barrier crashes. Several studies have been conducted to identify optimum values of barrier height. However, limited studies have investigated the monetary benefit associated with adjusting the barrier heights to the optimal values. In addition, few studies have been conducted to model barrier crash cost. This is because the crash cost is a heavily skewed distribution, and well-known distributions such as linear or poison models are incapable of capturing the distribution. A semi-parametric distribution such as asymmetric Laplace distribution can be used to account for this type of sparse distribution. *Method:* Interaction between different predictors were considered in the analysis. Also, to account for exposure effects across various barriers, barrier lengths and traffic volumes were incorporated in the models. This study is conducted by using a novel machine-learning-based cost-benefit optimization to provide an efficient guideline for decision makers. This method was used for predicting barrier crash costs without barrier enhancement. Subsequently the benefit was obtained by optimizing traffic barrier height and recalculating the benefit and cost. The trained model was used for crash cost prediction on barriers with and without crashes. *Results:* The results of optimization clearly demonstrated the benefit of optimizing the heights of road barriers around the state. *Practical Applications:* The findings can be utilized by the Wyoming Department of Transportation (WYDOT) to determine the heights of which barriers should be optimized first. Other states can follow the procedure described in this paper to upgrade their roadside barriers.

1. Introduction

Every year, more than a million people die due to traffic crashes worldwide. In addition, 50 millions more are severely injured on roadways (OECD. Publishing, 2017). In 2017 alone, more than 37,000 people died in the United States as a result of road crashes (National Highway Traffic Safety Administration, 2017). Hitting a fixed object account for a significant proportion of these fatalities: about 8,000 people died in 2016 alone in collision with fixed objects, which was 3% higher than the 2015 crashes (FARS & National Highway Traffic Safety Administration, 2016). Roadway departure crashes tend to be hazardous, especially in a mountainous areas like Wyoming with challenging roadways geometric characteristics (Rezapour, Wulff, & Ksaibati, 2019). However, various mitigating strategies are available that could be implemented to reduce the severity of hitting a fixed object on the roadsides. These methods include transferring the hazards out of the clear zones or a more practical approach, which is interposing traffic barriers. Still, the traffic barrier itself is hazardous and its installation would be recommended only if it could redirect drivers from more hazardous objects.

When the installation of traffic barriers is warranted, to minimize the severity of these crashes special attention should be paid to the geometric characteristics of traffic barriers to prevent underride or override crashes. Extensive research has been conducted with the help of simulated or field crash tests to identify an

* Corresponding author.
E-mail address: mrezapou@uwyo.edu (M. Rezapour).

https://doi.org/10.1016/j.jsr.2021.06.002

optimum height to prevent the risks of override or underride crashes. Short-height barriers poses a risk of override vehicle crashes. For instance, barrier heights between 24 and 26 inches pose a possibility of an increased risk of vehicle override crashes (Wiebelhaus, Lechtenberg, Sicking, Faller, & Rosenbaugh, 2013). On the other hand, a barrier with a height of 36 inches, would increase the risk of underride crashes (Julin, Asadollahi Pajouh, Stolle, Reid, & Faller, 2017). However, the impacts of these measures (e.g., various barrier heights) have not been quantified by the majority of past studies. As no study has been conducted on barrier optimization based on crash cost or severity, a few studies would be welcome on the application of various optimization techniques on transportation problems and other issues.

A simultaneous consideration of traffic safety risks and the cost burden related to the appropriate planning and design was conducted in a past study (Li, Ding, & Zhong, 2019). The optimization results showed that the method is capable of generating high-quality solution.

In another study, a model for highway alignment optimization that integrates GIS with genetic algorithm was presented (Jha & Schonfeld, 2004). The objective of this study was to examine the effects of costs on alignment selection. The results indicated that travel-time cost is one of the factors that significantly impacts the alignment optimization.

An innovative wrong-way driving (WWD) countermeasure optimization was discussed to help decision makers identify the optimal deployment locations based on available resources (Sandt & Al-Deek, 2019). The WWD model used non-crash events, interchange design, and traffic volumes to predict the frequency of crashes. It was found that, based on some of the scenarios, the decision makers could expect more than 30% of WWD crash risk reduction by equipping some of the ramps.

Several studies have been conducted on application of machine learning technique for optimizing non-transportation problems. For instance, machine learning techniques were used for optimizing solar water heater performance (Li et al., 2017).

In addition, studies have been conducted to investigate the impact of barrier geometric on the crash severity (Rezapour & Ksaibati, 2018). Despite the efforts, the impacts of various barrier height on the barrier crash cost is still missing in the literature. What is the estimated cost of barrier crashes based on barrier heights? Which barriers should be upgraded first to secure the highest benefit based on limited available budget? And how much money would be saved by optimizing the traffic barriers geometric to its optimal value to maximize benefits over a few years? These are just some of the questions that are of crucial importance for decision makers to be answered. One of the challenges of addressing the above questions is that measuring the traffic barriers geometric characteristics are costly. Also, the past study highlighted the importance of incorporating traffic geometric characteristics in the safety analysis. In addition, the past study has shown that the impact of traffic barrier height on barrier severity should be evaluated in combination with the impact of shoulder width (Rezapour et al., 2019; Rezapour & Ksaibati, 2020a; 2020b).

The Wyoming Department of Transportation (WYDOT) collected information on 1.3 million of linear feet of traffic barriers on the state and interstate systems in Wyoming. The information included measurements of traffic barrier height, types of barrier, and shoulder width.

Thus, this study took advantage of the provided information with the help of machine learning techniques to model traffic barriers crash cost as response before and after optimizing barriers heights. The trained model was implemented on barriers with and without crashes to determine the cost effectiveness of optimizing barriers heights.

In summary, the Contributions of this study are as follows:

1. WYDOT has a limited budget to upgrade barriers so it is important to prioritize barriers optimization. Thus, the barriers would be prioritized to determine which ones are most cost effective to address first, based on budget availability.
2. Since a crash is a random/rare event, an optimization is needed not only on barriers with historical crashes but also on barriers that did not experience any crash. That is especially important as there is a significant proportion of barriers with no crashes in Wyoming due to low traffic volumes and randomness of crashes. Thus, these barriers need to be optimized if they are outdated or outside the recommended heights.
3. Although barrier crash cost is dependent on various factors, this study only considers predictors that are common across both barriers with and without crashes so a trained model could be implemented on barriers with no crashes.
4. Interaction terms were considered in this study to account for the heterogeneity across crashes by allowing the influence of predictors to vary by crashes.

2. Data

In order to optimize the performance of traffic barriers, they should be up-to-date based on a recommended height. Since updating all barriers in one shot is cost prohibitive, selecting which barriers should be upgraded first is really essential. The following section summarizes the challenges for the optimization process.

First, upgrading all the barriers would be very costly and decision makers often have limited budget. Secondly based on the literature review, barrier crash severity depends on many factors such as human behavior, roadway conditions, and roadway and barrier designs. But sometimes barriers do not receive any crashes due to their scarcity and randomness. Consequently, there would be no response (e.g., crash costs) or driver behavior predictors for barriers without crashes. This makes it hard for conducting any statistical analysis. This situation would be much direr if the barriers with no crashes are outdated and they need immediate attention. Third, for optimization process, the cost of crash should be known after modifying barrier heights based on predicting models.

Thus, in order to have a model that could be implemented on a barrier with no crashes, both datasets (barriers with and without crashes) should incorporate similar predictors. So in the modeling portion, we incorporated only variables that were available for the two datasets such as barrier geometric characteristics and shoulder width. Also to be unbiased across barriers with different traffic volumes and lengths, those exposure parameters were incorporated in the model. To account for the structure of the dataset resulting from incorporating different barrier types, the interaction between barrier types and significant predictors were considered. A machine learning technique was implemented to address the issue of finding predicted crash costs with and without enhancement for both barriers with and with no crashes. Table 1 presents those predictors that were found to be significant in the final analysis.

It should be noted that the number of barriers experienced crashes in the state highway system in Wyoming is higher than those barriers without crashes (836 vs. 374). This difference is related to the fact that only barriers outside the recommended heights were incorporated in the optimization process, and mostly state highway barriers were outdated. Table 1 highlights the differences across these two groups of barriers. As can be seen from Table 1, while the average of barrier heights experienced crashes is 29.2 inches, this value is significantly lower for barriers with no crashes (22.7 inches). This highlights the importance of studying the barriers without historical crashes.

It is interesting to see although the second barrier group did not experience any crash, the average annual daily traffic (AADT) is very close to the first group. Moreover, the posted speed limit is

Table 1
Descriptive analysis of barriers with and without crashes.

Variables	Mean	St.dev	Max	Min
Barriers with crashes (number = 836, length = 138 miles)				
Real crash cost in dollar per barrier	328,895	1,449,023	19,244,066	34,612
Barrier height (in)	29.2	3.3	40.8	<12
Barrier length (ft)	874	1816.78	35,47	14
Traffic (AADT)	1,313	1,313	8,854	27
Side slope, 0 if it is flat,1 otherwise	0.75	0.431	1	0
Shoulder width	<=5.5′	538	——	——
	>5.5′	298	——	——
Posted speed limit	54.89	15.310	70	20
Barrier types	Box beam	386	——	——
	W-beam	450	——	——
Barriers with no crash (number = 374, length = 29miles)				
Crash cost in dollar per barrier	0	0	0	0
Barrier height (in)	22.7	2.828	25.2	<12
Barrier length (ft)	402	412.903	2500	14
Traffic (AADT)	1,177	1191	5878	27
Side slope, 0 if it is flat,1 otherwise	0.87	0.328	1	0
Shoulder width	<=5.5′	267	——	——
	>5.5′	107	——	——
Posted speed limit	64.170	10.847	70	20
Barrier types	**Box beam**	123	——	——
	W-beam	251	——	——

higher for barriers with no crash compared to barriers with crashes. For this study, W-beam accounts for a higher number of barriers for both barrier datasets. Shoulder width converted to a binary predictor as it was found that a cutting point of 5.5 feet is important for prediction of crash cost. In addition, this value splits shoulder width into almost two equal categories.

For barriers with crash dataset, the crashes between 2007 and 2016 were used in the analysis. Data were filtered to include only single- vehicle barrier crashes. Cable barriers and concrete barriers crashes were excluded from the dataset due their low number of frequency and crashes. As discussed earlier, only barriers at a critical conditions, above 35 or below 27 inches, were included in the analysis.

3. Method

The distributions (such as Laplace and Legendre) suggest that the minimization of absolute deviation is preferable to least square when some sample observation are poorly distributed (Koenker & Bassett Jr, 1978). Also, Laplace distribution has proven that in the simple bivariate regression, the least absolute error estimator has a smaller asymptotic variance than the least square estimator if the model error law has variance, σ^2, and density at the median, $f(0)$, satisfying $[2f(0)]^{-1} < \sigma$. So this distribution with a long tail could be a good option for modeling sparse crash cost distribution. Fig. 1 highlights the sparsity of crash cost aggregated over various barriers.

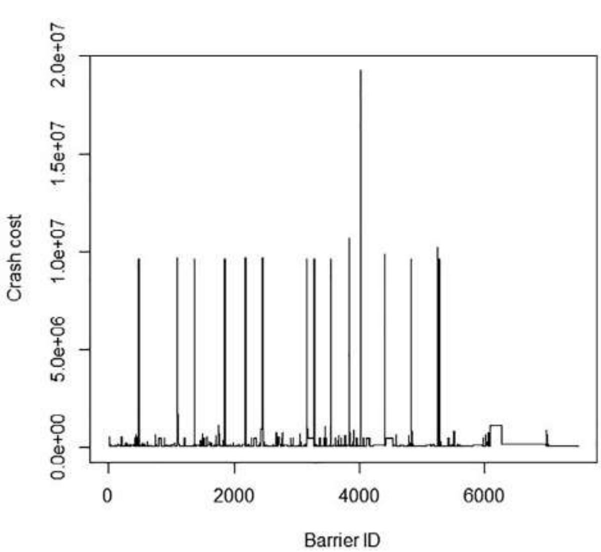

Fig. 1. Sparsity of crash cost versus barrier ID.

Linear quantile regression model has been proposed to model the poorly distributed performance model (Koenker & Hallock, 2001). This model distribution is based on Laplace distribution, which has often been used for modeling sparse dataset such as cost data. The interpretation of quantile regression is very similar to ordinary least square (OLS) with the difference that instead of predicting the mean of the dependent variable, this model looks at the quantile of the dependent variable.

In contrast with least square estimation, where the objective is to minimize the sum of squared residuals, the objective function for quantile regression model is to minimize sum of absolute deviation of residuals. The minimization objective function for θth regression quantile, $0<\theta<1$, could be written as any solution to the below equation (Koenker & Bassett Jr, 1978):

$$b \int^{\min} \mathbb{R} \left[\sum_{t \in \{t: y \geqslant x_t b} \theta |y_t - x_t b| + \sum_{t \in \{t: y < x_t b} (1-\theta)|y_t - x_t b| \right] \tag{1}$$

where b is the linear function of included variables.

The analysis of this study conducted in R (Geraci, 2014). The response is crash cost, which is calculated based on the following equation:

$$\begin{aligned} y_i = {} & 4612 \times PDO_i + 9604727 \times Fatal_i + 132181 \\ & \times Suspected\ serious\ injury_i + 132181 \\ & \times Suspected\ \mathrm{minor}\ injury_i + 149551 \times Unknown_i \end{aligned} \tag{2}$$

The response, crash cost, was highly sparse (see Fig. 1), which could not be modeled by any known distribution such as Gaussian, or gamma distribution. It should be noted that various extreme transformations were conducted to determine if Gaussian or

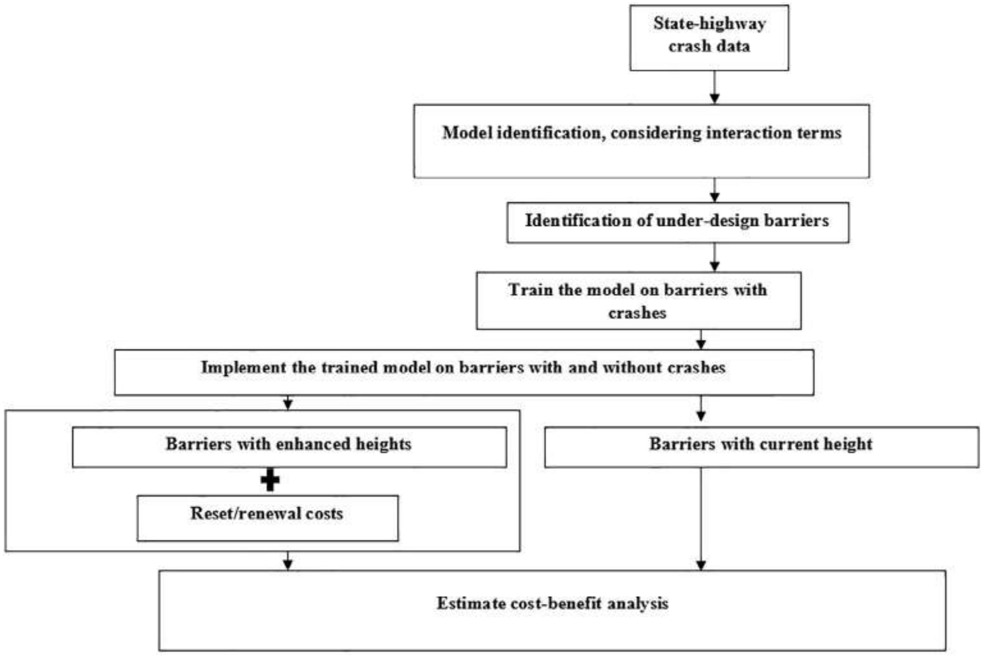

Fig. 2. Methodological steps for performing the analysis.

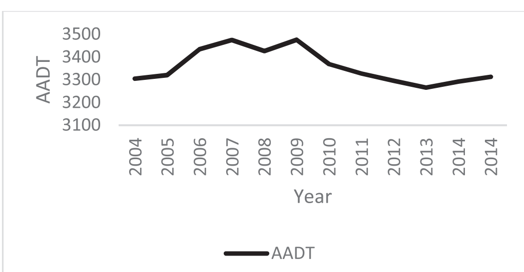

Fig. 3. Distribution of Average annual daily traffic over the last 10 years on Wyoming state highway system.

Gamma distribution could be applied. However, even those transformations did not result in an improvement of the residual distribution. For optimization process, the steps in Fig. 2 were performed. As mentioned earlier, one of the main concerns of WYDOT is to optimize barriers without, in addition to those with crashes, so after a model was trained on barriers with crashes, it was implemented on barriers with no crash (Fig. 3).

All the highway barriers with crashes were used for training the model. Standard quantile regression model does not account for the hierarchy of the dataset structure, so in order to account for the structure of datasets resulted from different barriers types, w-beam versus box beam, interaction terms of the barrier types, and various important predictors were incorporated in the model. This would be accomplished by letting the crash cost vary by changing the barrier types along with other predictors. Only predictors that were available for the barriers with no crashes were incorporated in the trained model so the trained model could be implemented on barriers with no crashes.

Various predictors were identified as important, such as barrier height and shoulder width. Although a model trained across all barriers, regardless of whether they were within the recommended height or not, the trained model was implemented on the dataset being filtered to incorporate only the over-height or under-height barriers based on the dimensions: the barriers above 35″or below 27″ were incorporated in the optimization process only (Transportation Officials. Task Force for Roadside Safety, 2011) as these barriers are more likely to result in override and underride barrier crashes.

A value of 27 inches was chosen in this study to optimize all the barriers to that value (Transportation Officials. Task Force for Roadside Safety, 2011). After predicting cost by implementing the trained model on under-design barriers with and without crashes, the barrier heights were optimized to 27 inches to predict costs based on new heights. Barriers would be reset or newly installed based on their conditions and types of their posting. Barriers with wooden posts cannot be reset but need to be replaced. As almost all W-beams in the state included wooden post, they would be replaced while the box-beam would be reset to its optimum value of 27 inches without replacement. The cost of these changes are included in Table 2. The values would be multiplied by the length of barriers in feet to come up with a total cost. This cost would be taken into account for only the first year, and would be added up to the predicted cost with enhancement only once (see equation (4)).

Like every optimization process, the optimization process consists of an objective function and constraints. The objective function of this study was to minimize the cost of barrier crashes over the next 10 years based on budget availability. The main constrain is to only optimize those barriers being outside the recommended high (below or above the threshold height). Another constrain is to keep all other predictors constant during the 10 years period, except for barrier height. Barriers would be upgraded to a fixed value highlighted as 27 inches.

3.1. Barrier ranking criteria

Total benefit is calculated from the following equation:

$$\begin{aligned} \textit{Total benefit in } 10 \textit{ years} &= \textit{benefit over the } 1st \textit{ year} + 9 \\ &\quad * \textit{benefit over consecutive years} \end{aligned} \tag{3}$$

Table 2
Cost of changing barriers in Wyoming based on bid price.

Barrier types	Brand new installation$ per foot	Reset bid price$ per foot
Box beam	50.25	10.45
W beam	50.89	–

On the other hand, the benefit over the first year would be calculated as follows:

$$\textit{Benefit over the 1st year} = \textit{predicted cost with no optimization} - (\textit{predicted cost with optimization} + \textit{reset cost}) \quad (4)$$

Based on equation (3), various barriers across two datasets would be ranked based on a higher total benefits. Reset cost would be calculated from length × values obtained in Table 2. It should be noted that although real current crash cost was available for barriers with crashes, predicted crash costs without barriers enhancement were estimated by implementing the trained model over the dataset. This is because the scale of prediction and real values are different, and current real cost cannot be used to be compared with predicted crash cost after enhancement.

As for predicted cost after enhancement, the only unconstrained predictor was traffic barrier height. The barrier heights were changed to an optimum value of 27 inches, and then the trained model was implemented on a new dataset, with a new barrier height to come up with a predicted future cost after optimization of barriers height. Since traffic volumes on the state highway system have been almost constant over the last 10 years (Fig. 2), it was assumed this value would remain constant for the next 10 years. So after calculating the cost for the first year, the benefit would be multiplied by 9 and added up with the first year to come up with the total benefit (see equation (3)). The same process is implemented on barriers with no crashes.

4. Results

The Lqm function in R was used for training the model on barriers with crashes. This model has a form of linear regression so the format of the model needs no further explanation. The significant identified predictors are presented in Table 3. As can be seen from Table 3, although this model does not account for the structures of data resulting from different types of barriers, the interaction between barrier types and important predictors were considered in the model. Also, traffic and length of barriers were incorporated in the model to account for the exposure. In addition, all interaction terms between significant predictors were considered for the model. As expected, the exposure variables (including barrier length and traffic volume) resulted in a higher barrier crash cost. The interaction between shoulder width and barrier height was also found to be important.

It was found that there is important interaction terms between barrier types and side slope, as well as barrier types and posted speed limit. These interaction terms might result from a decision of WYDOT in selection of various barrier types based on traffic and side slopes. Variance of real crash cost and predicted crash cost, and root mean square error (RMSE) of crash cost were presented to visualize the performance of the model. As the objective of this study was not identification of contributory factors but optimization, no further details would be presented for the modeling results.

As discussed earlier, to predict the crashes after enhancement, all the variables in Table 3 were kept constant and the changes were made on traffic barrier heights only. This is due to limitations of changing other parameters such as shoulder width due to roadside limitation being clarified by WYDOT.

Table 4 presents the top 25 ranked cost-effective barriers with historical crashes as identified in the optimization process. Table 5, on the other hand, shows the top 25 barriers without any historical crashes. The barriers were sorted based on the highest benefits resulting from reductions in predicted crash cost due to only changes in barrier heights. As can be seen from Table 4, the majority of cost effective barriers were W-beam barriers. This is despite the fact that these barriers have higher initial costs related to barrier height change costs (see Table 2).

Also, it can be seen from Table 4, the top nine most economical barriers are those that are above the threshold height of 35 inches, followed by very low barriers. In addition, the above threshold barriers are accompanied by wider shoulder width, while short barriers are accompanied by a shorter shoulder width. This resulted from the interaction terms identified in Table 4.

Similar to Table 5, despite the higher initial cost of W-beam, it was found that the highest improvement could be made on W-beam compared to box-beam on barriers with no crash. As expected from the identified results in Table 5, the highest impact/benefit is observed for the lowest-height barriers when they are located at lower shoulder widths. It should be noted (as discussed earlier in the manuscript) that barrier length and traffic were incorporated in the modeling results to account for barriers with different heights and various traffic. For instance, while barrier ID 109 was identified as the most economical barrier to be optimized, that barrier is among the barriers with lowest length, and with average AADT, highlighting the importance of incorporating AADT and barrier lengths in the model.

As explained earlier, the model was trained to include only variables that are common across both datasets, barriers with crashes

Table 3
Results of regression quantile model, 95% quantile.

	Value	Std. Error	Lower bound	Upper bound	Pr(>|t|)
Intercept	8.81E + 05	6.54E + 05	−4.32E + 05	2,194,828	0.183859
Barrier height	−1.93E + 05	2.47E + 05	−6.88E + 05	302582.4	0.437842
AADT	2.34E + 02	7.67E + 01	7.99E + 01	388.28	0.003672
Length of barrier	2.65E + 02	1.56E + 02	−4.89E + 01	−578.76	0.052001
Side slope	−1.32E + 05	8.25E + 04	−2.98E + 05	33666.75	0.11571
Shoulder width	−2.29E + 06	9.45E + 05	−4.19E + 06	−393499	0.01899
Barrier type	−2.30E + 05	4.55E + 04	−3.22E + 05	−138574	6.46E-06
Posted speed limit	−5.17E + 03	3.74E + 03	−1.27E + 04	2355.46	0.173718
Side slope : type of barrier	6.26E + 04	2.87E + 04	5.06E + 03	120226.2	0.033615
Barrier height: shoulder width	9.25E + 05	3.94E + 05	1.32E + 05	1,717,153	0.023088
Barrier type: Posted speed limit	3.82E + 03	1.37E + 03	1.06E + 03	6585.42	0.007661

Variance (crash real cost) = 2e + 12,variance (predicted cost with no enhancement) = 3.576208e + 11, RMSE = 1.47e + 06.

Table 4
The top critical state highway-system barriers: sorted based on highest benefit,

Barrier ID	Shoulder width	AADT	Barrier height	Type of barrier	Length of barrier (ft)	Predicted crash cost without barrier enhancement	Predicted crash cost with enhancement	barrier height change cost	Benefit in 10 years	Rank
2744	1	1536	40.8	W Beam	152	−1,123,297	−342,233	−7,712	7,802,925	1
4231	1	363	37.2	W Beam	66	−638,563	−61,255	−3,372	5,769,706	2
5324	1	657	37.2	Box Beam	723	−714,252	−136,944	−7,552	5,765,526	3
1443	1	2189	37.2	W Beam	665	−1,111,476	−534,168	-33,851	5,739,228	4
3055	1	719	36	W Beam	312	−941,238	−431,848	−15,878	5,078,015	5
3055	1	719	36	W Beam	312	−941,238	−431,848	−15,878	5,078,015	6
3056	1	719	36	W Beam	712	−1,067,924	−558,534	−36,245	5,057,648	7
3056	1	719	36	W Beam	712	−1,067,924	−558,534	−36,245	5,057,648	8
3888	1	659	36	W Beam	884	−1,074,108	−564,718	−45,005	5,048,888	9
576	0	955	<12	Box Beam	169	−667,164	−275,992	−1,761	3,909,954	10
2745	0	588	<12	Box Beam	461	−1,988,261	−1,597,088	−4,822	3,906,894	11
2745	0	588	<12	Box Beam	461	−1,988,261	−1,597,088	−4,822	3,906,894	12
4739	0	5234	18	Box Beam	169	−818,620	−688,229	−1,762	1,302,142	13
5453	0	1417	19.2	W Beam	901	−865,220	−752,215	−45,866	1,084,184	14
5196	0	687	19.2	W Beam	1062	−969,475	−856,469	−54,056	1,075,995	15
3847	0	4810	19.2	W Beam	1834	−1,160,352	−1,047,346	−93,315	1,036,735	16
3678	0	4810	20.4	Box Beam	214	−1,551,362	−1,455,742	−2,232	953,964	17
366	0	467	20.4	Box Beam	257	−685,951	−590,331	−2,686	953,511	18
796	0	75	21.6	W Beam	87	−610,229	−531,994	−4,451	777,891	19
1365	0	1264	21.6	Box Beam	492	−483,973	−405,738	−5,140	777,202	20
1859	0	739	21.6	Box Beam	1204	−620,043	−541,809	−12,586	769,757	21
3069	0	1045	21.6	Box Beam	1267	−590,369	−512,134	−13,245	769,098	22
3993	0	659	21.6	W Beam	634	−539,124	−460,890	-32,282	750,061	23
498	0	1677	21.6	W Beam	1088	−1,026,365	−948,130	−55,357	726,986	24
5453	0	1417	21.6	W Beam	1201	−675,170	−596,935	−61,116	721,226	25

All values, except for barrier height, are rounded.
All the values, except for barrier height are rounded.

Table 5
The tope critical highway-system barriers, sorted based on highest benefit, barriers with no crash.

Barrier ID	Shoulder width	AADT	Barrier height	Type of barrier	Length of barrier (ft)	Predicted cost with no enhancement	Predicted cost after enhancement	barrier height change cost	Benefit	Rank
109	1	955	<12	Box Beam	97	−584,529	−193,358	−1,016	3,910,694	1
575	1	955	<12	Box Beam	102	−650,957	−259,786	−1,068	3,910,642	2
1048	0	530	<12	W Beam	64	−423,022	−31,850	−3,245	3,908,475	3
1047	0	530	<12	W Beam	65	−423,323	-32,152	−3,302	3,908,408	4
112	1	955	<12	Box Beam	361	−718,907	−327,736	−3,774	3,907,936	5
5200	0	687	12	W Beam	2087	−1,344,888	−1,127,570	−106,212	2,066,968	6
15	0	128	13.2	W Beam	364	−522,201	-322,269	−18,504	1,980,816	7
7771	0	1176	14.4	W Beam	124	−760,922	−578,375	−6,308	1,819,162	8
5202	0	687	14.4	W Beam	201	−798,679	−616,132	−10,224	1,815,246	9
5454	0	1417	14.4	W Beam	631	−706,157	−523,611	-32,100	1,793,360	10
5198	0	687	14.4	W Beam	725	−940,739	−758,192	−36,886	1,788,584	11
4573	0	141	15.6	W Beam	201	−681,109	−515,948	−10,240	1,641,370	12
5197	0	687	15.6	W Beam	913	−974,317	−809,155	−46,451	1,605,169	13
3614	0	228	16.8	W Beam	127	−703,060	−555,284	−6,450	1,471,310	14
5201	0	687	16.8	W Beam	201	−763,834	−616,058	−10,210	1,467,550	15
5462	0	1417	16.8	W Beam	338	−710,382	−562,606	−17,222	1,460,538	16
5199	0	687	16.8	W Beam	537	−855,078	−707,302	−27,335	1,450,425	17
1205	1	1176	18	Box Beam	159	−287,212	−156,821	−1,657	1,302,253	18
852	1	739	18	Box Beam	324	−291,908	−161,518	−3,384	1,300,516	19
3615	0	228	18	W Beam	126	−685,550	−555,160	−6,427	1,297,473	20
4572	0	141	18	W Beam	202	−646,482	−516,092	−10,267	1,293,633	21
5204	0	687	18	W Beam	249	−759,662	−629,271	−12,690	1,291,220	22
21	0	128	18	W Beam	299	−671,647	−541,257	−15,210	1,288,690	23
5205	0	687	18	W Beam	300	−773,295	−642,905	−15,248	1,288,652	24
62	0	551	18	W Beam	338	−379,416	−249,026	−17,189	1,286,711	25

All values, except for barrier height, are rounded.

and barriers with no crashes. Special attention was also made to incorporate barrier length and traffic count in the model to normalize the barriers. After the model was trained on barriers with crashes, it was implemented on barriers with no crashes. Table 6 presents the summary description of optimizations process across barriers with and without crashes.

The steps for obtaining values in Table 6 were discussed in the method section. Values for number of economical barriers means considering equation (3), the benefit for that barrier ID is positive versus those barriers that changing the barrier height is not cost effective. As can be seen from Table 6, for barriers with crashes, WYDOT could save more than a predicted value of 52 million dol-

Table 6
Summary statistics of cost benefit analyses across various barriers.

Barrier types		Number of barriers	Total length in Miles	# of Economical barriers	Length of economic barriers	# of non-economic barriers	Length of non-economic - barriers	Reset/new installation cost for the 1st year	Total benefit
Barriers with crashes	W-beam	87	19	60	8	27	11	−5,165,219	37,951,532
	Box-beam	62	6.5	44	4	18	2.5	−358,664	14,676,278
	Sum	149	26	104	12.3	45	13.4	−5.523,884	52,627,809
Barriers without crashes	W-beam	251	19	207	16	43	3	−5,210,783	8,851,006
	Box-beam	123	10	59	5	64	5	−543,742	No benefit
	Sum	374	29	266	21	107	8	−5,754,526	8,851,006
Total benefit across all barriers	1046	55	370	33	152	21	−10,987,410	61,478,815	

lars. Although on average, both W-beam and box-beam were found to be cost effective, the W-beam benefits outperform box-beam benefits. This might be due to a higher length of barriers of this type on highway system and also more associated benefits as discussed earlier.

As far as barriers with no crashes, Table 6 shows that while W-beam optimization was found to be cost effective, just changing barriers height for box-beam barrier was found to be not-cost-effective. This might be due to the fact that there are other confounding factors resulting in a high predicted cost that cannot be addressed only by changing the barrier height. Another reason is that, for this category, shoulder width needs to be changed along with barrier height as the interaction between these two predictors indicated in Table 3.

In summary, by WYDOT investing more than 10 million dollars for upgrading the recommended height barriers in the state, in 10 years, and just by focusing on barrier heights change, they could expect not only to recover the invested money but gain more than 60 million dollars through reducing the barrier crash severity costs.

4.1. Change of shoulder width along with change of barrier height

As discussed in the previous section, although most of the barriers in the highway system were cost effective to be optimized solely based on their height, it was found that optimizing only the height of box-beam barriers with no crashes is not cost effective. As can be seen from Table 3, the impact of barrier height on crash cost cannot be explained just based on barrier heights, but also based on the interaction of both barrier heights and shoulder width.

It was interesting to see that while an increase in barrier height, on average, enhanced the safety of Wyoming highway systems, the combination/interaction of shoulder width and barrier height for box-beams with no crashes necessitates changing these two predictors together. In the original analysis, reset cost was considered for box beam. However, when optimization involved enhancing the shoulder width as well, changing the whole barrier needs to be considered.

The cost of widening the shoulder width is calculated based on the following formula:

$$\begin{aligned}&\mathit{Cost\ of\ shoulder\ widening\ in\ Wyoming}\\&= 100,000 \times \mathit{width\ of\ shoulder\ width\ needs\ to\ be\ enhanced}\ (\mathit{ft})\\&\quad \times \mathit{length\ of\ the\ roadway}\ (\mathit{Mile})\end{aligned} \tag{5}$$

For Wyoming highway systems, there are 59 box-beams that did not experience any crash, and they are suffering from extremely short height. These short barriers are accompanied by short shoulder width, (less than 5.5 feet). Based on the interaction terms obtained from the modeling results (Table 3), in order to optimize the short barrier height accompanied by a short shoulder width, both of these predictors need to be enhanced to obtain cost effective optimization results. Thus, after optimizing both predictors, a benefit of more than 13 million dollar could be expected (see Table 7).

4.2. Barrier optimum height

It is worth discussing the chosen value of 27 inches for an optimum value of all traffic barriers. Based on the results in Table 3, the linear programing for identifying traffic barrier height with various constrains would be as follows:

$$\begin{cases}\min : \mathit{barrier\ height} * (-22.8E+05) - 2.29E\\ \quad +06 * (\mathit{shoulder\ width}) + 111E + 05 * \mathit{barrier\ height}\\ \quad * \mathit{shoulder\ width\ objective\ function}\\ 27 \leqslant W - \mathit{beam\ height} \leqslant 31,\ \mathit{constrain}\\ \mathit{Box\ beam} = 27\mathit{inches},\ \mathit{constrain}\end{cases} \tag{6}$$

Table 7
Optimization results by enhancing shoulder width and barrier height just for box-beam barriers that did not have crashes.

Barrier types	Number/length (miles) barrier needing shoulder changes	Cost of change in shoulder width + barrier height	Predicted cost with no optimization	Predicted cost after the 1st year, for every year, after barriers enhancement	Predicted benefit for the first year after enhancement	Total saving in 10 years
Barriers	without crashes	Box-beam		−3,089,453	12,799,612	
			59/5.19			
	−11,491,920	1,781,761	13,550,989			
Sum of the benefit across all barriers in highway system = 75,029,804						

The above linear programing does not need any modeling programming skill and an optimum value can be calculated with simple algebra. Based on the above equation, when the shoulder width is at a higher value, greater than 5.5 feet (being set as 1), an optimum value of the above equation for w-beam could be achieved by setting barrier height to its minimum value of 27, and box-beam to the value of 27 inches.

On the other hand, based on above formula, if the shoulder width is set as a value less than 5.5 feet, being 0, the optimum value of traffic barrier height could be achieved by setting the value at 30 inches, and box-beam at its fixed value of 27 inches. It should be noted that various values are in the literature review regarding optimal values of W-beam, between 27 and 31, and box-beam for a value of 27 inches (American Association of State Highway and Transportation Officials, 2011; Fang, Gutowski, Li, & DiSogra, 2013; Teng, Liang, & Tran, 2015). So in this study, in order to be consistent and to prevent confusion, one value of 27 was chosen as a value that all the barriers would be set to be optimized.

4.3. Summary and conclusions

The main goal of traffic barriers is to minimize the cost of run-off-road crashes in terms of severity levels. There are a variety of factors that could impact the severity of barrier crashes. These include various environmental, human, roadway, and barrier geometric characteristics. State DOTS do not have the resources to address all the factors immediately, but they need to have a process to optimize the required resources to bring barrier geometric characteristics to the standard dimensions.

This study highlighted the importance of considering the interaction of shoulder width and barrier height in predicting traffic barrier crash severity. However, due to roadway conditions and budget limitation, especially in a mountainous area like Wyoming, it is not possible to optimize the shoulder in most parts, so shoulder width was initially constrained in this study to be constant despite its important interaction with barrier height. As a result, the most viable option that DOTs have is to optimize barrier heights only.

Thus, this study conducted a cost-benefit analysis of changing barrier heights to the optimal values. As a result, only barriers were included in this study that are outside the recommended height range: there is a possibility of override or underride barrier crashes. Also, to account for the heterogeneity resulted from the data structure (various barriers types), possible interaction terms between barrier types and different important predictors were considered. The rarity of a barrier crash makes it hard for decision makers to identify the real hazard for barriers, especially if they have not experienced any crash. This situation is more challenging for areas with low traffic like Wyoming with low traffic.

Only predictors were incorporated in the analysis that are common across both groups of barriers: with and without crash. A novel machine learning technique (quantile regression model), was implemented in this study to account for the sparse nature of crash cost. After training the model over barriers with crashes, the trained model was employed on barriers with no crashes to predict a current crash cost based on the available predictors. Then, barrier heights were changed to an optimal value and consequently the cost with enhancement was predicted. A comparison was made between predicted cost with enhancement and barrier optimization cost, and predicted cost with no barrier enhancement to come up with a benefit that can be obtained by optimization traffic barriers. A detailed description of the steps were discussed in the content of this study.

In summary, it was found that while WYDOT needs to spend more than 10 million dollars to bring all the state-highway barriers to optimal values. However, in 10 years they could expect not only to recover the money being spent, but gain more than 60 million dollars by preventing the predicted crash cost by barrier height enhancement. An additional scenario was discussed in this study in which a shoulder constraint was set free for only box-beam barriers with no crashes. For these barriers, shoulder width enhancement, along with barrier heights, resulted in a significant increase in total benefit.

The optimization approach taken in this study could help decision makers in Wyoming to take appropriate measures to utilize the limited resources by targeting first those barriers that are more cost effective to be optimized. Although the literature contains several optimization studies in the transportation area, this is one of the first studies implemented on traffic barrier data to estimate barrier dimensions enhancement benefits based on a machine learning technique. This study also estimated the benefit of enhancement to not only barriers with crashes, but also to those barriers without any historical crashes. The WYDOT would implement the recommendations of this study to prioritize the enhancement of barriers in the state. Other states can follow similar procedure to prioritize their barrier enhancements.

For future optimizations, more flexibility is needed in terms of variables that could be changed, especially for barriers with no crashes. For instance, it is recommended to set the constraint off the shoulder width for all barriers, especially barriers with no crashes, so the highest benefit could be achieved while minimizing costs. Also based on the literature review, human factors are one the most important factors that can impact crash severity. It would be interesting to identify the impact of human factors on crash cost in future studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge that this work is part of project #RS03218 funded by the Wyoming Department of Transportation (WYDOT).

References

American Association of State Highway and Transportation Officials. (2011). Roadside design guide.
Fang, H., Gutowski, M., Li, N., & DiSogra, M. (2013). No title. Performance Evaluation of NCDOT W-Beam Guardrails Under MASH TL-2 Conditions.
FARS, N. & National Highway Traffic Safety Administration. (2016). Fatality analysis reporting system. On-Line: Http://Www-Fars.Nhtsa.Dot.Gov/Main/Index.Aspx,
Geraci, M. (2014). Linear quantile mixed models: The lqmm package for laplace quantile regression. *Journal of Statistical Software, 57*(13), 1–29.
Jha, M. K., & Schonfeld, P. (2004). A highway alignment optimization model using geographic information systems. *Transportation Research Part A: Policy and Practice, 38*(6), 455–481.
Julin, R. D., Asadollahi Pajouh, M., Stolle, C. S., Reid, J. D., & Faller, R. K. (2017). *Maximum Mounting Height for Midwest Guardrail System (MGS)*.
Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society, 46*(1), 33. https://doi.org/10.2307/1913643.
Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives, 15*(4), 143–156.
Li, C., Ding, L., & Zhong, B. (2019). Highway planning and design in the Qinghai-Tibet plateau of china: A Cost-Safety balance perspective. *Engineering, 5*(2), 337–349.
Li, W., Pu, H., Schonfeld, P., Yang, J., Zhang, H., Wang, L., & Xiong, J. (2017). Mountain railway alignment optimization with bidirectional distance transform and genetic algorithm. *Computer-Aided Civil and Infrastructure Engineering, 32*(8), 691–709.

National Highway Traffic Safety Administration. (2017). 2016 fatal motor vehicle crashes: Overview. Traffic Safety Facts-Research Note (DOT HS 812 456).Online Verfügbar Unter Https://Crashstats.Nhtsa.Dot.Gov/Api/Public/Publication/812456, Zuletzt Geprüft Am, 27, 2018.
OECD. Publishing. (2017). Road safety annual report 2017 OECD Publishing.
Rezapour, M., & Ksaibati, K. (2018). Application of multinomial and ordinal logistic regression to model injury severity of truck crashes, using violation and crash data. *Journal of Modern Transportation, 26*(4), 268–277.
Rezapour, M., Wulff, S. S., & Ksaibati, K. (2019). Examination of the severity of two-lane highway traffic barrier crashes using the mixed logit model. *Journal of Safety Research, 70*, 223–232.
Rezapour, M., & Ksaibati, K. (2020a). Two-Component Bayesian Hierarchical Models for Cost-Benefit Analysis of Traffic Barrier Crash Count. *Algorithms, 13*(8), 179.
Rezapour, M., & Ksaibati, K. (2020b). Application of Bayesian Hierarchical Negative Binomial Finite Mixture Model for Cost-Benefit Analysis of Barriers Optimization. *Accounting for Severe Heterogeneity. Algorithms, 13*(11), 288.
Sandt, A., & Al-Deek, H. (2019). An optimization approach for deployment of intelligent transportation systems wrong-way driving countermeasures. *Journal of Intelligent Transportation Systems*, 1–14.
Teng, T., Liang, C., & Tran, T. (2015). Effect of various W-beam guardrail post spacings and rail heights on safety performance. Advances in Mechanical Engineering, 7(11), 1687814015615544.
Transportation Officials. Task Force for Roadside Safety. (2011a). Roadside design guide AASHTO.
Transportation Officials. Task Force for Roadside Safety. (2011b). Roadside design guide AASHTO.
Wiebelhaus, M. J., Lechtenberg, K. A., Sicking, D. L., Faller, R. K., & Rosenbaugh, S. K. (2013). Cost-effective treatment of existing guardrail systems.

Mahdi Rezapour, PhD, is a post doc researcher at the WYT2. He is passionate about statistics and machine learning technique.

Khaled Ksaibati, PhD, P.E., obtained his BS degree from Wayne State University and his MS and PhD degrees from Purdue University. Dr. Ksaibati worked for the Indian Department of Transportation for a couple of years prior to coming to the University of Wyoming in 1990. He was promoted to an associate professor in 1997 and full professor in 2002. Dr. Ksaibati has been the director of the Wyoming Technology Transfer Center since 2003.

Journal of Safety Research 78 (2021) 146–154

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Changes in driving performance after first and second eye cataract surgery: A driving simulator study

Lynn B. Meuleners [a,*], Jonathon Q. Ng [b], Michelle L. Fraser [a], Dewi Tjia [a], Ying Ru Feng [a], Nigel Morlet [b]

[a] Western Australian Centre for Road Safety Research, School of Population and Global Health, The University of Western Australia, Perth, Western Australia, Australia
[b] Eye & Vision Epidemiology Research (EVER) Group, Perth, Western Australia, Australia

ARTICLE INFO

Article history:
Received 25 September 2020
Received in revised form 27 January 2021
Accepted 28 April 2021
Available online 24 May 2021

Keywords:
Cataract surgery
Driving performance
Driving simulator
Visual measures

ABSTRACT

Introduction: This study investigated the separate impact of first eye and second eye cataract surgery on driving performance, as measured on a driving simulator. *Method:* Forty-four older drivers with bilateral cataract aged 55+ years, awaiting first eye cataract surgery participated in a prospective cohort study. They completed a questionnaire, visual tests and a driving simulator assessment at three time points: before first eye, after first eye, and after second eye cataract surgery. Generalized Estimating Equation Poisson or linear regression models were undertaken to examine the change in four driving outcomes of interest after adjusting for cataract surgery and other potential confounders. *Results:* The rate of crashes/near crashes decreased significantly by 36% (incidence rate ratio (IRR) 0.64, 95% CI 0.47–0.88, $p = 0.01$) after first eye surgery and 47% (IRR 0.53, 95% CI 0.35–0.78, $p < 0.001$) after second eye surgery, compared to before first eye cataract surgery, after accounting for confounders. The rate of crashes/near crashes also decreased with better contrast sensitivity (IRR 0.69, 95% CI 0.48–0.90, $p = 0.041$). A separate model found that time spent speeding 10 kilometers per hour or more over the limit after second eye surgery was significantly less (0.14 min, $p = 0.002$), compared to before first eye surgery, after accounting for confounders. As contrast sensitivity improved, the duration of speeding also decreased significantly by 0.46 min ($p = 0.038$). There were no statistically significant changes in lane excursions or speed variation. *Practical applications:* The findings highlight the importance of timely first and second eye cataract surgery to ensure driver safety, especially as older drivers wait for second eye cataract surgery. It also provides further evidence that contrast sensitivity is probably a better predictor of driving ability in older drivers with cataract than visual acuity, the measure on which driver licensing requirements are currently based, and should also be used when assessing fitness to drive.

1. Introduction

As the population ages, the number of older drivers on the road is also increasing. The past 10 years has seen the number of licensed drivers aged 65+ years increase by 44%, compared to the average total population increase of 17% (Bureau of Infrastructure Transport and Regional Economics (BITRE), 2014). Over 60% of those aged 75+ years hold a driver's license, a 54% increase from a decade ago (Bureau of Infrastructure Transport and Regional Economics (BITRE), 2014). Driving provides a form of independence and social inclusion for many older adults, with driving cessation linked to depression, decreased participation in social activities, declines in physical health, cognitive decline, institutionalization in care-facilities, and mortality (Chihuri et al., 2016). Although many age-related medical conditions can decrease driving ability (Dobbs, 2005), visual impairment is regarded as having a major impact on driving performance and presents a significant economic and public health challenge.

Cataract is a leading cause of visual impairment worldwide with an estimated prevalence of 75% by age 80 (Australian Institute of Health and Welfare (AIHW), 2005). However visual impairment due to cataract can be successfully treated with cataract surgery, which is routinely performed on each eye

Abbreviations: ANOVA, analysis of variance; CI, confidence interval; GEE, Generalized Estimating Equations; ETDRS, Early Treatment Diabetic Retinopathy Study; IRR, incidence rate ratio; km, kilometers; km/h, kilometers per hour; logMAR, logarithm of the minimum angle of resolution; MMSE, Mini Mental State Examination; SD, standard deviation; WA, Western Australia.

* Corresponding author at: Western Australian Centre for Road Safety Research, School of Population and Global Health, The University of Western Australia, Clifton Street Building, Nedlands, WA 6009 Australia.

E-mail address: lynn.meuleners@uwa.edu.au (L.B. Meuleners).

https://doi.org/10.1016/j.jsr.2021.04.006
0022-4375/

separately. Since driving is heavily dependent on vision, it is likely that driving performance may change after both first-eye and second-eye cataract surgery.

Despite growing evidence for decreased crash risk with cataract surgery (Meuleners, Brameld, Fraser, & Chow, 2019; Owsley et al., 2002), there is less understanding of the impact of cataract surgery on specific aspects of driving performance. A meta-analysis by Subzwari et al. found that self-reported driving difficulty reduced by 88% following cataract surgery, however several of the observational studies included in the review did not define whether participants had undergone surgery on the first or second eye or they analyzed both eyes together (Subzwari et al., 2008). Three other studies of visually normal participants used goggles to simulate cataract and found that participants wearing the goggles performed poorer on video-based hazard perception tests (Marrington, Horswill, & Wood, 2008), closed-circuit daytime driving (Wood, Chaparro, & Hickson, 2009), and closed-circuit night-time driving (Wood, Chaparro, Carberry, & Chu, 2010). Only one study examined the closed-road driving performance of older drivers with bilateral cataract before first eye surgery and following surgery in both eyes (Wood & Carberry, 2006). This study reported significant improvements in driving performance after both eyes were operated on in terms of overall driving score, road sign recognition, road hazards recognized, and road hazards avoided (Wood & Carberry, 2006). However, people with bilateral cataract may need to wait six months to a year between operations in the Australian public health system and may continue to drive (Australian Institute of Health and Welfare, 2010). While first eye surgery brings about significant improvements in vision, frequent problems are reported while awaiting second-eye surgery, most likely due to differences in vision between the operated and un-operated eyes (Comas, Castells, Acosta, & Tuñí, 2007). Also, many cataract patients continue to drive with their old spectacles until after the second eye surgery has been completed, which may impact their driving performance between the first and second eye surgery (Fraser, Meuleners, Lee, Ng, & Morlet, 2013).

There is a significant gap in the evidence surrounding the *separate impacts* of first and second eye cataract surgery on *specific aspects* of driving performance. It is essential to comprehensively understand these separate impacts so that bilateral cataract patients can be advised on whether they are safe to drive while waiting for the second eye surgery and also what specific aspects of driving performance they may have difficulty with while waiting for the first and second eye surgery. This would also allow patients to then be advised on measures they could take to make driving safer, particularly between the first and second eye surgery when they may have increased confidence in their driving abilities due to the first eye surgery, but are still waiting for the second eye surgery.

Driving simulators offer a safe and effective method for examining driving performance as they represent an approach that is repeatable and easily adaptable, including the ability to quickly alter driving scenarios and expose drivers to hazardous situations in a systematic way (Godley, Triggs, & Fildes, 2002). They also have external face validity, can distinguish safe from unsafe drivers (Lee, Cameron, & Lee, 2003), and can be programmed to test general measures of driving as well as driving tasks thought to be problematic for people with cataract. Examples of these measures include speed, lateral position, braking response, inattention, and risky driving behavior (Blana, 1996; Hoskins & El-Gindy, 2006). Therefore, the aim of this study was to investigate the separate impact of first eye and second eye cataract surgery on specific changes in driving performance such as non-compliance to speed limit, speed variation, lane excursions, and the number of crashes/near crashes in a driving simulator.

2. Material and methods

2.1. Study design, participants and sample size

A prospective cohort study of older adults with bilateral cataract awaiting first eye cataract surgery was undertaken in Perth, Western Australia (WA) as part of the larger Cataract Extraction and Driving Ability Research study (Meuleners et al., 2015). The current paper is based on the driving simulator component of the study and a sample of drivers who agreed to undertake the three driving simulator assessments. A power analysis using the Gpower computer program (Faul et al., 2009) indicated that a total sample of 43 people would be needed to detect medium effect with 95% power at alpha 0.05 based on weekly kilometers traveled obtained from the in-vehicle monitoring device.

Informed written consent was obtained from all participants and the study was conducted in accordance with the tenets of the Declaration of Helsinki. Ethics Committee approval was obtained from all participating hospitals and from Curtin University's Human Research Ethics Committee (HR 29/2014).

The recruitment of participants for the larger study was carried out by a letter of invitation sent to the participant or directly through an ophthalmologist, at one of three recruiting hospitals between December 2014 and February 2017. Inclusion criteria included that a participant be aged 55+, have a diagnosis of bilateral cataract but no other significant eye conditions (such as glaucoma, macular degeneration or diabetic retinopathy), a current WA driver's license, drive at least twice a week, and a score of $\geq$ 24 on the Mini Mental State Examination (MMSE). Participants were excluded from the study if they were wheelchair-bound and/or diagnosed with one of the following conditions: dementia, Alzheimer's disease or Parkinson's disease, did not speak English, lived in a residential care facility, had previously undergone cataract surgery, history of vomiting/seasickness, previous head injury, or an upper respiratory tract infection at the time of driving simulator assessment.

2.2. Data collection

Information was collected at three-time points: the month before first eye surgery, one to three months after first eye surgery, and at least one month after second eye surgery. Information collected for the study included a researcher-administered questionnaire, three objective visual measurements, and a driving simulator assessment.

2.2.1. Questionnaire

Socio-Demographic Data: Information on age, gender, marital status, country of birth, education level, living arrangements, self-reported prescription medications, co-morbid medical conditions, retirement status and driving were obtained from the researcher-administered questionnaire.

Cognitive Assessment: Participants completed the MMSE (Folstein, Folstein, & McHugh, 1975) to measure cognitive status. A MMSE score of $\geq$24 was required for inclusion into the study and indicated normal cognitive function (score range is 0–30).

2.2.2. Visual measures

Three visual assessments were undertaken at each of the three time-points in the study. These assessments were made under the guidance of an ophthalmologist in standard conditions with constant luminance and no mydriasis (dilation of the pupil). Participants wore their habitual corrective lenses or glasses for the visual testing. Monocular and binocular visual acuity was obtained using the letter by letter scoring method on an Early Treatment

Diabetic Retinopathy Study (ETDRS) acuity chart calibrated for a distance of three meters (Ferris, Kassoff, Bresnick, & Bailey, 1982). The logarithm of the minimum angle of resolution (logMAR) indicated the visual acuity score. Monocular and binocular contrast sensitivity were obtained using the Mars Letter Contrast Sensitivity Test (© Mars Perceptrix), calibrated at a distance of 50 centimeters. Stereopsis was assessed using the Titmus Fly Stereotest (Stereo Optical Co., Inc.), measured in log seconds of arc.

2.3. Driving simulator

A PC-based driving simulator was used for the driving assessment (Fig. 1). The simulator consisted of a car with automatic transmission, a driver's seat, a steering wheel, three-dimensional visuals, an accelerator and brake pedals. The visual display system consisted of three color monitors spanning a 180-degree field of view synchronized to display a realistic view of a computer-generated road environment. It also included an audio system that provided realistic traffic sounds. The driving simulator was previously validated against an on-road driving assessment in WA based on the type and number of driving errors, which included maintaining correct speed for the road, adjusting speed as approaching/negotiating intersections, maintenance of correct lane and near crashes (Meuleners & Fraser, 2015).

2.3.1. Driving scenario

The 10 minute simulated driving scenario consisted of a three-dimensional model, which represented approximately 10 kilometers (km) of generic WA road in the metropolitan area, during the daytime. The simulator scenario required drivers to negotiate two give way signs (one left, one right turn), one stop sign (continue straight), one merge, three uncontrolled intersections (one left, two right turns), five sets of traffic lights (two right, one left turn and two continue straight), and five roundabouts (two right turns and three continue straight). Traffic was programmed throughout the route with cars, buses and motorcycles traveling in the same and opposite directions to the driver. Traffic was also programmed at most intersections to allow similar opportunities for driving errors as on road. Other programmed events (hazards) were a bus stopping at a bus stop, a truck stopped by the side of the road and a pedestrian crossing at a designated pedestrian crosswalk. They were included to at least partially simulate the extra cognitive load that accrues from the requirement of drivers to monitor and negotiate the road with other road users. It was anticipated that these events would test driving performance in potential problem areas for drivers at the three different stages of their cataract surgery. Speed limits varied depending on the area so participants were assessed under several different road conditions.

Fig. 1. Driving simulator.

2.3.2. Driving simulator assessment

At each of the three assessments, participants were instructed to operate the simulator as they would normally drive their own car. They were given the opportunity to drive a practice circuit for approximately five minutes in order to familiarize themselves with the vehicle control dynamics, road environment, and simulator tasks (e.g., able to use turn signals; side mirrors; accelerator and brake pedal). This also provided participants with an opportunity to ask any questions before the start of their drive and to make any necessary adjustments to the vehicle so that they were comfortable prior to commencing. The practice circuit consisted of a straight section of the road.

2.3.3. Driving simulator measures

The OKTAL SCANeRTM studio software package was used to simulate the driving experience and record the driving data. Images were displayed in full high definition resolution of 1920×1080 pixels per channel and updated at a frame rate of 120 Hz. The use of simulation allows researchers to measure changes in a variety of surrogate measures that may indicate crash risk (Freund, Gravenstein, Ferris, & Shaheen, 2002; Lee & Lee, 2005). The following surrogate measures were recorded: non-compliance to speed limits, speed variation, lane excursions, and number of crashes/near crashes.

Non-compliance to speed limits: As speed limits were different throughout the simulated drive it was reasonable to assess whether drivers complied with the speed limit. This was measured in minutes spent $\geq$ 10 kilometers per hour (km/h) over the posted speed limit and was used instead of mean speed. It was hypothesized that non-compliance to the speed limit would reflect difficulty with sign recognition and visualizing the speedometer (Wood, 2002; Wood & Carberry, 2006).

Speed Variation: Speed variation of participants over the entire scenario was also assessed, which was measured as the standard deviation of speed in km/h. Speed variation is important to include as it is considered a general measure of driving ability. It was also hypothesized that higher speed variation may indicate harsh braking and/or acceleration and this may be due to late detection of signs, hazards, or road infrastructure due to vision.

Lane Excursions: To measure lateral control, it was considered that stability measures such as lane deviation standard deviation would not be an appropriate measure on a road that contained both straight and curved sections. Young and Stanton found lateral stability can be misrepresentative of proper driving techniques on curved road sections (Young & Stanton, 2002). Therefore, time spent with any part of the vehicle out of the lane (minutes) was used to evaluate lateral control, with the assumption that good driving performance was characterized by less time out of the lane (Wood, 2002). Lane excursions are a commonly used measure of general driving performance. It was also hypothesized that visual impairment due to reduced contrast sensitivity may result in difficulty detecting the edge of the road and lane markings and more time spent out of the lane (Mäntyjärvi & Tuppurainen, 1999; West et al., 2002).

Crashes/Near Crashes: There were three hazards presented during the simulated drive and the number of resulting crashes or near crashes (combined) was measured. Near crashes were defined as any circumstance that required a rapid evasive maneuver by the driver to avoid a crash (Virginia Tech Transportation Institute, 2015). An evasive maneuver may include steering, braking, or accelerating. Near crashes and crashes were identified by the observing researcher. The researcher was trained to identify crashes and near crashes as per this definition. The researcher recorded a detailed description of each event that was reviewed by a second experienced researcher. The two researchers then came to an agreement on whether the crash or near crash fit the criteria for inclusion in the study. It was hypothesized that older adults with cataract may have poorer hazard recognition and avoidance (Wood & Carberry, 2006), resulting in more crashes/near crashes.

2.4. Statistical analysis

Descriptive analyses were undertaken to describe the demographic and driving profile of the cohort. A repeated-measures analysis of variance (ANOVA) was undertaken to analyze the changes in visual measurements, objective driving exposure, and three of the driving simulator measures (non-compliance to speed limit, speed variation, and lane excursions) during each stage of cataract surgery (before first eye, after first eye, and after second eye surgery).

Four separate Generalized Estimating Equation (GEE) Poisson or linear regression models were undertaken for the four driving simulator outcomes of interest (non-compliance to speed limit, speed variation, lane excursions and number of crashes/ near crashes), depending on whether the outcome was categorical or continuous. The GEE method is suitable for longitudinal or repeated measures study designs where observations within each participant are not independent (Zeger & Liang, 1986). GEEs permit specification of a certain working correlation matrix that accounts for this within-subject correlation, providing more robust regression coefficients. Stepwise variable selection was performed using backwards elimination, resulting in the final models.

The cataract surgery time points (before first eye, after first eye and after second eye) were included as an independent variable in each of the four models. Other potential confounders considered for inclusion in the GEE models, based on the literature and our previous experience were: gender (female, male), age group (55–69, 70+ years), marital status (single/separated/divorced/widowed versus married/de facto), prescription medication/s (no, yes), retirement status (not retired, retired), number of years driving, and weekly driving exposure (km traveled). To account for weekly driving exposure in the analysis, all participants had an in-vehicle monitoring device installed in their car at the time of the driving simulator assessment. It measured their driving exposure (km driven) over a one-week period. Visual acuity, contrast sensitivity, and stereopsis measures were also included in the models. Refractive management of visual impairment was not included as a confounding factor as all visual assessments were undertaken with habitual eyewear where applicable. Since 98% of the cohort reported a co-morbid medical condition, this was also not included in the model. All analyses were carried out using R statistical software. *P*-values less than 0.05 were considered statistically significant.

3. Results

A total of 44 participants completed the three driving assessments which provided a total of 132 observations for the analyses.

3.1. Socio-demographic characteristics

The mean age of participants before first eye cataract surgery was 73.2 years (SD 8.3) with a range from 56 to 88 years. Most participants were male ($n = 23$, 52.3%), married/ de facto ($n = 29$, 65.9%), did not live alone ($n = 25$, 56.8%), had higher than secondary school education ($n = 25$, 56.8%), were not born in Australia ($n = 29$, 65.9%), were on at least one prescription medication ($n = 38$, 86.4%), had at least one co-morbid medical condition ($n = 43$, 97.7%), and were retired ($n = 35$, 79.6%) (Table 1). At baseline, the mean MMSE score was 27.6 (SD 2.36). The driving simulator assessment after first eye cataract surgery occurred between 9 and 417 days with a mean of 99.6 days (SD 73.7). The driving simulator assessment after second eye cataract surgery occurred between 29 and 238 days, with a mean of 112.3 days (SD = 40.6). This range in length of time between cataract surgery and assessment is unlikely to impact on participants' vision since the artificial lenses do not deteriorate.

3.2. Visual measures

Mean binocular visual acuity significantly improved from 0.15 (SD 0.15) logMAR at baseline, to 0.09 (SD 0.22) logMAR after the first eye surgery, and then to −0.01 (SD = 0.20) logMAR after the second eye surgery) ($p < 0.001$). Binocular contrast sensitivity was similar between baseline, 1.65 (SD 0.14) log units and after the first eye surgery, 1.66 (SD 0.28) log units, however significantly improved to 1.75 (SD 0.08) log units after the second eye cataract surgery ($p = 0.026$). Stereopsis worsened from 2.11 (SD 0.64) log seconds of arc at baseline to 2.26 (SD 0.73) after the first eye surgery, and significantly improved to 1.91 (SD 0.60) log seconds of arc after the second eye surgery ($p = 0.008$) (Table 2).

3.3. Self-reported driving ability and objective driving exposure

The study found that over 66% ($n = 29$) of participants self-reported the quality of their driving as excellent or very good before the first eye cataract surgery. After the second eye surgery this improved to 75% ($n = 33$). In terms of objective driving exposure, which was recorded using the in-vehicle driving monitoring

Table 1
Socio-demographic characteristics of the study population before first eye cataract surgery ($n = 44$).

Characteristic	*N*	%
Gender		
Female	21	47.7
Male	23	52.3
Age Group (years)		
55–69	12	27.3
70–84	29	65.9
85+	3	6.8
Marital Status		
Single (Single, separated, divorced, widowed)	15	34.1
Married/de facto	29	65.9
Living Arrangements		
Alone	19	43.2
Not alone	25	56.8
Education Level		
Primary or secondary school	19	43.2
Higher education	25	56.8
Country of Birth		
Not Australia	29	65.9
Australia	15	34.1
Prescription Medication/s		
No	6	13.6
Yes	38	86.4
Co-morbid Medical Condition/s		
No	1	2.3
Yes	43	97.7
Retired		
No	9	20.5
Yes	35	79.6

Table 2
Visual measurements for participants before first eye, after first eye, and after second eye cataract surgery ($n = 44$).

Visual measure	Before first eye surgery Mean (SD)	After first eye surgery Mean (SD)	After second eye surgery Mean (SD)	*P*-value*
Visual Acuity (logMAR)				
Better eye	0.17 (0.15)	0.10 (0.24)	0.02 (0.21)	<0.001
Worse eye	0.37 (0.26)	0.35 (0.27)	0.11 (0.21)	<0.001
Binocular	0.15 (0.15)	0.09 (0.22)	−0.01 (0.20)	<0.001
Contrast Sensitivity (log units)				
Better eye	1.58 (0.14)	1.62 (0.31)	1.69 (0.12)	0.052
Worse eye	1.41 (0.31)	1.48 (0.29)	1.62 (0.13)	<0.001
Binocular	1.65 (0.14)	1.66 (0.28)	1.75 (0.08)	0.026
Stereopsis (log seconds of arc)				
Binocular	2.11 (0.64)	2.26 (0.73)	1.91 (0.60)	0.008

Abbreviations: SD: standard deviation, logMAR: logarithm of the minimum angle of resolution.
* *p*-values calculated using repeated ANOVA.

device, the mean km traveled per week before the first eye cataract surgery was 99.0 km (SD 92.3), which decreased to 84.4 km (SD 87.9) after the first eye cataract surgery, but increased to 125.1 km (SD 129.8) after the second eye surgery.

3.4. Driving simulator outcomes

The total number of crashes/near crashes decreased after the first and second eye cataract surgery (Table 3). In terms of non-compliance to the speed limits (total minutes spent ≥ 10 km/h over the speed limit), there was a significant decrease in speeding after the first and second eye surgery ($p = 0.009$) (Table 4). The total time spent ≥ 10 km/h over the speed limit was 0.73 min (44 s) before the first eye surgery and reduced to 0.66 min (40 s) after the first eye surgery and 0.41 min (25 s) after the second eye surgery. There was no significant difference in lane excursions (time spent out of lane) ($p = 0.230$) or speed variation (standard deviation of speed) ($p = 0.098$) after the first and second eye surgery (Table 4).

3.5. First and second eye cataract surgery and crashes/near crashes

The results of the GEE Poisson model found that the rate of crashes/near crashes significantly decreased by 36% (IRR 0.64, CI 0.47–0.88, $p = 0.01$) after the first eye surgery and 47% (IRR 0.53, CI 0.35–0.78, $p < 0.001$) after the second eye surgery, compared to before the first eye cataract surgery, after accounting for all relevant confounders. Post-hoc testing also found that the rate of crashes/near crashes significantly decreased after second eye cataract surgery compared to after first eye cataract surgery ($p = 0.03$). Males had over twice (IRR 2.12, CI 1.43–3.14) the rate of crashes/near crashes compared to females ($p = 0.002$). Retired drivers also had a significantly higher rate of crashes/near crashes compared to non-retired participants (IRR 1.77, CI 1.04–3.03, $p = 0.045$). The rate of crashes/near crashes decreased with better contrast sensitivity (IRR 0.69, CI 0.48–0.90, $p = 0.041$). Lastly, as driving exposure (km/week) increased, the rate of crashes/ near crashes in the driving simulator significantly decreased (IRR 0.90, CI 0.90–0.91, $p = 0.043$) (Table 5).

Table 3
Number of crashes/near crashes in the driving simulator before first eye, after first eye and after second eye cataract surgery ($n = 44$).

Number of crashes/ near crashes	Before first eye surgery *N* (%)	After first eye surgery *N* (%)	After second eye surgery *N* (%)
1	21 (42.9%)	16 (32.7%)	12 (24.5%)
2	8 (42.1%)	7 (36.8%)	4 (21.1%)
≥3	4 (57.1%)	2 (28.6%)	1 (14.3%)
Total	33 (44.0%)	25 (33.3%)	17 (22.7)

3.6. First and second eye cataract surgery and non-compliance to speed limit

The results of the GEE linear regression model found that the amount of time speeding after the second eye surgery was significantly less (0.14 min), compared to before the first eye surgery, after accounting for confounders in the model ($p = 0.002$). However, there was no significant change in speeding from before to after the first eye surgery ($p = 0.426$). Post-hoc testing also found no significant differences in the amount of time speeding after the first and second eye cataract surgery ($p = 0.52$). Males reported significantly more time speeding (0.62 min), compared to females ($p = 0.003$). As expected, those aged 70+ reported significantly less time speeding (0.51 min), compared to those aged 55–69 years ($p = 0.030$). As contrast sensitivity in both eyes improved, the amount of time speeding significantly reduced by 0.46 min ($p = 0.038$) (Table 6).

There were no significant changes in the lane excursions and speed variation outcomes over the cataract surgery process so the multivariate modeling results have not been reported.

4. Discussion

This is one of the first studies to use a driving simulator to examine the separate impacts of the first and second eye cataract surgery on driving performance. No previous study has specifically quantified driving performance between the first and second eye cataract surgery but have either combined both surgeries or examined driving after the first eye or second eye surgery only.

The study found that the rate of crashes/near crashes in the simulator decreased significantly by 36% after the first eye cataract surgery and 47% after the second eye surgery, compared to before the first eye surgery. In addition, crashes/ near crashes also decreased significantly after the second eye surgery compared to after the first eye surgery. These findings demonstrate the positive effects of the first eye surgery and the important additional benefits of the second eye surgery for those with bilateral cataract, in terms of road safety. These findings build on those from a previous population-based Australian study, which showed that real world crash risk reduced after both the first and second eye cataract surgery (Meuleners et al., 2019). An earlier cohort study conducted in the United States also reported significantly reduced crash risk following cataract surgery, although this study combined those who had unilateral and bilateral eye surgery in the analysis (Owsley et al., 2002). The observed reduction in crashes/near crashes may be due to improved hazard perception ability as the result of vision improvements from the first and second eye cataract surgery. A closed-road study also reported marked improvements in the ability to detect and avoid hazards after bilateral cataract surgery (Wood & Carberry, 2006).

Table 4
Speed limit compliance, lane excursions and speed variation in the driving simulator before first eye, after first eye and after second eye cataract surgery ($n = 44$).

Driving outcome	Before first eye surgery Mean (SD)	After first eye surgery Mean (SD)	After second eye surgery Mean (SD)	*P*-value*
Speed limit compliance†	0.73 (0.91)	0.66 (0.79)	0.41 (0.70)	0.009
Lane excursions‡	1.19 (2.30)	1.12 (1.69)	0.79 (0.16)	0.230
Speed variation§	13.20 (2.96)	13.12 (2.45)	12.52 (2.26)	0.098

Abbreviations: SD: standard deviation.
* *p*-value calculated using repeated ANOVA.
† Minutes ≥10 km/h over posted speed limit.
‡ Time spend out of lane (minutes).
§ Standard deviation of speed (km/h).

Table 5
GEE Poisson model of the impact of first and second eye cataract surgery on number of crashes/near crashes in the driving simulator ($n = 44$).

Variable	IRR	95% CI		*P*-Value
Cataract Surgery				
Before first eye surgery	1.00			
After first eye surgery	0.64	0.47	0.88	0.01
After second eye surgery	0.53	0.35	0.78	<0.001
Gender				
Female	1.00			
Male	2.12	1.43	3.14	0.002
Retirement Status				
Not retired	1.00			
Retired	1.77	1.04	3.03	0.045
Contrast Sensitivity† (log units)	0.69	0.48	0.90	0.041
Driving exposure (km/week)	0.90	0.90	0.91	0.043

Abbreviations: GEE: Generalized Estimating Equations, IRR: Incidence rate ratio, CI: confidence interval.
† Contrast sensitivity measured with both eyes.

Table 6
GEE linear regression model of the impact of first and second eye cataract surgery on speed limit compliance* in the driving simulator ($n = 44$).

Variable	Parameter Estimate	95% CI		*P*-Value
Cataract Surgery				
Before first eye surgery (reference)				
After first eye surgery	−0.04	−0.25	0.18	0.426
After second eye surgery	−0.14	−0.30	−0.02	0.002
Gender				
Female (reference)				
Male	0.62	0.23	0.89	0.003
Age Group				
55–69 years (reference)				
70+ years	−0.51	−0.89	−0.12	0.030
Contrast Sensitivity‡ (log units)	−0.46	−0.87	−0.06	0.038

Abbreviations: CI: confidence interval.
* Minutes spent ≥ 10 km/h over the speed limit.
‡ Contrast sensitivity measured with both eyes.

Time spent speeding ≥ 10 km/h over the posted speed limit also significantly reduced after the second eye surgery compared to before the first eye surgery. This is encouraging as speeding is one of the most common factors contributing to fatal crashes (National Center for Statistics and Analysis, 2018). Although study participants only spent a short amount of time speeding, it should be noted that the simulator scenario was also short. Any amount of time spent speeding is important since even small increases in speed substantially increase the risk of crash involvement, with speeding by 10 km/h in a 60 km/h zone found to increase casualty crash risk by four times (Kloeden, Mclean, Moore, & Ponte, 1997). The improvement in adherence to the speed limit amongst this cohort may be due to an improved ability to read and recognize speed limit signs, as previously reported following bilateral cataract surgery (Wood & Carberry, 2006). Interestingly, this study revealed that the improvement in speeding was only found after the second eye cataract surgery. This may be explained by the significant improvement in binocular contrast sensitivity following second but not first eye cataract surgery in the cohort, as well as the fact that many cataract patients wait until after both surgeries to purchase new spectacles (Fraser et al., 2013). This suggests that access to low-cost temporary spectacles during the waiting period between the first and second eye surgery may be beneficial for the safety of bilateral cataract patients who continue to drive. These results also highlight the importance of a timely second eye surgery for improving driver safety. Participants' driving exposure also significantly decreased after the first eye surgery but increased after the second eye surgery to levels higher than before surgery. This shows the importance of both the first and second eye surgery for restoring older adults' mobility and independence through driving.

This study did not find any significant change in lane excursions or speed variation after the first or second eye cataract surgeries. Possible explanations for these findings may be that driving simulator steering performance is robust to severe blur (Brooks, Tyrrell, & Frank, 2005), which may be experienced with cataract, and that lane keeping difficulties are more common among drivers with visual field defect (Kasneci et al., 2014), which is not a specific symptom of cataract.

Improved contrast sensitivity in both eyes was significantly associated with fewer crashes/near crashes and reduced speeding behavior. This result is not surprising because if the contrast between the target image (e.g., vehicle or pedestrian) and the background is below the driver's ability to detect it, this would increase the risk of a crash (Guo, Fang, & Antin, 2015). A study in the United States also showed that drivers with poor contrast sensitivity due to cataract were eight times more likely to be involved in an at-fault crash (Owsley, Stalvey, Wells, Sloane, & McGwin, 2001). Contrast sensitivity was also shown to be a better predictor of visual recognition of road signs, obstacles, and pedestrians than visual acuity (Wood & Owens, 2005), and a better predictor of self-reported driving difficulty (Fraser et al., 2013). The current findings provide further evidence that contrast sensitivity is probably a better predictor of driving ability in older drivers with cataract than visual acuity, the measure on which driver licensing requirements are currently based. Therefore, contrast sensitivity should also be used when assessing fitness to drive. It should also be noted that the vision of participants in the current study was better than those in previous studies at baseline (Fraser et al., 2013; Owsley et al., 2002; Wood & Carberry, 2006), yet cataract surgery still significantly improved their driving ability. This highlights the importance of timely first and second eye cataract surgery for cataract patients who drive, even at lower levels of visual impairment.

Before the first eye surgery, 66% of participants rated the quality of their driving as excellent or very good and drove an average of 99 km per week. However, 75% experienced crashes/near crashes in the simulator before the first eye surgery. This discrepancy between perceived driving ability and simulator performance may be partially due to differences between simulator and real-world driving, as well as participants being presented with more hazards than they would usually encounter on a short journey. However, it is also possible that drivers do not recognize their reduced driving ability and hazard perception, due to gradual nature of cataract impairment. Previous studies have found that older drivers routinely over-estimate their driving ability (Horswill, Anstey, Hatherly, Wood, & Pachana, 2011; Ross, Dodson, Edwards, Ackerman, & Ball, 2012; Wood, Lacherez, & Anstey, 2013), so this lack of awareness may also have explained the observed discrepancy.

Strengths of this study include the three distinct time points in which objective measures of driving performance and vision were assessed. Most previous research has not measured the specific effects of first and second eye cataract surgery separately (Owsley et al., 2002; Subzwari et al., 2008; Wood & Carberry, 2006). The use of the driving simulator in the study, which was previously validated against an on-road driving assessment (Meuleners & Fraser, 2015) allowed each participant's driving performance to be tested under identical conditions and exposed drivers to hazardous situations in a systematic way that would be impossible, dangerous, and unethical to obtain from on-road driving assessments. In addition, the collection of objective driving exposure information using an in-vehicle device allowed us to control for participants' real-world driving experience at each time point.

Limitations of the study include the sample size, which may not have been large enough to detect all associations. The strict inclusion criteria may also have influenced the generalizability of the results as the sample represented a healthy and functioning group living in the community and cannot be considered representative of all older drivers. A high proportion of participants were not born in Australia (66% vs. 37% of the general older population) and this may also affect generalizability. This is likely because people born outside of Australia are over-represented in the public health system, where participants were recruited from. As well, since some participants waited long periods between the first and second eye surgery, the negative effects of the aging process on driving ability may have contributed to some of the non-significant results. It should also be acknowledged that some improvement in driving performance over the three assessments could be due to learning effects on the simulator. Future studies should use a healthy control group to evaluate these potential learning effects as well as allowing between-group comparisons of driving performance. Other visual measures such as visual field and disability glare were not collected in this study. However, the driving simulator does not replicate the real-world effects of glare. In addition, visual measures were performed using participants' habitual corrective lenses, rather than best-corrected measures (using correct prescription for refraction). It is well known that many cataract patients continue to drive with their old spectacles until after their second eye surgery, which means they may be driving with less than optimal vision during the waiting period between surgeries. The authors believed it was important to measure vision under the natural circumstances in which the participants were actually driving (with or without post-surgery refraction), rather than their optimal vision. The small sample size also meant we could not include all the visual measures in the model so were limited to including bilateral contrast sensitivity only. It would be useful for future studies with larger sample sizes to examine how specific visual measures in the first and second operated eye are associated with changes in driving performance.

5. Conclusion

In conclusion, this study reported fewer crashes/near crashes following the first and second eye cataract surgery, and improved adherence to the speed limit following the second eye surgery in a driving simulator. This highlights the importance of timely first and second eye cataract surgery even for those with lower levels of visual impairment, and the provision of temporary spectacles while waiting for the second eye surgery for cataract patients who drive. The current findings provide further evidence that contrast sensitivity is probably a better predictor of driving ability in older drivers with cataract than visual acuity, the measure on which driver licensing requirements are currently based, and should also be used when assessing fitness to drive. The results of this study may provide opportunities for overall cost savings to the community by funding timely cataract surgery in both eyes to avoid the costs associated with crashes and injury.

Declarations of interest

None.

Acknowledgments

The authors would like to acknowledge John Hess for his work on programming and maintaining the driving simulator used in this study.

Data availability

The datasets generated during and/or analyzed during the current study are not publicly available due personal participant information but are available from the corresponding author on reasonable request.

Funding

This work was supported by an Australian Research Council Discovery Grant #DP140101299. The funding body played no role in the design of the study, the collection, analysis or interpretation of data, or the writing of the manuscript.

Disclosure

The author reports no conflicts of interest in this work.

Authors' contributions

LM is responsible for the conception and design, acquisition, analysis and interpretation of the data, and drafting of the manuscript. JN was involved in the conception, design of the study and drafting of the manuscript. MF was involved in the conception, design of the study and drafting of the manuscript. DT was involved in the analysis and interpretation of the data. YF was involved in the analysis and interpretation of the data. NM was involved in the conception, design of the study and drafting of the manuscript. All authors read and approved the final manuscript.

References

Australian Institute of Health and Welfare (2010). *Australian hospital statistics 2009–2010: Emergency department care and elective surgery waiting times*. Canberra, ACT, Australia: AIHW.

Australian Institute of Health and Welfare (AIHW). (2005). Vision problems among older australians. Bulletin no. 27. AIHW cat. No. Aus 60. AIHW, Canberra, ACT, Australia.

Blana, E. (1996). Driving simulator validation studies: A literature review. Technical report, Institute of Transport Studies, University of Leeds, Working Paper 480.

Brooks, J. O., Tyrrell, R. A., & Frank, T. A. (2005). The effects of severe visual challenges on steering performance in visually healthy young drivers. *Optometry and Vision Science, 82*(8), 689–697. https://doi.org/10.1097/01.opx.0000174722.96171.86.

Bureau of Infrastructure, Transport and Regional Economics (BITRE) (2014). *Road safety of older Australians: Recent statistics*. Canberra: BITRE, Australian Government.

Chihuri, S., Mielenz, T. J., DiMaggio, C. J., Betz, M. E., DiGuiseppi, C., Jones, V. C., & Li, G. (2016). Driving cessation and health outcomes in older adults. *Journal of the American Geriatrics Society, 64*(2), 332–341. https://doi.org/10.1111/jgs.13931.

Comas, M., Castells, X., Acosta, E. R., & Tuñí, J. (2007). Impact of differences between eyes on binocular measures of vision in patients with cataracts. *Eye, 21*(6), 702–707. https://doi.org/10.1038/sj.eye.6702305.

Dobbs, B. (2005). Medical conditions and driving: A review of the scientific literature (1960-2000). Association for the Advancement of Automotive Medicine/National Highway Traffic Safety Administration, Washington, DC.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Ferris, F. L., 3rd, Kassoff, A., Bresnick, G. H., & Bailey, I. (1982). New visual acuity charts for clinical research. *American Journal of Ophthalmology, 94*(1), 91–96.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Fraser, M. L., Meuleners, L. B., Lee, A. H., Ng, J. Q., & Morlet, N. (2013). Which visual measures affect change in driving difficulty after first eye cataract surgery? *Accident Analysis and Prevention, 58*, 10–14. https://doi.org/10.1016/j.aap.2013.04.015.

Freund, B., Gravenstein, S., Ferris, R., & Shaheen, E. (2002). Evaluating driving performance of cognitively impaired and healthy older adults: A pilot study comparing on-road testing and driving simulation. *Journal of the American Geriatrics Society, 50*(7), 1309–1310. https://doi.org/10.1046/j.1532-5415.2002.50325.x.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis and Prevention, 34*(5), 589–600. https://doi.org/10.1016/S0001-4575(01)00056-2.

Guo, F., Fang, Y., & Antin, J. F. (2015). Older driver fitness-to-drive evaluation using naturalistic driving data. The Journal of Safety Research, 54, 49.e29-54. doi: https://doi.org/10.1016/j.jsr.2015.06.013.

Horswill, M. S., Anstey, K. J., Hatherly, C., Wood, J. M., & Pachana, N. A. (2011). Older drivers' insight into their hazard perception ability. *Accident Analysis and Prevention, 43*(6), 2121–2127. https://doi.org/10.1016/j.aap.2011.05.035.

Hoskins, A., & El-Gindy, M. (2006). Technical report: Literature survey on driving simulator validation studies. *International Journal of Heavy Vehicle Systems, 13*. https://doi.org/10.1504/IJHVS.2006.010020.

Kasneci, E., Sippel, K., Aehling, K., Heister, M., Rosenstiel, W., Schiefer, U., & Papageorgiou, E. (2014). Driving with binocular visual field loss? A study on a supervised on-road parcours with simultaneous eye and head tracking. *PLoS ONE, 9*(2). https://doi.org/10.1371/journal.pone.0087470 e87470.

Kloeden, C., Mclean, A., Moore, V., & Ponte, G. (1997). *Travelling speed and the risk of crash involvement*. Adelaide, South Australia: NHMRC Road Accident Research Unit, The University of Adelaide.

Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention, 35*(5), 797–803. https://doi.org/10.1016/S0001-4575(02)00083-0.

Lee, H. C., & Lee, A. H. (2005). Identifying older drivers at risk of traffic violations by using a driving simulator: A 3-year longitudinal study. *American Journal of Occupational Therapy, 59*(1), 97–100. https://doi.org/10.5014/ajot.59.1.97.

Mäntyjärvi, M., & Tuppurainen, K. (1999). Cataract in traffic. *Graefes Archive for Clinical and Experimental Ophthalmology, 237*(4), 278–282. https://doi.org/10.1007/s004170050233.

Marrington, S. A., Horswill, M. S., & Wood, J. M. (2008). The effect of simulated cataracts on drivers' hazard perception ability. *Optometry and Vision Science, 85*(12), 1121–1127. https://doi.org/10.1097/OPX.0b013e31818e8d00.

Meuleners, L., & Fraser, M. (2015). A validation study of driving errors using a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour, 29*, 14–21. https://doi.org/10.1016/j.trf.2014.11.009.

Meuleners, L. B., Agramunt, S., Ng, J. Q., Morlet, N., Keay, L., McCluskey, P., & Young, M. (2015). The cataract extraction and driving ability research study protocol: Characterisation of deficits in driving performance and self-regulation among older drivers with bilateral cataract. *Injury Prevention, 21*(6), 424–429. https://doi.org/10.1136/injuryprev-2014-041423.

Meuleners, L. B., Brameld, K., Fraser, M. L., & Chow, K. (2019). The impact of first- and second-eye cataract surgery on motor vehicle crashes and associated costs. *Age and Ageing, 48*(1), 128–133. https://doi.org/10.1093/ageing/afy156.

National Center for Statistics and Analysis. (2018). Speeding: 2016 data (traffic safety facts. Dot hs 812 480). National Highway Traffic Safety Administration., Washington, DC.

Owsley, C., McGwin, G., Jr., Sloane, M., Wells, J., Stalvey, B. T., & Gauthreaux, S. (2002). Impact of cataract surgery on motor vehicle crash involvement by older adults. *Journal of the American Medical Association, 288*(7), 841–849. https://doi.org/10.1001/jama.288.7.841.

Owsley, C., Stalvey, B. T., Wells, J., Sloane, M. E., & McGwin, G. Jr., (2001). Visual risk factors for crash involvement in older drivers with cataract. *Archives of Ophthalmology, 119*(6), 881–887. https://doi.org/10.1001/archopht.119.6.881.

Ross, L. A., Dodson, J. E., Edwards, J. D., Ackerman, M. L., & Ball, K. (2012). Self-rated driving and driving safety in older adults. *Accident Analysis and Prevention, 48*, 523–527. https://doi.org/10.1016/j.aap.2012.02.015.

Subzwari, S., Desapriya, E., Scime, G., Babul, S., Jivani, K., & Pike, I. (2008). Effectiveness of cataract surgery in reducing driving-related difficulties: A systematic review and meta-analysis. *Injury Prevention, 14*(5), 324–328. https://doi.org/10.1136/ip.2007.017830.

Virginia Tech Transportation Institute. (2015). Researcher dictionary for safery critical event video reduction data, version 4.1. Virginia Tech Transportation Institute. Blacksburg, VA.

West, S. K., Rubin, G. S., Broman, A. T., Muñoz, B., Bandeen-Roche, K., Turano, K., & SEEPT (2002). How does visual impairment affect performance on tasks of everyday life?: The SEE project. *Archives of Ophthalmology, 120*(6), 774–780. https://doi.org/10.1001/archopht.120.6.774.

Wood, J., Chaparro, A., Carberry, T., & Chu, B. S. (2010). Effect of simulated visual impairment on nighttime driving performance. *Optometry and Vision Science, 87*(6), 379–386. https://doi.org/10.1097/OPX.0b013e3181d95b0d.

Wood, J., Chaparro, A., & Hickson, L. (2009). Interaction between visual status, driver age and distracters on daytime driving performance. *Vision Research, 49*(17), 2225–2231. https://doi.org/10.1016/j.visres.2009.06.017.

Wood, J. M. (2002). Age and visual impairment decrease driving performance as measured on a closed-road circuit. *Human Factors, 44*(3), 482–494. https://doi.org/10.1518/0018720024497664.

Wood, J. M., & Carberry, T. P. (2006). Bilateral cataract surgery and driving performance. *British Journal of Ophthalmology, 90*(10), 1277–1280. https://doi.org/10.1136/bjo.2006.096057.

Wood, J. M., Lacherez, P. F., & Anstey, K. J. (2013). Not all older adults have insight into their driving abilities: Evidence from an on-road assessment and implications for policy. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 68*(5), 559–566. https://doi.org/10.1093/gerona/gls150.

Wood, J. M., & Owens, D. A. (2005). Standard measures of visual acuity do not predict drivers' recognition performance under day or night conditions. *Optometry and Vision Science, 82*(8), 698–705. https://doi.org/10.1097/01.opx.0000175562.27101.51.

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors, 44*(3), 365–375. https://doi.org/10.1518/0018720024497709.

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*(1), 121–130. https://doi.org/10.2307/2531248.

Lynn Meuleners, PhD, is a Professor and the Director of the Western Australian Centre for Road Safety Research. She is an injury/road safety epidemiologist with a

special interest in visual impairment and older drivers, data linkage studies, driving simulation research and evaluation of road treatments.

Jonathan Ng, PhD**,** is an ophthalmologist with a special interest in cataract and driving**.** In addition to his medical qualifications and clinical experience, Dr Ng also has academic qualifications relevant to public health and policy.

Michelle Fraser, PhD, is a research associate at the Western Australian Centre for Road Safety Research. She is an epidemiologist with a special interest in visual impairment and older drivers, cycling safety and naturalistic studies.

Dewi Tjia, PhD, is a research associate at the Western Australian Centre for Road Safety Research. Her background is engineering and statistics. Her recent research includes projects related to statistical methods used in road safety.

Ying Ru Feng is a PhD student at Western Australian Centre for Road Safety Research. Her research topic is older drivers using a naturalistic study design.

Nigel Morlet is an ophthalmologist with a special interest in cataract and driving. He has also been involved with several epidemiological studies using linked data with large patient cohorts.

Journal of Safety Research 78 (2021) 96–104

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Cycle Aware: Piloting a module for novice drivers

Jennifer Bonham [a,*], Marilyn Johnson [b,c], Narelle Haworth [d]

[a] School of Social Sciences, University of Adelaide, South Australia, Australia
[b] Institute of Transport Studies, Monash University, Clayton, Victoria, Australia
[c] Amy Gillett Foundation, Melbourne, Victoria, Australia
[d] Centre for Accident Research and Road Safety – Queensland, Queensland University of Technology (QUT), Queensland, Australia

A R T I C L E I N F O

Article history:
Received 13 August 2020
Received in revised form 24 December 2020
Accepted 22 April 2021
Available online 7 May 2021

Keywords:
Novice driver education
Driver license
Cycling
Cyclist safety
Online learning

A B S T R A C T

Introduction: In low-cycling countries, motor-vehicle traffic and driver behavior are well known barriers to the uptake of bicycles, particularly for utility cycling. Lack of separation between cyclists and faster-moving traffic is one key issue, while attitudes of drivers toward and/or harassment of cyclists is another. Cyclist-related driver education has been recommended as a means to improve driver-cyclist interactions. *Methods:* The driver licensing process provides an opportunity for such education. The *Cycle Aware* module was developed to test and enhance novice drivers' knowledge of interacting safely with cyclists. It was piloted across three Australian jurisdictions targeting both novice and experienced drivers. Participants were asked to complete the *Cycle Aware* module and an accompanying survey. A total of 134 novice and 97 experienced drivers completed the survey with 42 novice and 50 experienced drivers going on to complete the module. *Results:* Both groups of drivers scored equally well in the module but the very youngest and very oldest participants were more likely to have some incorrect responses. We did not find any relationship between correct module scores and attitudes toward cyclists. Survey results showed both novice and experienced drivers had somewhat positive attitudes toward cyclists. The two cohorts differed on several attitude questions. Sixty percent (60%) of novices compared to 30% of experienced drivers reported feeling concerned when sharing the road with cyclists, and novices were less likely to agree that cyclists had a right to use the roads. *Conclusions and practical applications:* The analysis suggests novices need to be better equipped to share roads confidently with cyclists and to recognize cyclists as legitimate traffic participants.

1. Introduction

Influencing change to improve bicycling safety from a social, cultural, or attitudinal perspective is difficult. However, for some people, particularly in countries with low levels of cycling participation, negative driver attitudes and harassment of cyclists are key barriers to cycling (Horton, 2007; Heesch, Sahlqvist, & Garrard, 2011). While the benefits of cycling in terms of health (personal, public) and environmental and sustainable transport are well evidenced, (e.g., Andersen, Schnohr, Schroll, & Hein, 2000; Kingham & Tranter, 2015; Patterson et al., 2020), exploiting these benefits and fostering a mode shift toward cycling and other forms of active travel is proving more difficult.

Cyclist-related driver education has been recommended as a means to improve driver-cyclist interactions (Garrard, Greaves, & Ellison, 2010). Researchers in the United States and Australia have developed classroom-based material (Dutt et al., 2018; Bonham & Johnson, 2017, 2018), however the broader driver licensing process is an obvious place to start (Bonham & Johnson, 2018; Bonham, Johnson, & Haworth, 2018).

A review of the driver licensing process in Australia reported only two of the eight jurisdictions (South Australia and the Australian Capital Territory, totaling less than 10% of Australia's population) had compulsory testing of cyclist related road rules and no jurisdictions required novice drivers to demonstrate knowledge of interacting safely with cyclists during a practical test (Bonham et al., 2018). There was limited cyclist-related content in Australian driver licensing resources and that material was inconsistent across jurisdictions with cyclists often represented as problematic road users (Bonham et al., 2018, 2020). The *Cycle Aware* module has been produced to begin addressing this issue.

The *Cycle Aware* module was developed to enhance and test novice driver knowledge of interacting safely with cyclists. The module can be incorporated into the existing requirements for

* Corresponding author at: Department of Geography, Environment & Population, University of Adelaide, Adelaide 5005, South Australia, Australia.

E-mail address: jennifer.bonham@adelaide.edu.au (J. Bonham).

https://doi.org/10.1016/j.jsr.2021.04.003
0022-4375/

novice drivers to pass on their way from pre-learner to fully licensed driver. It includes cyclist-related educational material relevant in all Australian jurisdictions. Further, the imprimatur of government on the module and supporting educational material will confer legitimacy on cyclists as road users.

A key element in ensuring drivers interact safely with cyclists is embedding cycling and safe interactions with cyclists as societal norms. Producing driver education material that normalizes cycling and safe driver behavior, rather than representing cyclists as problems, is a step toward creating that broader societal norm. The *Cycle Aware* module included positive messaging that constituted cyclists as everyday road users and safe interactions as a 'normal' part of driving. The current paper reports on the module pilot and, in particular, the findings of the attitudinal survey that accompanied the pilot. We continue with a review of the existing research into driver attitudes towards cyclists.

2. Literature review

A small but growing body of work examines the overall relationship between drivers and cyclists and, in particular, drivers' attitudes toward and interactions with cyclists (e.g., Basford, Reid, Lester, Thomson, & Tomie, 2002; Rissel, Campbell, Ashley, & Jackson, 2002). In their UK study of drivers' perceptions of cyclists, Basford et al. (2002) reported that during focus group discussions and in-depth interviews, cyclists did not figure among motorists' concerns. This 'invisibility' of cyclists is a concern in itself. More to the point in the current context, it was only when the researchers drew attention to cyclists that motorists acknowledged their presence on the roads and often expressed negative attitudes toward them. Drivers' negative attitudes have been related to cyclists' characteristics and behaviors such as visibility, speed (slow), rule following, and predictability (e.g., Basford et al., 2002; Wood, Lacherez, Marszalek, & King, 2009; Johnson, Oxley, Newstead, & Charlton, 2014; Goddard, Dill, & Monsere, 2016).

Despite some drivers claiming cyclists do not know the road rules, a recent study by Briant, Haworth, and Twisk (2020) reported drivers who do not cycle had poorer knowledge of cyclist-related road rules than cyclists who were also drivers. Rissel et al. (2002) associated drivers' lack of knowledge with drivers' negative views of cyclists. Driver attitudes toward cyclists have also been explained in terms of drivers' participation in cycling. Drivers who regularly ride a bicycle or know someone who rides have more positive attitudes toward cyclists than drivers who do not ride or know anyone who rides (Johnson et al., 2014a; Fruhen & Flin, 2015). Oldmeadow, Povey, Povey, & Critchley (2019) argue drivers who cycle regularly are more likely to identify with cyclists and regard cycling as a normal activity. As noted, these drivers also have better knowledge of cyclist-related road rules and know what cyclists 'should' be doing (and can perhaps anticipate and understand why cyclists don't follow road rules at particular times and places). These 'driver-cyclists'[1] consider cyclists as legitimate road users and, as such, they regard cyclists more positively (Oldmeadow, Povey, Povey, & Critchley, 2019).

A key question is whether negative attitudes about cyclists translate to unsafe driver behavior on the road (e.g., intimidation, harassment). The results are mixed. Basford et al. (2002) reported scant evidence of drivers' negative attitudes toward cyclists being expressed in 'actively hostile' behavior. They argue it is not necessarily attitudes that translate to unsafe behaviors, but these behaviors may be related to drivers' 'perceived behavioral control' in a given situation and/or their perception of broader 'social norms' (Basford et al., 2002). Further, Fruhen and Flin (2015) found drivers who perceived aggressive warnings (e.g., blasting a car horn) toward cyclists as a 'social norm' were more likely to report engaging in this type of behavior (see also Fruhen, Rossen, & Griffin, 2019). This point was highlighted in the naturalistic study of Johnson et al. (2014b). Their study was underway when two well-known comedians made an anti-cyclist rant on a popular commercial television program and captured cyclists' experiences:

> *"The worst night of getting harassed while riding my bike, that I can remember, was the Friday after the Magda Szubanski and Julia Morris anti-cyclist rant... people yelling at me from their cars... [another] car 'jumped' a little, as if the driver had deliberately jumped on the accelerator and then the brake. The driver... was looking at me with a look of pride or aggressiveness or superiority. I associate that night's harassment with the Szubanski-Morris skit because I've never had three separate harassment incidents on one ride."* (p36-37)

Oldmeadow et al. (2019), Fruhen and Flin (2015) and Johnson et al. (2014a) all identify the role of 'social norms' in enabling negative attitudes toward cyclists. The question of how something comes to be constituted as a social norm relates to broader questions of processes of normalization.

Some authors argue drivers with negative attitudes toward cyclists identify them as members of an out- or minority-group and extrapolate the behavior of one cyclist to all members of the group (Basford et al., 2002; Fruhen & Flin, 2015; Prati, Puchades, & Pietrantoni, 2017). Unfortunately, these researchers do not examine the formation of categories, like drivers and cyclists, the mechanisms by which various characteristics and behaviors become attached to these categories, the processes by which identities are 'pressed upon' people or the procedures through which people are encouraged to identify with a particular category (Butler, 1990; Bonham & Bacchi, 2017). Existing research on driver attitudes toward and interactions with cyclists, risks naturalizing these categories and their emergence. Certainly, the categories of driver and cyclist have circulated in contemporary societies for many decades. But institutions, including government departments and research institutes, have lent authority to these categories as they have taken them up in transport, road safety, and public health research, policies and programs, etc.

Importantly, the characteristics and behaviors attributed to cyclists in Australian driver licensing materials are precisely the same as those mentioned by drivers in the research, that is, they are 'hard to see,' slow, unpredictable, and rule breakers (Bonham et al., 2018, 2020). It is possible that drivers' concerns about cyclists are derived in part from the licensing process itself. Alternative characterizations of cyclists might elicit different responses.

In piloting the *Cycle Aware* module, we focused on drivers' knowledge of interacting safely with cyclists and related this to demographic variables, whether they know cyclists or are cyclists themselves and their attitude toward cyclists. The role of the module itself in normalizing cycling/cyclists and developing drivers' appreciation of cyclists as 'just another road user' is a long-term project. Once the module has been rolled out nationally, it will be possible to compare attitudes of novice drivers reported in this study with those who have completed the module as part of their drivers license education and training.

[1] The term 'driver-cyclist' is a variation on 'cyclist-driver' used by Johnson et al (2014b) as a mechanism to break down hard categories of road users. It de-privileges specific modes as it acknowledges people use different modes in different contexts or they may use multiple modes within the same journey (like public transport-cyclists, or public transport-pedestrian).

3. Method: *Cycle Aware* module development and pilot

3.1. Module development

Development of the *Cycle Aware* module was informed by an assessment of online driver education and training materials, novice driver focus-groups, stakeholder interviews, and crash data analysis. The format incorporated Mayer's (2003) principles for interactive multimedia instruction.

Online driver education and training materials: Eleven resources were reviewed from Australia ($n = 7$) and internationally (United Kingdom, $n = 2$; Scotland, $n = 1$; The Netherlands, $n = 1$). Resources were assessed for inclusion of anticipatory driving strategies that did not problematize cyclists as hazards to drivers, rather than legitimate users of the road. The "what happens next" approach (GoSafeGlasgow, Scotland) was identified as a good starting point.

Focus groups were conducted with participants aged from mid-teens to early twenties in South Australia and the Northern Territory. Participants provided feedback on format options (animation, computer generated images (CGI), live-action). Live-action (video) was preferred, considered realistic and appropriate for the seriousness of the topic. Following the focus groups, the "what happens next" approach was adapted to "what will you do?" to encourage a more personalized and self-reflective response.

Stakeholder interviews were conducted with representatives of motoring and cycling organizations, insurance companies, road safety, and transport departments. They provided insight into everyday interactions that motorists and/or cyclists found challenging.

Cyclist crash data were analyzed (Haworth et al., 2019) to identify the most frequent crash types involving novice drivers (age was used as proxy for experience) with most crashes occurring at intersections, on lower speed roads, during the day. This informed the intersection focused situations in the *Cycle Aware* module.

Mayer's principles of interactive multimedia instruction informed the development of the *Cycle Aware* module to ensure learners focused on the instructional content and were not distracted by non-essential information that required extraneous processing and greater cognitive load.

Cycle Aware module format. The *Cycle Aware* module consists of 12 common interactions or situations with cyclists that drivers should expect as usual on the roads. It starts with an introductory video from a cyclist's point of view and includes all 12 situations but in a different sequence. The video is narrated and the voice-over explains the cyclist's behavior and potential issues faced by cyclists in different road environments. Participants then proceed through the 12 situations (Table 1).

Each situation starts with a video establishing the scene and stops at a decisive moment to ask "What will you do?" Multiple choice responses were set out below the video and participants were required to answer before moving on. Each response had an optional voice over for people with reading difficulties. The 'check response' button started a second video and this second section was narrated including the correct response and a key message slide ("Remember").

A progress bar tracked completion and participants could move between the situations in the order generated by researchers, or choose the order. While the module is designed to allow participants to change their responses at any time, this was disabled during the pilot as we were checking initial responses.

3.2. Module pilot

Although targeted at novice drivers, the module was piloted with novice and experienced drivers. In Australia, learner drivers must be accompanied by a supervising driver so experienced driver knowledge of interacting safely with cyclists plays an important role in novice driver development. Research protocols were approved by the University of Adelaide (H-2016-076).

3.2.1. Recruitment

A convenience sampling method was used. A link to the pilot was distributed to novice and experienced drivers through Partner Organization[2] networks and a closed, university student-only Facebook page with 70,000 members. Databases of Partner Organizations indicate that more than 1,000 novices and over 2,000 experienced drivers were invited to participate. The chance of winning a $200 cash prize was offered as an incentive to novice drivers to complete all stages of the pilot (i.e., survey + module and repeat module after 30 days). Experienced drivers were not offered a cash incentive.

3.2.2. Pilot process

The module was piloted in three stages (Fig. 1) and aimed to determine:

- comprehensibility
- overall level of difficulty
- particular situations that cause problems
- level of engagement with the module
- driver characteristics associated with module scores
- the overall efficacy of the module

Stage 1 was designed to determine the comprehensibility of the module. Three novice drivers were observed as they completed the module. Researchers focused on how participants worked their way through the module, in particular: identifying any issues with module instructions; ensuring progression through the module was intuitive; and, determining whether features were obvious (e.g., progress bar). After finishing, participants provided feedback on the module. Comments informed refinements (e.g., clarifying instructions) for the pilot roll-out and for potential longer-term improvements.

Stage 2 targeted novice drivers and was divided into two phases. In *Phase 1,* participants completed a survey that captured data on demographics, driver experience, mode of transport normally used, social networks, advice received on interacting safely with cyclists, and attitudes toward cyclists. Table 2 lists the attitudinal questions used that were adapted from the Bell Dignam scale used by Rissel et al. (2002). A five (5) point Likert scale was used to respond to 10 common statements relating to cyclists. An overall attitude score was developed using nine (9) of the statements (statements 2–10). Each statement was determined to be positive or negative toward cyclists and the Likert response designated accordingly (extreme negative: 9; neutral: 27.5; extreme positive: 45). A Cronbach's alpha test was applied to ensure internal consistency in the weighting of the designation and weighting of responses. Cronbach's alpha for the experienced drivers was 0.819 (based on the 97 valid responses), compared to 0.810 for the novice drivers (based on the 134 valid responses), indicating that the attitude scoring scale has high internal consistency for both groups.

After completing the survey, novices were invited to proceed to the Module. While participants completed the module, backend data were recorded including: responses selected for each multiple-choice question, in/correct responses, time taken to respond. Ideally, participants would have repeated the survey

[2] Partner organizations on the project represented insurers, cyclists, motorists, state and local government authorities. These organizations are listed in the acknowledgements.

Table 1
Cycle Aware Situations (* indicates correct response).

Situation	Multiple response options	Video image and Reminder slide
1. You are walking to the back of the car about to go for a drive. What will you do?	A) Walk to the driver's door, open it and get in. B) Look for vehicles, wait for all vehicles, including cyclists, to pass, then walk to the driver's door, open it and get in.* C) Check for motor vehicles, walk to the driver's door, open it and get in because you don't have to wait for cyclists.	
2. You are driving in the left lane and want to pass a cyclist ahead. What will you do?	A) Stay behind the cyclist until there is enough room to pass safely. B) Pass the cyclist as closely as possible so you don't annoy drivers in the next lane. C) If the next lane is free, change lanes to pass the cyclist safely. D) Either (A) or (C) above.*	
3. A cyclist is traveling ahead of you in the bike lane next to the parked cars. What will you do?	A) Pass the cyclist quickly before any oncoming traffic arrives. B) Move up close behind the cyclist to hurry her along. C) Slow down and stay behind the cyclist until you pass at a safe distance.*	
4. A cyclist is passing a parked vehicle by riding away from the car doors. What will you do?	A) Pass the cyclist quickly before any oncoming traffic arrives. B) Move up close behind the cyclist to hurry them along. C) Slow down and stay behind the cyclist until you pass at a safe distance.*	
5. You would like to turn left at an intersection and there is a cyclist in front riding straight ahead. What will you do?	A) Put on your indicator, accelerate to get ahead of the cyclist as quick as possible and turn left in front of them. B) Put on your indicator, stay behind until the cyclist has cleared the intersection and then turn left.* C) Put on your indicator, and sound your horn so the cyclist knows your there and can wait for you to turn.	
6. You pass a cyclist and want to turn left at the next intersection. What will you do?	A) Head check left, put on the indicator and turn left. B) Put on the indicator and turn left. C) Head check left, put on the indicator, head check again and turn left if you are far enough ahead of the cyclist.*	
7. You are about to enter a roundabout and you see a vehicle on the roundabout about to come off at the next exit. What will you do?	A) Stop completely as it's the law. B) Speed up and enter the roundabout as there is a gap. C) Be prepared, check once for cars and twice for bikes.*	
8. You are approaching a roundabout and there is a cyclist ahead. What will you do?	A) Stay behind and allow the cyclist to 'take the lane' so they can ride safely through the roundabout. B) Rev the car engine to warn the cyclist to hurry up to get out of your way. C) Speed up to beat the cyclist to the roundabout.	
9. You are about to leave a property. What will you do?	A) Enter the road if there aren't any cars coming. B) Check for pedestrians on the footpath and cars on the road. When clear, enter the road. C) Check for pedestrians and cyclists on the footpath, check for all vehicles, including cyclists, on the road. When clear enter the road.*	
10. You are at an intersection and have indicated to turn right. What will you do?	A) Turn as soon as the car has passed. B) Wait until the car has passed and check for any cyclists obscured by the car.* C) Turn in front of the oncoming car.	
11. There are two cyclists riding side by side. You would like to overtake. What will you do?	A) Prepare to slow down and stay behind until you can pass safely.* B) Beep your horn because they should be riding in single file. C) Speed up and pass as quickly and closely as possible so you don't inconvenience other drivers	
12. You have finished your drive and would like to get out of the car. What will you do?	A) Check your side mirror. If you don't see a vehicle approaching, open the door and exit. B) Open the car door and exit the vehicle. C) Check your side mirror then make a head check. If you don't see any vehicles, including cyclists, approaching open the door and exit carefully.*	

attitude statements after completing the module. However, the interface between the module and the external survey made this option unworkable. *Phase 2* tested the efficacy of the module. Participants were asked to repeat the module after 30 days (Palmer & Devitt, 2014) and it was assumed that an improved score would demonstrate the efficacy of the module.

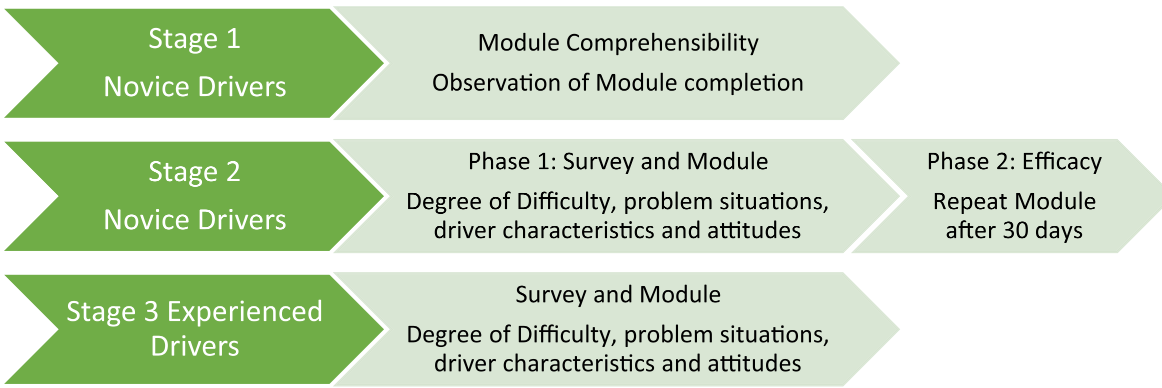

Fig. 1. Stages of the Cycle Aware Pilot.

Table 2
Cyclist Related Statements.

1. I feel worried when sharing the road with cyclists
2. Cyclists should be able to ride on main roads (that don't have bike lanes) at peak hour
3. Most cyclists take notice of road rules
4. Many cyclists on the road have not learned to ride properly
5. Cyclists have as much right to use the road as motorists
6. Drivers are not trained to look out for cyclists
7. Most cyclists are courteous to motorists on the road
8. Cyclists should pay registration fees or road taxes
9. More people cycling means less people driving and that reduces congestion
10. Drivers should change lanes when passing cyclists rather than veering around or squeezing past them

Stage 3 targeted experienced drivers. It followed the process used for the novice drivers but did not ask drivers to repeat the module. The experienced driver survey asked the additional question of whether respondents felt confident in showing others how to interact with cyclists in a range of common crash situations. This information could assist in developing educational resources for both novice and experienced (potentially supervisory) drivers.

3.2.3. Limitations

As noted, we did not repeat the attitudinal questions after the module so we could not determine the influence of the module content on participant attitudes. Further, we did not have a mechanism to capture how the introductory video impacted on module performance – the introductory video provided important information on why cyclists behaved in particular ways and offered clues on how the driver would ideally respond.

4. Findings and discussion

In total, 231 surveys were included in the analysis after 38 surveys were excluded due to incomplete responses. The module was completed by 101 drivers but only 92 drivers completed *both* the survey and the module (Table 3).

4.1. Module analysis

The 42 novice drivers who responded to *both* the survey and module were almost evenly divided between males (20 or 47%) and females (22 or 52%). In contrast, of the 50 experienced drivers who completed *both* the survey and the module, 70% were male, 26% female, and 4% did not state their gender or identified as gender non-specific. The mean ages of respondents were 21.7 years (range 16–36) for novice drivers: and 66.4 years (range 30–86) for experienced drivers.

Over 60% of novice and experienced drivers who completed the module identified the correct responses for all 12 situations and almost 40% had at least one incorrect response (Table 4). High scores (11 or 12 correct answers) were recorded by a higher proportion of female drivers (novice: 91%; experienced: 92%) compared to male drivers (novice: 80%; experienced: 83%). There was some relationship between age and module score at the 95% confidence level. Both the very youngest and the very oldest participants were more likely to have some incorrect responses.

Time taken to complete the module was used as a proxy measure for level of engagement. The entire module ran for 424 s (7 min and 6 s). This included the introductory video from the cyclist's point of view (57 s), watching both videos for all 12 situations and the 'Remember' slide. Experienced drivers were more likely than novice drivers to spend 7 min or more on the entire module (Fig. 2, Fig. 3).

Negotiating a roundabout was the situation with the most frequent incorrect responses for novices and experienced drivers. In both cases, drivers believed it was the law to stop at the roundabout rather than simply being prepared to stop. The other situation with most incorrect responses was cyclists riding side by side where just under 10% in each group believed the cyclists should be riding in single file.

In contrast to findings by Rissel et al. (2002) that drivers with better knowledge of cyclist-related road rules also had more positive attitudes toward cyclists, our analysis did not find any relationship between attitudes and correct module responses. This difference is likely due to differences in what is being tested and how. Much previous research focuses on testing knowledge of cyclist-related road rules using spoken or written prompts. Our module is underpinned by road rules but focuses on best practice in high-crash situations and situations drivers and/or cyclists find particularly challenging. Using video, rather than spoken or written prompts provides rich contextual information.

4.2. Survey analysis: driver attitudes

Echoing other studies (e.g., Johnson et al., 2014a; Fruhen & Flin, 2015), in the survey data analysis both experienced and novice driver attitudes toward cyclists were more positive if the driver used a bike on a regular basis (driver-cyclist) or knew someone who cycled. They were most positive if the bike rider they knew was a family member, friend, or work colleague. There was also a relationship between drivers' attitudes toward cyclists and the person who provided them with advice on sharing the road with cyclists. For both groups, where advice was provided by family

Table 3
Cycle Aware Module and Survey Summary Data.

Survey	Novice drivers (frequency)		Experienced Drivers (Frequency)		Total
Module responses	47		54		101
Module and survey responses	42		50		92
Survey responses	134		97		231
Criteria	**Module and survey responses**	**Survey responses**	**Module and survey responses**	**Survey responses**	
Gender					
Female	22		13		35
Male	20		35		55
Non-specific			1		1
Not stated			1		1
Age					
15–19	13	49		0	62
20–24	19	65		0	84
25–29	7	10		0	17
30–39	3	5	1	4	13
40–49			2	5	7
59–59			10	15	25
60–69			12	25	37
70–79			21	34	55
80+			3	4	7
Not stated		5	1	10	16

Table 4
Correct responses to traffic situations for novice and experienced drivers.

Driver type	Score					Total*
	12/12	11/12	10/12	9/12	8/12	
Novice	29 (62%)	11 (23%)	5 (11%)	1 (2%)	1 (2%)	47 (100%)
Experienced	33 (61%)	13 (24%)	6 (11%)	1 (2%)	1 (2%)	54 (100%)

*Includes ALL module respondents regardless of whether they also completed the survey.

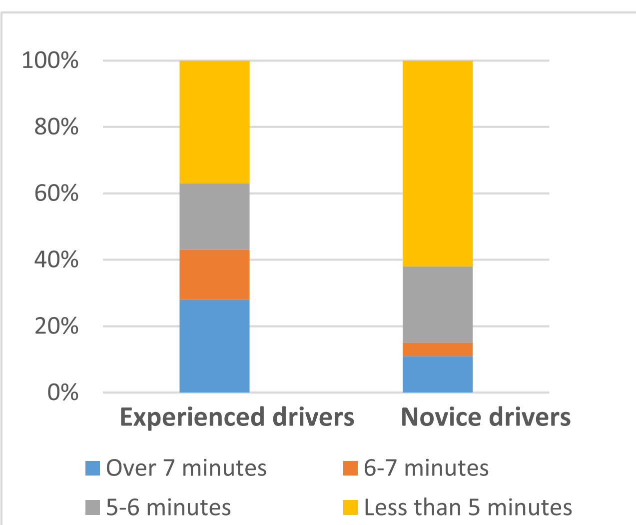

Fig. 2. Time completing module questions only.

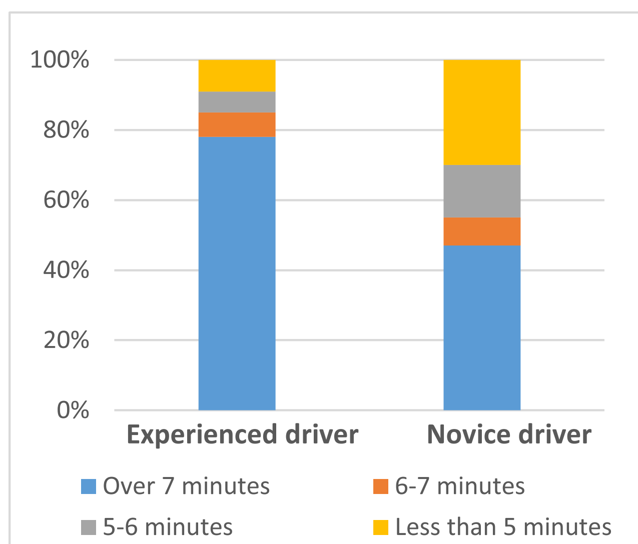

Fig. 3. Time completing entire module.

members, drivers had much more positive attitudes toward cyclists. This suggests that those giving advice are likely to be cyclists so it captures the importance of knowing someone who cycles.

Overall, both cohorts had positive attitudes toward cyclists (i.e., above the neutral score of 27.5) but experienced drivers were slightly more positive (30.05) than novice drivers (29.20) (Table 5). We anticipated novice drivers, as a younger cohort, might be more positive toward cyclists than experienced drivers. This view was based on anecdotal reports from motoring organizations that they receive most complaints about cyclists from experienced drivers.

However, elaborating the point made earlier it seems there is a relationship between where novice drivers receive advice about interacting with cyclists and their attitudes toward cyclists. While 54% reported getting advice from family members, 46% reported getting that advice via driver training guides and websites. It is possible that some of the negative messaging (see Bonham et al., 2018) about cyclists in driver education and training resources has in fact influenced novice driver attitudes toward cyclists. This hypothesis warrants further investigation.

While overall attitudes of novice and experienced drivers toward cyclists were similar, there were substantial differences

Table 5
Novice and Experienced Driver Attitudes to Cyclists.

Driver Type	Range	Median	Mean
Experienced	13–45	29	30.05
Novice	9–44	29	29.20

in some of the detailed responses. Novice drivers felt far less confident than experienced drivers in sharing the road with cyclists (Fig. 4). Disaggregated by mode, both experienced drivers who did not cycle and novice drivers who cycled were concerned about sharing the road with cyclists (see bolded numbers in Supplementary Table A).

In terms of attitude statements, there was considerable difference by driver experience to Statement 2 "Cyclists should be able to ride on main roads (that don't have bike lanes) in peak hour." Responses were divided among novice drivers with only half agreeing with this statement (49%) compared to two thirds of experienced drivers (65%) (Fig. 4). There was also a difference in responses among respondents who also cycled. While most experienced driver-cyclists (94%) agreed with Statement 2, only two thirds of novice driver-cyclists agreed (66%) (see bolded numbers in Supplementary Table).

Novice drivers were also less likely than experienced drivers to agree with Statement 5 that "Cyclists have as much right to use the road as motorists." Half of novice drivers (53%) compared to the majority of experienced drivers (75%) believed cyclists had the same right to use the road as drivers (Fig. 4). The finding for experienced drivers is in line with the 2001 results from Rissel et al. (2002) (75% of drivers agreed). Novice drivers who regularly cycled (64%) were less likely to agree with the statement than experienced drivers who regularly cycled (97%).

The response of novice drivers to Statements 2 and 5 could be related to concerns about sharing the road with cyclists. As they gain experience and confidence in interacting with cyclists, novice drivers may become more accepting of cyclists on the road.

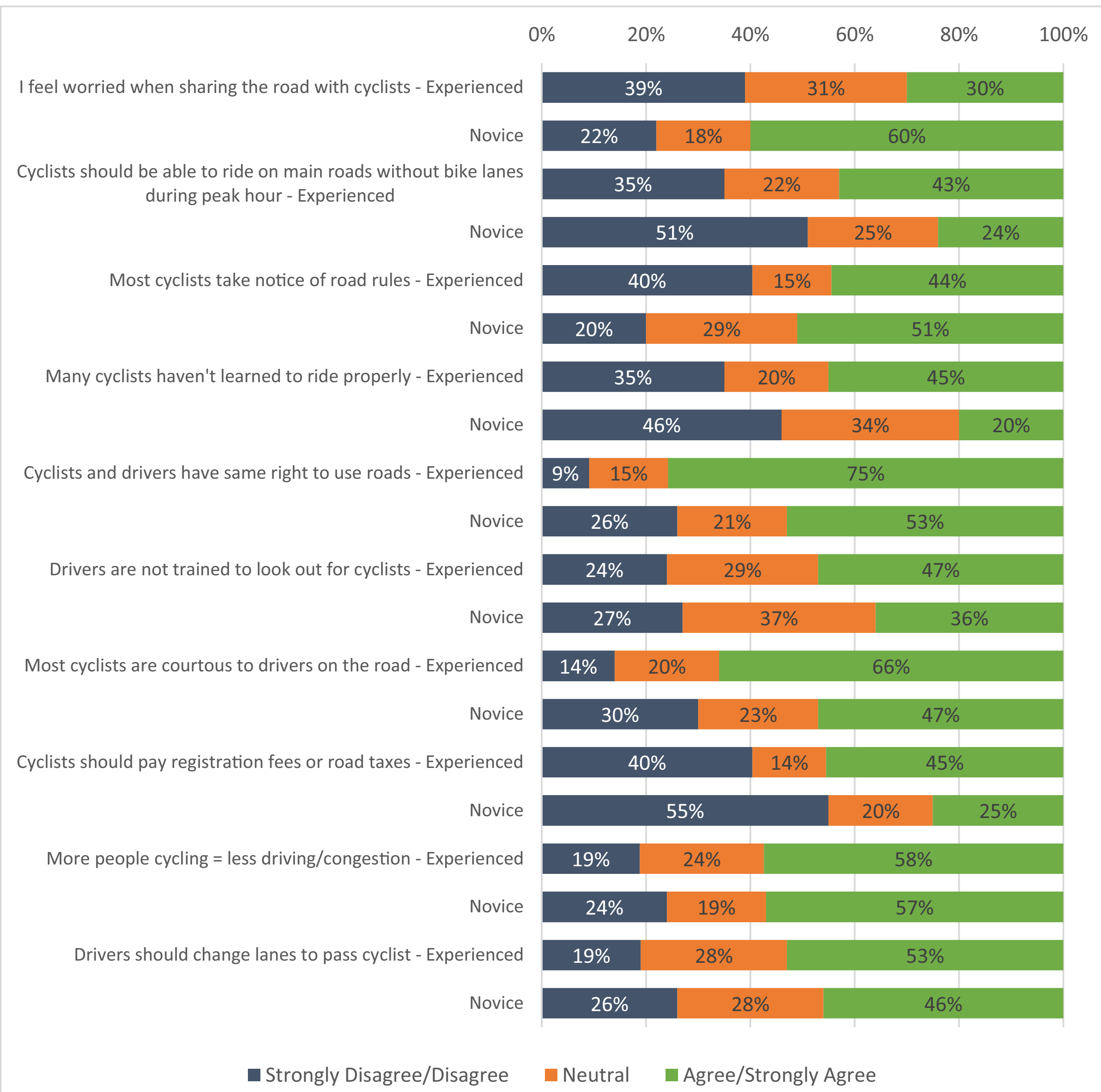

Fig. 4. Responses to attitude statements by drivers – experienced and novice.

Table 6
Average Attitude to Cyclists Score by Driver/Cyclist Status.

Knows someone who cycles	Driver/Cyclist Status			Total
	Drivers who are not cyclists		Driver-cyclists * (Frequency)	
	May know someone who cycles (Frequency)	Uses Public Transport, Shared Modes (Frequency)		
No, doesn't know anyone who cycles	24.07 (14)	28.71 (14)	NA (0)	28
Yes, knows someone who cycles	26.69 (35)	27.84 (90)	33.96 (78)	203
Family member	29.27 (15)	29.03 (29)	35.27 (44)	88
Friend	27.89 (18)	27.64 (50)	34.39 (59)	127
Work colleague	24.00 (4)	29.67 (21)	36.57 (30)	55
Someone in social network	29.70 (10)	28.26 (35)	34.82 (39)	84
Someone else (other)	20.00 (4)	30.60 (10)	35.11 (9)	23

* Drivers who also cycle 'usually, often, sometimes'.

None-the-less, these differences between novice and experienced drivers do raise concerns and suggest the need to improve driver education in relation to interacting with cyclists. In support of this, a high proportion of drivers, except novices who did not cycle, agreed that drivers were not trained to look out for cyclists.

We also tested the relationship between attitudes toward cyclists and mode of transport regularly used. Arguably, people who do not rely exclusively on a car but use several modes, including public transport, might be more accepting or accommodating of other road users. As indicated in Table 6, our results suggest this broader transport experience does influence drivers' attitudes toward cyclists (and possibly the broader range of road users). While the attitude scores for those who use public transport but do not know a cyclist are higher than for those who use public transport and know a cyclist, the difference could be related to the overall number of respondents in each cohort.

5. Conclusion

The *Cycle Aware* module addresses a substantial gap in the driver licensing process in Australia and is a roll-out ready resource to train and test novice drivers about sharing the roads safely with cyclists. By improving or reinforcing novice drivers' knowledge, the module seeks to contribute to the longer-term goal of normalizing cycling/cyclists and developing drivers' appreciation of cyclists as 'just another road user.' Piloting of the module assessed its comprehensibility, degree of difficulty, problem situations, and driver characteristics and attitudes.

Results from the pilot show the very youngest and very oldest drivers were more likely to respond incorrectly to some situations. As a training tool, the module can assist these drivers in developing or refreshing their knowledge. The high level of correct responses recorded during this pilot study is heartening but it requires further investigation. One explanation might be that only people confident or knowledgeable about interacting with cyclists responded to the module, particularly given the information provided in the introductory video from the cyclist's point of view. Alternatively, it could mean the multiple-choice options were too easy. In either case, the module can confirm and reinforce safer interactions with cyclists.

The differences between novice and experienced drivers in the amount of time spent on the module suggests experienced drivers were either more engaged or that novice drivers were more adept at using online resources. The low response rate for the module might indicate drivers are confident in their knowledge of interacting safely with cyclists or drivers lack interest in or concern for interacting safely with cyclists. It seems that without compulsion (e.g., as a requirement of a theory or practical driving test) there is little motivation for drivers to ensure they know safe cyclist-motorist interactions.

The piloting revealed that some novice drivers have already developed negative attitudes towards cyclists. A worrying proportion of novice drivers reported concerns about sharing the road with cyclists, stating cyclists shouldn't be on certain roads at certain times and that cyclists do not have as much right to use the road as motorists. These attitudinal results raise important issues around the current provision of cyclist-related advice to novices. It suggests the need for better education and training on interacting safely with cyclists and debunking claims that help delegitimize cyclists' use of road space. Drivers who received advice from family members appear to have more positive attitudes toward cyclists and this could be related to those family members also riding a bike. Further research is warranted into the relationship between source of advice and attitudes toward cyclists. It seems particularly important that driver education and training resources provide positive advice about interacting with cyclists.

Currently, the focus of driver education and training is on drivers interacting with other motor vehicles. For example, drivers are advised to maintain the speed of the motor vehicles around them. This type of advice encourages novices to view driving as a set of interactions between motor vehicles, rather than focusing on the potential risks that their vehicle can pose to all users of the road (Bonham et al., 2018, 2020). In failing to incorporate the broad range of road users and communicate their point of view there are likely to be ongoing tensions between drivers and cyclists. Understanding the qualities of different road users and the road environment they must navigate is likely to improve safety for all road users.

Declarations of interest

None.

Acknowledgements

This research was funded by the Australian Government through the Australian Research Council (LP150100071) in collaboration with the Motor Accident Commission of South Australia, Amy Gillett Foundation, Australia, Royal Automobile Association of South Australia, Adelaide City Council, Australia, Department of Planning, Transport and Infrastructure (South Australia), Department of Infrastructure, Planning and Logistics (Northern Territory, Australia). The views expressed herein are those of the authors and are not necessarily those of the Australian Government, Australian

Research Council or the partner organizations. We would also like to thank Cathy Jervis and John Davey for assistance with research and data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2021.04.003.

References

Andersen, L., Schnohr, P., Schroll, M., & Hein, H. (2000). All-cause mortality associated with physical activity during leisure time, sports, work, and cycling to work. *Archives of Internal Medicine, 160*(11), 1621–1628.

Basford, L., Reid, S., Lester, T., Thomson, J., & Tomie, A. (2002). Driver's perceptions of cyclists. TRL Report TRL549. London: Department for Transport.

Bonham, J., & Bacchi, C. (2017). Cycling 'subjects' in ongoing-formation: The politics of interviews and interview analysis. *Journal of Sociology, 53*(3), 687–703.

Bonham, J., Johnson, M. & Haworth, N. (Under review). Interacting safely with cyclists: development of an education module for novice drivers.

Bonham, J., & Johnson, M. (2017). *Learning to drive with bikes: Identifying ways to increase safe driving practices*. Clayton, Victoria: Monash Institute of Transport Studies.

Bonham, J., & Johnson, M. (2018). Cyclist-related content in novice driver education and training. *Accident Analysis & Prevention, 111*, 321–327.

Bonham, J., Johnson, M., & Haworth, N. (2018). Cycling related content in the driver licensing process. *Transportation Research Part A, 117*, 117–126.

Bonham, J., Johnson, M., & Haworth, N. (2020). On constituting cyclists as hazards. *Journal of Transport Geography, 84*. https://doi.org/10.1016/j.jtrangeo.2020.102675.

Briant, O., Haworth, N., & Twisk, D. (2020). Driver knowledge of cycling-related road rules in Queensland, Australia. *Traffic Injury Prevention, 21*(7), 464–469. https://doi.org/10.1080/15389588.2020.1793332.

Butler, J. (1990). *Gender Trouble: Feminism and the subversion of identity*. New York, NY: Routledge.

Dutt, K., Pinkos, B., Bopp, M., Doyle, K., Fisher, B., Nelson, A., & Meek, P. (2018). A driver education campaign to increase bicyclist safety and awareness. *The Chronicle for Driver Education Professionals*, 12–16.

Fruhen, L. S., & Flin, R. (2015). Car driver attitudes, perceptions of social norms and aggressive driving behaviour towards cyclists. *Accident Analysis and Prevention, 83*, 162–170.

Fruhen, L. S., Rossen, I., & Griffin, M. (2019). The factors shaping car drivers' attitudes towards cyclist and their impact on behaviour. *Accident Analysis and Prevention, 123*, 235–242.

Garrard, J., Greaves, S., & Ellison, A. (2010). Cycling injuries in Australia: Road safety's blindspot? *Journal of the Australasian College of Road Safety, 21*(3).

Goddard, T., Dill, J., & Monsere, C. (2016). Driver attitudes about bicyclists: Negative evaluations of rule-following and predictability. Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2016.

Haworth, N., Legge, M., Twisk, D., Bonham, J., O'Hare, T. & Johnson, M. (2019). Young driver crashes with cyclists: Identifying training opportunities. Transportation Research Record, at <10.1177/0361198119860118> (viewed 23 August 2019).

Heesch, K. C., Sahlqvist, S., & Garrard, J. (2011). Cyclists' experiences of harassment from motorists: Findings from a survey of cyclists in Queensland, Australia. *Preventive Medicine, 53*, 417–420.

Horton, D. (2007). Fear of cycling. In D. Horton, P. Rosen, & P. Cox (Eds.), *Cycling and society* (pp. 133–152). Hampshire: Ashgate.

Johnson, M., Oxley, J., Newstead, S., & Charlton, J. (2014a). 'Safety in numbers? Investigating Australian driver behaviour, knowledge and attitudes towards cyclists'. *Accident Analysis and Prevention, 70*, 148–154.

Johnson, M., Chong, D., Carroll, J., Katz, R., Oxley, J., & Charlton, J. (2014b). Naturalistic cycling study: identifying risk factors for cyclists in the Australian Capital Territory. MUARC Report No. 322. https://www.monash.edu/__data/assets/pdf_file/0019/217306/Naturalistic-Cycling-Study-Identifying-Risk-Factors-for-Cyclists-in-the-Australian-Capital-Territory.pdf.

Kingham, S., & Tranter, P. (2015). Cycling and sustainable transport. In J. Bonham & M. Johnson (Eds.), *Cycling futures* (pp. 131–152). Adelaide: University of Adelaide Press.

Oldmeadow, J. A., Povey, S., Povey, A., & Critchley, C. (2019). Driver anger towards cyclists in Australia: Investigating the role of the perceived legitimacy of cyclists as road users. *Transportation Research Part F: Psychology and Behaviour, 63*, 240–251.

Palmer, E., & Devitt, P. (2014). The of a stassessment ructured online formative assessment program: A randomised controlled trial. *BMC Medical Education, 14*, 8. https://doi.org/10.1186/1472-6920-14-8.

Patterson, R., Panter, J., Vamos, E. P., Cummins, S., Millett, C., & Laverty, A. A. (2020). Associations between commute mode and cardiovascular disease, cancer, and all-cause mortality, and cancer incidence, using linked Census data over 25 years in England and Wales: A cohort study. *Lancet Planet Health, 4*(5), e186–e194.

Prati, G., Marín Puchades, V., & Pietrantoni, L. (2017). Cyclists as a minority group. *Transportation Research Part F: Traffic Psychology and Behaviour, 47*, 34–41.

Rissel, C., Campbell, F., Ashley, B., & Jackson, L. (2002). Driver road rule knowledge and attitudes toward cyclists. *Australian Journal of Primary Health, 8*(2), 66–69.

Wood, J., Lacherez, P., Marszalek, R., & King, M. (2009). Drivers' and cyclists' experiences of sharing the road: Incidents, attitudes and perceptions of visibility. *Accident Analysis and Prevention, 41*, 772–776.

Dr Jennifer Bonham is a Senior Lecturer specializing in urbanization and cultural practices of travel. Jennifer has 25 years of research experience examining road user identities. Since 2005 her work has focused on cycling. She has published on intra-urban differences in cycling, the role of cycling in dispersed cities, gendering of cycling and the representation of cycling in public awareness safety campaigns.

Dr Marilyn Johnson is a Senior Research Fellow with the Institute of Transport Studies, in the Department of Civil Engineering. Marilyn is also the Research and Policy Manager at the Amy Gillett Foundation. Marilyn's research expertise is in cycling including cyclist safety, cyclist-driver interactions and electric bike use in Australia.

Professor Narelle Haworth works at CARRS-Q, the Centre for Accident Research and Road Safety-Queensland and has 30 years' experience in road safety research. Her focus has been on improving the safety of vulnerable road users – pedestrians, bicyclists and motorcyclists. She has also stressed the importance of the interaction between safety and environmental and other health outcomes for these groups. A major thrust of her work has been to make policymakers aware of the relevance and impact of research on road safety practice, and to reinforce the need to focus on the most effective measures, notably speed reductions and road infrastructure as safety improvements. Professor Haworth has led evaluations of driver-education programs in schools and motorcycle rider training.

Journal of Safety Research 78 (2021) 180–188

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Development of pedestrian- and vehicle-related safety performance functions using Bayesian bivariate hierarchical models with mode-specific covariates

Mankirat Singh [a], Wen Cheng [a,*], Dean Samuelson [b], Jerry Kwong [c], Bengang Li [a], Menglu Cao [a], Yihua Li [d]

[a] Department of Civil Engineering, California State Polytechnic University, Pomona, CA 91768, United States
[b] Traffic Safety Investigations Branch, Department of Transportation California, United States
[c] Division of Research, Innovation and System Information, Department of Transportation California, United States
[d] Department of Logistics Engineering, Logistics and Traffic College, Central South University of Forestry and Technology, Hunan 410004, China

ARTICLE INFO

Article history:
Received 20 October 2020
Received in revised form 15 February 2021
Accepted 21 May 2021
Available online 4 June 2021

Keywords:
Pedestrian-Vehicles crashes
Crash frequency Models
Bivariate models
Pedestrian count
Safety Performance Function

ABSTRACT

Introduction: Pedestrian safety is a major concern as traffic crashes are the leading cause of fatalities and injuries for commuters. Traffic safety research in the past has developed various strategies to counteract traffic crashes, including the safety performance function (SPF). However, there is still a need for research dedicated to enhancing the SPF for pedestrians from perspectives of methodological framework and data input. To fill this gap, this study aims to add to the current SPF development practice literature by focusing on pedestrian-involved collisions, while considering the typical vehicle ones as well. *Methods:* First, bivariate models are used to account for the common unobserved heterogeneity shared by the pedestrian- and vehicle-related crashes at the same intersections. Second, variable importance ranking technique is used, along with correlation analysis, to determine mode-specific feature input. Third, the exposure information for both modes, annual pedestrian count, and annual daily vehicles traveled are used for model development. Fourth, a recent Bayesian inference approach (integrated nested Laplace approximation (INLA)) was adopted for bivariate setting. Finally, different evaluation criteria are used to facilitate comprehensive model assessment. *Results:* The results reveal different statistically significant factors contributing to each of the modes. The offset intersection provides better safety performance for both pedestrians and drivers as compared to other intersection designs. The model findings also corroborate the sensibility of using the bivariate models, rather than the separate univariate ones. *Practical Applications*: The study shows that pedestrians are more vulnerable to various intersection features such as left-turn channelization, intersection control, urban and rural population group, presence of signal mastarm on the cross-street, and mainline average daily traffic. Greater focus should be directed toward such intersection features to improve pedestrian safety.

1. Introduction

Walking is considered an important travel mode due to its immense benefits. However, the percentage of total trips undertaken by pedestrians compared to other modes is very low. In the United States, the National Household Travel Survey (NHTS, 2018) reported that trips made by walking accounted for only 0.6% of total person-miles travel (PMT). Past studies reveal that there are many reasons for this, but one main reason is that pedestrians are among the most vulnerable and unsafe road users from the viewpoint of traffic crashes (De Hartog et al., 2010; Retting et al., 2003). According to national statistics (Governor Highway Safety Association, 2019), a total of 6,590 pedestrian fatalities and around 70,000 injuries were estimated in pedestrian-vehicle crashes in 2019. These consequences create an urgent need to prevent pedestrian-involved crashes by implementing better policies and strategies to provide a safe traffic environment for pedestrians.

Given this context, many researchers have investigated various factors, such as roadway-built characteristics (Mansfield et al., 2018; Miranda-Moreno et al., 2011), pedestrian behavior (Xu et al., 2018; Dommes et al., 2014), driver behavior (Baker et al., 1974; Geruschat & Hassan, 2005; Schroeder & Rouphail, 2011), traffic characteristics (Shi et al., 2007; Barton & Morrongiello, 2011), drug/alcohol use (Li et al., 2019; Plurad et al., 2006), social

* Corresponding author.
E-mail address: wcheng@cpp.edu (W. Cheng).

https://doi.org/10.1016/j.jsr.2021.05.008
0022-4375/© 2021 The Authors. Published by the National Safety Council and Elsevier Ltd.

and demographic attributes (LaScala et al., 2000; Tabibi et al., 2012) to highlight the crucial insights related to pedestrian-involved crashes. Among the distinct strategies, the development of safety performance functions (i.e., crash frequency models) is one of the most popular strategies to address traffic safety that not only aids in screening out the significant influential factors, but also predicts the crash counts for various purposes (Ukkusuri et al., 2011; Wu et al., 2018; Harwood et al., 2008). Poisson regression models were initially widely adopted due to the popularity of Poisson distribution for discrete outcomes (Miranda-Moreno, 2006). However, Poisson regression models are not able to provide reliable and unbiased results in the case of over-dispersion (i.e., variance greater than the associated mean). In response, researchers employed alternate model formulations such as Poisson gamma or negative binomial (Hauer, 2001), Poisson lognormal (Park & Lord, 2007), zero-inflated models (Aguero-Valverde, 2013), which can better address over-dispersion issues and, hence, provide more valid inferences.

Among the above-mentioned models, the univariate model framework has seen widespread applications in traffic safety studies due to its ease of implementation with only one dependent variable being involved (Anarkooli et al., 2019). However, the univariate model is not capable of addressing the unobserved heterogeneities shared by various crash types or severities occurring in the same locations or situations (Mannering & Bhat, 2014). To overcome this issue, multivariate models have been proposed due to their enhanced capabilities to tackle the common heterogeneity among different crash types via the explicit consideration of correlated random effects (Lee et al., 2015; Park & Lord, 2007). As a special case of multivariate setting, the bivariate model is dedicated to the crashes of two categories and also enjoys frequent applications. For example, Russo et al. (2014) used bivariate framework to examine the factors pertaining to crash injury severities involved in angled collisions. The results demonstrate that bivariate models provide more insightful findings related to the factors influencing the propensity of crashes. Zheng and Sayed (2019) developed bivariate models to integrate the traffic conflict indicators for crash estimation. The finding showed that the bivariate model improved the crash estimation precision and accuracy.

Another dimension of model classification resides in the transportation modes. For instance, as previously stated, the SPF can be divided into motor vehicle-oriented, non-motorist-centered, and so on. Overall, the vehicle-related SPF dominates the SPF development given the largest proportion of such mode in the current transportation system. However, with the constant promotion of active transportation provided by various levels of government agencies in the past decades, ever-increasing interests were directed toward the SPF development for pedestrians (Wier et al., 2009; Rasciute & Downward, 2010). For example, Thomas and DeRobertis (2013) investigated road crashes involving pedestrians and bicyclists based on a very short period of volume count data available for these two modes. The results showed that the risk for an individual pedestrian/bicyclist to be involved in a crash decreases with an increase in the number of pedestrians/bicyclists. Subsequently, McArthur et al. (2014) conducted a study to develop SPF to estimate the pedestrian crashes using five-year data, including socioeconomic and demographic characteristics. Another study done by Tulu et al. (2015) developed SPF to investigate the pedestrian crashes on two-way two-lane rural roads by incorporating the short period counts of daily pedestrian crossing volumes in Ethiopia. Gates et al. (2016) developed SPFs for pedestrian and bicyclist crashes at road segments and intersections. The results demonstrated that pedestrian and bicycle crashes tend to increase when vehicle traffic volumes increase investigated two pedestrian crash types (total pedestrian crashes at intersections and a subset of intersection crashes involving pedestrian-motorists collision), by developing SPFs and incorporating many factors such as roadway, built environment, census, and activity measures. A common theme among these papers is the lack of or very limited exposure information directly related to active transportation modes such as pedestrian counts. The potential reasons for such data scarcity are due to lack of definite paths or routes, followed by pedestrian and bicyclists, limited use of emerging technologies (e.g., crowd-sourcing), expensive data collection devices, and so on. To address these issues, different strategies have been used in the past. Some studies employed daily vehicle miles traveled as the proxy for active transportation modes based on the assumption that most non-motorist-pertinent collisions are related to vehicles (Cheng et al., 2018). Others relied on the formulation of the pedestrian volume models using predictors such as land use, transportation system attributes, and neighborhood socioeconomic characteristics. One recent example is the pedestrian count model developed by UC Berkeley researchers (Griswold et al., 2019), which can be used in the pedestrian SPF as a major estimation of pedestrian exposure, rather than other proxy information.

Building upon previous research, this paper aims to add to the current literature with microscopic level research pertaining to pedestrian-involved collisions, with some enhancements to current SPF development practice. First, bivariate models are used to account for the common unobserved heterogeneity shared by the pedestrian and vehicle-related crashes at the same intersections. Second, random forest (RF) method and Pearson's correlation test are employed to determine the variables to be included in the models, which are different for each of the joint models. Such practice leads to the proper covariates not only striking a balance between multi-collinearity and omitted variable bias issues, but also enhancing model flexibility with different inputs to specific transportation mode. Third, the exposure information for both modes, annual pedestrian count, and annual daily vehicles traveled is used for the model development. The former counts were generated based on research by Griswold et al. (2019). Including exposure of both pedestrians and vehicles are anticipated to enhance model estimate precision. Fourth, in comparison with the typical Bayesian hierarchical models based on the Markov chain Monte Carlo (MCMC) algorithm, the integrated nested Laplace approximation (INLA) approach is selected due to faster calculation and more robust results (Taylor & Diggle, 2014). Finally, for a comprehensive comparison of the predictive accuracy of the models, distinct goodness-of-fit measurements that include DIC (deviance information criterion), $\overline{D}$ (posterior mean deviance), P_D (effective number of parameters) and LPML (log pseudo marginal likelihoods) are employed. The results of this study are expected to provide additional insights into SPF development, especially for the pedestrian mode.

2. Materials and methods

The paper features Bayesian joint hierarchical models with different inputs of covariates, variable selection using both random forest (RF) for predictor importance ranking and Pearson's correlation test for variable correlation analysis, and a set of performance evaluation criteria. The following subsections cover the corresponding methodological details in order.

2.1. Bayesian joint hierarchical model specification

The paper employed Poisson lognormal model, which assumes the crash count to be Poisson distributed with the logarithm of Poisson rate following normal distribution. The model formulation is shown as follows (Cheng et al., 2018).

$$\boldsymbol{y}\ Poisson(\lambda) \tag{1}$$

$$\ln(\boldsymbol{\lambda}) = \boldsymbol{\beta}_0 + \boldsymbol{\beta X} + \boldsymbol{\varepsilon} \tag{2}$$

where $\boldsymbol{y}$ is a matrix consisting of crash counts of different modes and intersections, $\boldsymbol{\lambda}$ is a matrix consisting of the corresponding Poisson rates of different modes and intersections, $\boldsymbol{\beta_0}$ represents a global intercept vector for the two modes, $\boldsymbol{\beta}$ is a coefficient vector, $\boldsymbol{X}$ is the covariate matrix, and $\boldsymbol{\varepsilon}$ represents the white noise matrix.

To better describe the joint models with different predictor input, let $\lambda_{\mathbf{p}}$ and $\lambda_{\mathbf{v}}$ denote the pedestrian- and vehicle-involved Poisson rate vector, respectively. The model framework for each of the transportation modes can be expressed using the following equation.

$$\ln(\lambda_{\mathbf{p}}) = \beta_{0p} + \beta_{\mathbf{cp}}\mathbf{X_c} + \beta_{\mathbf{dp}}\mathbf{X_{dp}} + \boldsymbol{\varepsilon_p} \tag{3}$$

$$\ln(\lambda_{\mathbf{v}}) = \beta_{0v} + \beta_{\mathbf{cv}}\mathbf{X_c} + \beta_{\mathbf{dv}}\mathbf{X_{dv}} + \boldsymbol{\varepsilon_v} \tag{4}$$

where the subscripts v and p represent the vehicle and pedestrian modes, β_0 is the global intercept, $\beta_{\mathbf{c}}$ is the vector of coefficients for the independent variables common to both modes, $\mathbf{X_c}$ is the matrix of covariates common to both modes, $\beta_{\mathbf{d}}$ is the vector of coefficients for the independent variables that are different between the two modes, $\mathbf{X_d}$ is the corresponding covariate matrix, and $\boldsymbol{\varepsilon}$ is the vector of error term. The two models are developed simultaneously with the two error vectors, $\boldsymbol{\varepsilon_p}$ and $\boldsymbol{\varepsilon_v}$, following the bivariate normal distribution:

$$\boldsymbol{\varepsilon}\ MN(\boldsymbol{\mu}, \textstyle\sum) \tag{5}$$

$$\text{where} : \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_p \\ \boldsymbol{\varepsilon}_v \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \mu_p \\ \mu_v \end{pmatrix}, \sum = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \tag{6}$$

In the above equations, MN represents multivariate (or, bivariate in the present study) normal distribution, $\boldsymbol{\varepsilon}$ is the random effect matrix that captures the extra-Poisson heterogeneity among intersections,$\boldsymbol{\mu}$ is the vector of the mean values for the bivariate normal distribution, and $\sum$ is the variance–covariance matrix where the diagonal elements (i.e., σ_{11} and σ_{22}) in the matrix represent the variances of the random effects, while the off-diagonal element represents the covariance. The inverse of the variance–covariance matrix represents the precision matrix, which can be formulated using the Wishart distribution:

$$\textstyle\sum^{-1}\ Wishart(I,J) \tag{7}$$

where I is the J identity matrix (Congdon, 2006), and J is the degree of freedom, J = 2 herein representing two transportation modes. The non-informative specifications (Heydari et al., 2017) for various coefficients were specified with a normally distributed vague priors N (0,100). Such diffused normal distribution with zero mean and a large variance is commonly employed as a vague prior to posterior estimates due to the absence of sufficient knowledge of priori distribution (Cheng et al., 2018).

2.2. Random forest (RF), and various evaluation criteria

Decision tree is one of the predictive models that come up with an item's target value (leaves) via the observations about the item (branches). Based on the nature of the target value, the decision tree model can be used for both regression and classification purposes. Compared with the other typical regression techniques, the decision tree gained its popularity as it closely mirrors the human decision-making process. As implied by the name, random forest (RF) (Fatholahzade et al., 2018) consists of a collection of individual decision trees that operate as an ensemble. The method combines bagging and the random selection of features to construct different decision trees with a controlled variation. Using ensembles of predictors have proven to give more accurate results than using a single predictor. This technique has an advantage over the traditional decision trees in obtaining unbiased error estimates without separating the cross-validation test dataset. When a particular tree in the RF is grown from a bootstrap sample, usually one-third of the training cases are left out (also called out-of-bag, OOB, data) from the tree-growing. The OOB data are then used later for the determination of the optimum number of predictors for each tree and the optimum number of trees in the RF, which results in the minimum OOB error rate.

As a robust data mining technique with the implicit wisdom: "a large number of uncorrelated individuals operating as a committee will make a better decision than do these individuals," RF has seen wide applications in various fields including traffic safety (Harb et al., 2009). For the classical purposes of regression and classification, RF has been frequently used for determining the importance of various response variables, based on the mean decrease of either prediction accuracy or node purity with a specific variable being excluded from the model (Jiang et al., 2016). The multiple steps are involved when the former metric is implemented. First, the prediction accuracy of the OOB sample is estimated. Second, the values of the variable in the OOB sample are randomly shuffled, with all other variables remaining the same. Third, the decreased prediction accuracy on the shuffled data is calculated. Finally, the average drop of accuracy across all trees is reported for the variable. The more decreased prediction accuracy in the OOB data, the more predictive power the variable tends to have. The second method follows a similar process as does the first one, except that the prediction accuracy is replaced with the node purity (or, Gini), which has the largest value when only one single class or value is involved in the node. Compared with the predictive accuracy-oriented metric, the node purity one has the advantage of faster computation and is, therefore, chosen in the study (Nicodemus, 2011).

In addition to the variable importance ranking, this paper adopted various evaluation criteria that include Deviance Information Criterion (DIC), $\overline{D}$ (posterior mean deviance), P_D (effective number of parameters) and Log pseudo marginal likelihoods (LPML) for the assessment of predictive accuracy and goodness-of-fit.

As a hierarchical modeling generalization of the Akaike Information Criterion (AIC) which uses maximum likelihood estimates (Hurvich & Tsai, 1998), DIC has been used extensively to assess the complexity and goodness of fit of the Bayesian models based on the posterior mean. The calculation of DIC can be done via the following expression (Spiegelhalter et al., 2003):

$$\overline{\text{DIC} = D} + P_D \tag{8}$$

where, $\overline{D}$is the posterior mean deviance that measures the closeness of the fitted data to the original observations, and P_D denotes the effective number of parameters in a model representing the model complexity. In general, models with more parameters tend to overfit the data, resulting in smaller deviance. Therefore, the P_D term can be considered as compensation for this effect by favoring models with a smaller number of parameters.

Different than DIC and $\overline{D}$, which are based on within-sample predictive errors, other alternatives are based on the test data using cross-validation techniques. Nonetheless, the typical approaches of cross-validation are prone to selection bias related with data-splitting into subsets. To circumvent such bias, a robust conditional predictive ordinate (CPO) based on CV-1 (leave-one-out) was employed in this study (Pettit, 1990). Within the INLA framework, the estimate of CPO for each observation i can be calculated as (Gelfand, 1996; Liu & Sharma, 2017):

$$CPO_i = \left(\frac{1}{T} \sum_{t=1}^{T} \frac{1}{f\left(Y_i|\beta^{(t)}\right)} \right)^{-1} \quad (9)$$

where Y_i is the i^{th} observation (i = 1, 2, 3, n) for all intersections and β represents the estimated model parameters.

Based on the CPO, the Log pseudo marginal likelihoods (LPML) can be calculated and have been employed in recent safety literature (Heydari et al., 2017; Cheng et al., 2018a). The computation for LPML can be performed using the following equation:

$$\mathrm{LPML} = \sum_{i=1}^{n} log(CPO_i) \quad (10)$$

where i, n, and CPO are denoted in Equation (9).

The large value of LPML signifies a better predictive capability related to the candidate model.

2.3. Data description

The analysis in this study was based on data derived from California Traffic Accident Surveillance and Analysis System (TASAS). TASAS is a traffic records system that includes a crash database and infrastructure database consisting of highway segments, intersections, ramps, and other data. The study focused on crashes occurring at the intersections that have 73 variables available in the raw file in TASAS. Nonetheless, some of these variables were not associated with pedestrian or vehicle collisions like intersection location information (district, county, route, and milepost), date of intersection update (begin date of intersection update, the end date of intersection update), and so on. After data cleaning, 20 covariate variables were selected from a total of 6,198 intersections in the state routes, where the estimated annual pedestrian volume at each intersection was available through the pedestrian count model developed by Griswold et al. (2019). Overall, a total

Table 1
Descriptive Statistics of Collected Data.

Numerical Variables					
Variables	Description	Minimum	Maximum	Mean	S.D.
MNL	Mainline - number of lanes	2	8	3.33	1.39
MOL	Mainline - override length (buffer)	15	350	187.90	61.94
X-NL	Cross street - number of lanes	0	6	2.13	0.55
X-OL	Cross street - override length	0	250	2.25	22.81
MADT	Mainline - average daily traffic	180	125,000	20,198	15254.28
X-ADT	Cross street - average daily traffic	0	77,000	1,911	4272.05
IRG	Intersection rate group	1	29	17.91	7.61
APV	Estimated annual pedestrian volume (2016)	520	9,400,000	116,636	481942.50
Veh counts	Vehicle related accidents counts	0	137	6.88	11.79
Ped counts	Pedestrian related accidents count	0	6	0.09	0.39

Categorical Variables		
Variables	**Description**	**Details of categories (frequency, percentage)**
Highway Group	Highway group of mainline in the intersection	1-Divided Highway (3,294, 53.15%); 2-Undivided Highway (2,881, 46.48%); 3-Right or Left Independent Alignment (23, 0.37%)
Population Group	Population code of the intersection	-Urban (1,539, 24.83%); 2-Rural (1,278, 20.62%); 3-Urbanized (3,381, 54.55%)
Intersection Design	Intersection design	1-Four legged (2,328, 37.56%); 2- >Four legs (67, 1.08%); 3-Offset (349, 5.63%); 4- Tee (3,182, 51.34%); 5 - Wye (206, 3.32%); 6-Other (66, 1.06%)
Light Condition	Presence of light condition at Intersection	1-No Lighting (1,561, 25.19%); 2-Lighted (4,637, 74.81%)
Mastarm	Presence of signal mastarm on the mainline of the intersection	1-No Mastarm (4,901, 79,07%); 2-Yes, Mastarm (1,297, 20.93%)
Left Turn	Left turn channelization on mainline at the intersection	1-Curbed Median Left Turn Channelization (808, 13.04%); 2-No Left Turn Channelization (3,005, 48.48%); 3 - Painted Left Turn Channelization (2,355, 37.00%); 4 - Others (30, 0.48%)
Right Turn	Right turn channelization on mainline at the intersection	1-No Right Turn Channelization (5,579, 90.01%); 2-Others (617, 9.99%)
Traffic Flow	Traffic flow on the mainline of the intersection	1-Two-Way Traffic, No Left Turns Permitted (297, 4.81%); 2-Two-Way Traffic, Left Turn Permitted (5,839, 94.21%); 3 - Others (61, 0.98%)
X-Mastarm	Presence of signal mastarm on the cross-street of the intersection	1-No Mastarm (5,341, 86.17 %); 2-Yes, Mastarm (857, 13.83%)
X-Left Turn	Left turn channelization on the cross-street	1-Curbed Median Left Turn Channelization (131, 2.11%); 2-No Left Turn Channelization (5,421, 87.47%); 3-Painted Left Turn Channelization (622, 10.04%); 4-Others (24, 0.39%)
X-Right Turn	Right turn channelization on the cross-street.	1-No Right Turn Channelization (5,631, 90.85%); 2-Others (567, 9.15%)
X-Traffic Flow	Traffic flow on the cross-street of the intersection	1–Two Way Traffic, No Left Turns Permitted (269, 4.34%); 2–Two-Way Traffic, Left Turn Permitted (5,846, 94.32%); 3-Others (83, 01.34%)
Intersection Control Condition	Intersection control condition	1-No Control (210, 3,039%); 2-Stop signs on Cross Street Only (4,514, 72.83%); 3-Signals Pretimed (2 Phase) (152, 2.45%); 4- Signals Semi-Traffic Actuated, Two-phase (125, 2.02%); 5 - Signals Full Traffic Actuated, Multi-Phase (993, 16.02%); 6-Others (204, 3.29%)

Note. S.D. represents standard deviation.

of 43,705 pedestrian and vehicle collisions spanning over six years (2012 to 2017) were aggregated for research purposes. The detailed information for all data including variable names, description, and other descriptive statistics are illustrated in Table 1.

3. Results

To develop the bivariate SPFs, the distinct covariates were selected for pedestrians and vehicles by using RF metric and correlation analysis. Under the INLA framework, models were developed with the posterior mean serving as the estimate for the model parameters. Different evaluation criteria were used to assess the predictive accuracy of the models.

3.1. Feature importance ranking by random forest (RF)

The importance of variables was reported and ranked using the R package "randomforest" (Cutler et al., 2012). When estimating the RF model, m = 4 variables were randomly sampled as a candidate at each split, with the OOB error rate reaching a minimum value of 0.132 and 59.24% of data variability being explained by the model. The variable importance plots for both pedestrian and vehicles are shown in Fig. 1 with the decreasing order of "IncNodePurity," which represents the mean decrease of node purity in predictions on OOB samples with a given variable being excluded from the model.

3.2. Correlation analysis

Variable importance ranking was used along with the correlation of numerical variables for the determination of the covariate inputs to the model development. The correlation tests were conducted using the Harrell Miscellaneous package in R software, which allowed the calculation of Pearson's correlation coefficient and the accompanying p-values. The variables were observed to be correlated by using the popular cut line of 0.6 for the correlation coefficient and with a significance level of 0.05 were eliminated in multiple steps using engineering judgment to choose the minimum subset of variables, while maintaining the maximum data variability. In other words, the selection procedure strived to strike the balance between omitted variable bias and multi-collinearity issues. As shown in Table 2, the upper portion values are Pearson's correlation coefficient magnitudes and lower shaded cells represent the associated p-values. Based on the results of the correlation test, out of eight numerical variables, six of them that include MOL, X-OL, MADT, X-ADT, IRG, and APV were retained.

Combining both results from RF and correlation analysis, the final list of predictors to be included into subsequent model development can be found in Table 3. It is important to note that the variables "Right Turn" and "X-OL" were retained only for pedestrians, while "X-Mastarm" and "X-Right Turn" were included only for vehicles. They were not considered for both modes at the same time since they had little influence on one of the modes according to the variable importance ranking results via RF.

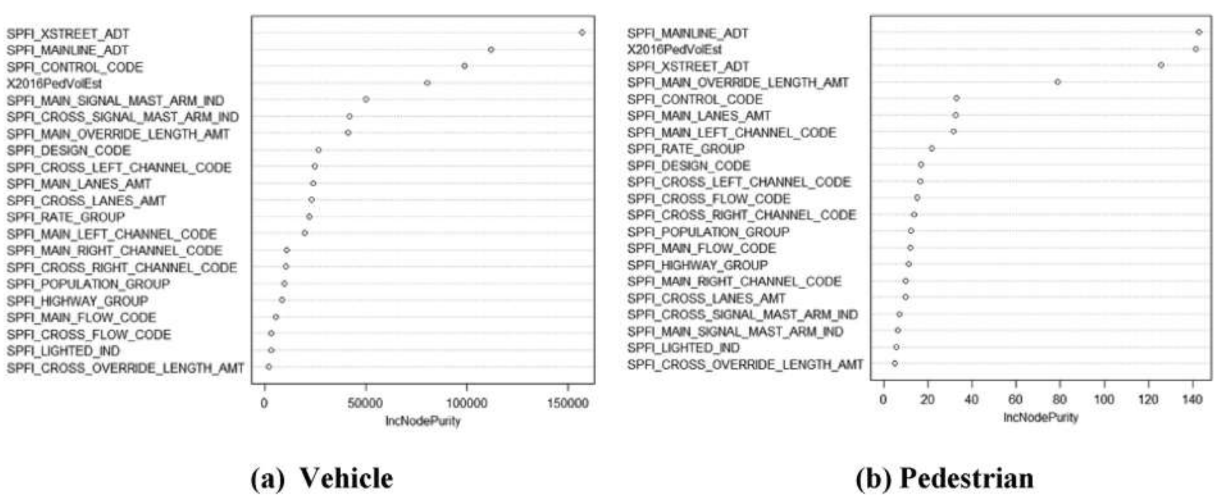

Fig. 1. A Variable Importance Plot for (a) Vehicle Crash Counts and (b) Pedestrian Crash Counts.

Table 2
Correlation Coefficients and P-Value for the Numerical Variables.

	MNL	MOL	X-NL	X-OL	MADT	X-ADT	IRG	APV
MNL	1.000	−0.116	0.230	0.001	**0.722**	0.273	0.118	0.254
MOL	0.000	1.000	0.046	−0.001	−0.088	0.032	−0.139	−0.152
X-NL	0.000	0.000	1.000	0.118	0.227	**0.600**	−0.087	0.112
X-OL	0.925	0.917	0.000	1.000	0.005	0.187	−0.034	0.194
MADT	0.000	0.000	0.000	0.683	1.000	0.269	0.154	0.257
X-ADT	0.000	0.014	0.000	0.000	0.000	1.000	−0.118	0.184
IRG	0.000	0.000	0.000	0.008	0.000	0.000	1.000	−0.046
APV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Notes.
1. The upper triangle of the matrix shows the correlation coefficients of the variables, and the gray grids in the lower triangle of the matrix shows the p-values.
2. Highly correlated estimate with correlation coefficient greater than 0.6 are marked as bold font.
3. Refer to Table 1 for details of variable definition.

Table 3
Description of Model Parameter Estimates.

Variables		β1 (Pedestrian)		β2 (Vehicle)	
		Mean	SD	Mean	SD
		Fixed Effects			
(Intercept)		**−5.222**	**0.772**	**3.788**	**0.735**
Highway Group	Highway Group 1 (Base)				
	Highway Group 2	−0.008	0.053	−0.082	0.141
	Highway Group 3	0.494	0.37	−0.941	0.831
Population Group	Population Group 1 (Base)				
	Population Group 2	**−0.834**	**0.092**	−0.38	0.321
	Population Group 3	**0.247**	**0.051**	0.088	0.141
Intersection Design	Intersection Design 1 (Base)				
	Intersection Design 2	−0.215	0.176	−0.34	0.448
	Intersection Design 3	**−0.341**	**0.085**	**−0.401**	**0.21**
	Intersection Design 4	**0.391**	**0.157**	**−1.761**	**0.485**
	Intersection Design 5	**0.485**	**0.189**	**−3.634**	**1.107**
	Intersection Design 6	**0.681**	**0.246**	**−1.994**	**0.779**
Light Condition	Light Condition (1) (Base)				
	Light Condition (2)	**0.307**	**0.054**	**0.747**	**0.213**
Mastarm	Mastarm 1 (Base)				
	Mastarm 2	−0.207	0.131	−0.087	0.271
Left Turn	Left Turn 1 (Base)				
	Left Turn 2	**−0.264**	**0.078**	−0.041	0.183
	Left Turn 3	0.048	0.063	−0.152	0.14
	Left Turn 4	**0.572**	**0.284**	**1.682**	**0.559**
Right Turn	Right Turn 1				
	Right Turn 2	N/A	N/A	**−0.246**	**0.153**
Traffic Flow	Traffic Flow 1 (Base)				
	Traffic Flow 2	**0.871**	**0.148**	−0.025	0.336
	Traffic Flow 3	0.338	0.256	0.784	0.455
X-Mastarm	X-Mastarm 1 (Base)				
	X-Mastarm 2	**0.251**	**0.077**	N/A	N/A
X-Right Turn	X-Right Turn 1 (Base)				
	X-Right Turn 2	0.038	0.068	N/A	N/A
X-Left Turn	X-Left Turn 1 (Base)				
	X-Left Turn 2	0.118	0.14	0.182	0.264
	X-Left Turn 3	0.159	0.136	0.232	0.258
	X-Left Turn 4	0.079	0.353	−1.311	1.125
X-Traffic Flow	X-Traffic Flow 1 (Base)				
	X-Traffic Flow 2	0.143	0.144	0.169	0.345
	X-Traffic Flow 3	0.292	0.224	0.234	0.473
Intersection Control Condition	Intersection Control Condition (1) (Base)				
	Intersection Control Condition (2)	**1.134**	**0.137**	0.708	0.503
	Intersection Control Condition (3)	**1.975**	**0.201**	0.922	0.554
	Intersection Control Condition (4)	**2.218**	**0.219**	**1.446**	**0.59**
	Intersection Control Condition (5)	**1.899**	**0.192**	**1.168**	**0.566**
	Intersection Control Condition 6	**1.611**	**0.186**	0.662	0.578
MOL		**0.002**	**0.001**	−0.001	0.001
X-OL		N/A	N/A	**0.002**	**0.001**
MADT		**2.997**	**0.215**	**1.845**	**0.497**
X-ADT		**3.581**	**0.388**	**2.598**	**0.718**
IRG		**−0.076**	**0.01**	0.048	0.031
APV		−0.146	0.451	0.429	0.775
		Random Effects			
Observation. ID		**0.643**	**0.018**	**2.297**	**0.462**
Goodness-of-fit Criteria					
DIC		29113.63			
$\overline{D}$		24753.94			
P_D		43596.89			
LPML		−32668.16			

Notes.
1. S.D. represents standard deviation; DIC represents deviance information criterion; $\overline{D}$ represents posterior mean deviance; P_D represents an effective number of parameters; LPML represents log pseudo marginal likelihood; NA means Not Applicable.
2. Refer to Table 1 for details of variable definition.
3. The bold fonts represent the variables with a statistically significant impact.

3.3. Model estimates

The posterior model estimates of model parameters across pedestrian and vehicle crash counts are shown in Table 3. The estimated coefficients for 10 influential variables including 'Intersection Design 2' (>four legs), 'Intersection Design 3' (offset), 'Intersection Design 4' (tee), 'Intersection Design 5' (wye), 'Intersection Design 6' (others), 'Light Condition (2)' (Lighted), 'Left turn 4' (No left turn channelization), 'Intersection Control Condition (4)' (signals semi-traffic actuated, two phase), 'Intersection Control Condition (5)' (signals full traffic actuated, multi-phase), 'MADT' (mainline-average daily traffic), and 'X-ADT' (crossline-average daily traffic), appeared to be statistically significant across both pedestrian and vehicle crash counts. Interestingly, among these significant covariates, five variables were found to have a negative impact, in which one variable ('Intersection Design 3') was common for both modes and other four variables were observed for vehicles crash count only (or, 'Intersection Design 4,' Intersection Design 5,' Intersection Design 6,' 'Right Turn 2'). It follows that, compared with the base condition of the four-legged intersection, offset intersection seems to be more advantageous in terms of traffic safety for both pedestrians and vehicle drivers. This finding is aligned with the previous studies (Bared & Kaisar, 2001; Jonsson et al., 2007; Cunningham et al., 2020), which illustrates that the implementation of offset intersection might help to decrease pedestrian-vehicle collisions. For vehicle drivers only, the tee and wye intersections and those without right-turn channels tend to provide more safety benefits compared with the base conditions of four-leg and intersection with right turn channels, respectively. The better safety performance associated with those without right turn channels is somewhat counterintuitive, which warrants further verifications from other studies.

At the individual mode level, 10 covariates that contain 'Population Group 2' (rural), 'Population Group 3' (urbanized), 'Left Turn 2' (no left turn channelization), 'Traffic Flow 2' (two-way traffic, left turn permitted), 'X-Mastarm 2' (presence of signal mastarm), 'Intersection Control Condition (2)' (stop signs on cross street only), 'Intersection Control Condition (3)' (signals pretimed), 'Intersection Control Condition 6' (others), 'MOL' (mainline - override length), and 'IRG' (intersection rate group) observed to be statistically significant for pedestrians. Similarly, for vehicles, there are two statistically significant variables that include 'Right Turn 2' (right turn channelization) and 'X-OL' (cross street - override length). Such phenomenon indicates that, relative to drivers, pedestrians are not only subject to more injury severities, but also sensitive to more intersection features such as left turn channelization, intersection control, and so on. This finding is consistent with previous studies (Pulugurtha & Sambhara, 2011; Zegeer & Bushell, 2012; Yue et al., 2020).

This study also employed various types of evaluation criteria including ($\overline{D}$) (measure of training errors), DIC (indirect measure of test errors), and LPML (measure of test errors) to assess the models from different perspectives. Under the close review of the evaluation results, it is obvious that DIC is the sum of $\overline{D}$ and P_D, where P_D serves as the correction term to the in-sample error so that DIC can approximate the out-of-sample error. Different from DIC, the LPML provide a direct cross validation-oriented error. Both values have relatively large magnitude (or, 29113.63 and −32668.16) due to the large sample size of the intersections (6,198) used in the study.

To better explore the suitability of using the bivariate setting, the random effects of the two transportations modes were also collected. Their correlation and covariance are shown in Table 4. The statistically significant correlation coefficient signifies the strong positive correlation between the two types of crashes, corroborating the importance of developing the joint models where the correlation between the two response variables was explicitly considered.

Table 4
Correlation and Covariance Matrix Between the Random Effects of Pedestrian and Vehicle Counts.

Observation. ID	β1 (Pedestrian)	β2 (Vehicle)
β1 (Pedestrian)	1.000	**0.899**
β2 (Vehicle)	0.344	1.000

Notes.
1. The lower triangle of the matrix shows the covariance, while the upper triangle of the matrix shows the associated correlation coefficient.
2. The bold font indicates the statistics are statistically significant at the significance level of 0.05.
3. Correlation coefficients are listed in the diagonal of the matrix at the same time.
4. Refer to Equation 6 for the definition of covariance of the two random effects.

4. Conclusions

Compared with the vehicle modes, much less research has been dedicated to the development of SPF for active transportation modes such as pedestrians. There are multiple reasons behind this, which include the dominant use of vehicle modes and the difficulty in obtaining exposure information of pedestrians. For enhancing current SPF development practice from perspectives of methodology and data usage, the paper aims to develop a joint SPF involving both pedestrian- and vehicle-related crashes.

The following conclusions were drawn based on the research results:

1. Compared with the base condition of four-legged intersection, offset intersection demonstrates better safety performance for both pedestrians and drivers. This finding suggests that implementation of offset intersection may help to reduce pedestrian-vehicle collisions significantly.
2. For drivers only, the tee and wye intersections and those without right-turn channels tend to provide more safety benefits compared with the base conditions of intersections with four-leg and right turn channels, respectively. The better safety performance associated with the intersections without right turn channels is relatively contradictory, which warrants further investigations from other studies.
3. There are much more statistically significant variables associated with pedestrians, suggesting that pedestrians are more sensitive to various intersection features than the vehicle drivers.
4. The correlation and covariance matrix between the random effects of both pedestrian and vehicle counts demonstrate the existence of strong correlation, indicating the sensibility of using the bivariate models, which explicitly consider the correlation between the two modes.

The aforementioned findings from this study reflect an improvement to current SPF development with mode-specific inputs of predictors and count model-estimated pedestrian exposure being used. Such improvement can help safety practitioners in reducing the pedestrian-related crashes by centralizing the resources toward potential improvements of intersection features. However, it is important to mention that the current findings are based on the empirical results obtained from the intersection-related crash data in California. Some of the model findings may not hold true when employing data at a different spatial level. The present paper employed Poisson-lognormal model for modeling the pedestrian-related crash data. Future studies may explore the other formulations such as zero-inflated models (both for the Poisson and negative binomial models), Gamma model, and Conway-Maxwell-Poisson model, which might lead to different findings from the present study. Moreover, only crashes of two modes are investigated. More modes involved might lead to different results, given more complex interrelationships are introduced among all crash outcomes. Finally, this study considered timely aggregated crashes only. The consideration of serial correlation among various years of crashes is also worth further investigation.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Wen Cheng, Mankirat Singh, Jerry Kwong, Dean Samuelson; data collection: Menglu Cao, Dean Samuelson, Jerry Kwong; analysis and interpretation of results: Bengang Li, Wen Cheng; draft manuscript preparation: Mankirat Singh, Wen Cheng. All authors reviewed the results and approved the final version of the manuscript.

Declarations of interest

None.

Acknowledgements

Financial support for this study was provided by Caltrans (Grant number: 65A0705). We are indebted to Safe Transportation Research and Education Center (SafeTREC) from UC Berkeley for the provision of estimated pedestrian counts. The efforts and comments from the anonymous reviewers are significantly appreciated as well.

References

Aguero-Valverde, J. (2013). Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: Comparing the precision of crash frequency estimates. *Accident Analysis & Prevention, 50*, 289–297.

Anarkooli, A. J., Persaud, B., Hosseinpour, M., & Saleem, T. (2019). Comparison of univariate and two-stage approaches for estimating crash frequency by severity-case study for horizontal curves on two-lane rural roads. *Accident Analysis & Prevention, 129*, 382–389.

Bared, J. G., & Kaisar, E. I. (2001). Advantages of offset T-intersections with guidelines. In Traffic safety on three continents: International conference in Moscow, Russia, 19-21 September, 2001 (Vol. 18, No. 2, p. 16).

Barker, S. P., O'neill, B., Haddon, W., & Long, W. B. (1974). The injury severity score: A method for describing patients with multiple injuries and evaluating emergency care. The Journal of Trauma: Injury, Infection, and Critical Care, 14 (3), 187-196.

Barton, B. K., & Morrongiello, B. A. (2011). Examining the impact of traffic environment and executive functioning on children's pedestrian behaviors. *Developmental Psychology, 47*(1), 182–191.

Cheng, W., Gurdiljot, S. G., Yongping, Z., & Zhong, C. (2018a). Bayesian spatiotemporal crash frequency models with mixture components for space-time interactions. *Accident Analysis & Prevention, 112*, 84–93.

Cheng, W., Gill, G. S., Ensch, J. L., Kwong, J., & Jia, X. (2018b). Multimodal crash frequency modeling: Multivariate space-time models with alternate spatiotemporal interactions. *Accident Analysis & Prevention, 113*, 159–170.

Congdon, P. (2006). Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. *Computational Statistics and Data Analysis, 50*, 346–357.

Cunningham, C. M., Baek, J., & Yang, G. (2020). Operational Applications of Signalized Offset T-Intersections-Final Report.

Cutler, A., Cutler, D., & Stevens, J. (2012). Random forests. Ensemble Machine Learning, 157-175.

de Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives, 118*(8), 1109–1116.

Dommes, A., Cavallo, V., Dubuisson, J.-B., Tournier, I., & Vienne, F. (2014). Crossing a two-way street: Comparison of young and old pedestrians. *Journal of Safety Research, 50*, 27–34.

Fatholahzade, N., Akbarizadeh, G., & Romoozi, M. (2018). Implementation of random forest algorithm in order to use big data to improve real-time. *Journal of Advances in Computer Engineering and Technology, 4*(2), 51–60.

Gates, T., Savolainen, P., Stapleton, S., Kirsch, T., & Miraskar, S. (2016). *Development of Safety Performance Functions and Other Decision Support Tools to Assess Pedestrian and Bicycle Safety*. Transportation Research Center For Livable Communities.

Gelfand, A. E. (1996). Model determination using sampling-based methods. *Markov chain Monte Carlo in practice*, 145–161.

Geruschat, D. R., & Hassan, S. E. (2005). Driver behavior in yielding to sighted and blind pedestrians at roundabouts. *Journal of Visual Impairment & Blindness, 99* (5), 286–302.

Governor Highway Safety Association - GHSA, 2019. Bicyclist and Pedestrian Safety.

Griswold, J. B., Medury, A., Schneider, R. J., Amos, D., Li, A., & Grembek, O. (2019). A pedestrian exposure model for the california state highway system. *Transportation Research Record, 2673*(4), 941–950.

Harb, R., Yan, X., Radwan, E., & Su, X. (2009). Exploring precrash maneuvers using classification trees and random forests. *Accident Analysis & Prevention, 41*(1), 98–107.

Harwood, D. W., Bauer, K. M., Richard, K. R., Gilmore, D. K., Graham, J. L., Potts, I. B., ... & Hauer, E. (2008). Pedestrian safety prediction methodology (No. NCHRP Project 17-26).

Hauer, E. (2001). Overdispersion in modeling accidents on road sections and in empirical bayes estimation. *Accident Analysis and Prevention., 33*(6), 799–808.

Russo, B. J., Savolainen, P. T., Schneider, W. H., & Anastasopoulos, P. C. (2014). Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. *Analytic Methods in Accident Research, 2*, 21–29.

Heydari, S., Fu, L., Miranda-Moreno, L. F., & Jopseph, L. (2017). Using a flexible multivariate latent class approach to model correlated outcomes: A joint analysis of pedestrian and cyclist injuries. *Analytic Methods in Accident Research, 13*, 16–27.

Hurvich, C. M., & Tsai, C. L. (1998). A crossvalidatory AIC for hard wavelet thresholding in spatially adaptive function estimation. *Biometrika, 85*(3), 701–710.

Jiang, X., Abdel-Aty, M., Hu, J., & Lee, J. (2016). Investigating macro-level hotzone identification and variable importance using big data: A random forest models approach. *Neurocomputing, 181*, 53–63.

Jonsson, T., Ivan, J. N., & Zhang, C. (2007). Crash prediction models for intersections on rural multilane highways: Differences by collision type. *Transportation Research Record, 2019*(1), 91–98.

LaScala, E. A., Gerber, D., & Gruenewald, P. J. (2000). Demographic and environmental correlates of pedestrian injury collisions: A spatial analysis. *Accident Analysis & Prevention, 32*(5), 651–658.

Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accident Analysis & Prevention, 78*, 146–154.

Li, Z., Chen, X., Ci, Y., Chen, C., & Zhang, G. (2019). A hierarchical Bayesian spatiotemporal random parameters approach for alcohol/drug impaired-driving crash frequency analysis. *Analytic Methods in Accident Research, 21*, 44–61.

Liu, C., & Sharma, A. (2017). Exploring spatio-temporal effects in traffic crash trend analysis. *Analytic Methods in Accident Research, 16*, 104–116.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22.

Mansfield, T. J., Peck, D., Morgan, D., McCann, B., & Teicher, P. (2018). The effects of roadway and built environment characteristics on pedestrian fatality risk: A national assessment at the neighborhood scale. *Accident Analysis & Prevention, 121*, 166–176.

McArthur, G. A., Chapman, P. B., Robert, C., Larkin, J., Haanen, J. B., Dummer, R., et al. (2014). Safety and efficacy of vemurafenib in BRAFV600E and BRAFV600K mutation-positive melanoma (BRIM-3): Extended follow-up of a phase 3, randomised, open-label study. *The Lancet Oncology, 15*(3), 323–332.

Miranda-Moreno, L. F. (2006). *Statistical models and methods for the identification of hazardous locations for safety improvements* Doctoral dissertation, Ph. D. Thesis. Department of Civil Engineering, University of Waterloo.

Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis & Prevention, 43*(5), 1624–1634.

National Highway Traffic Administration – NHSTA, 2018. Pedestrian Safety.

Nicodemus, K. K. (2011). Letter to the editor: On the stability and ranking of predictors from random forest variable importance measures. *Briefings in Bioinformatics, 12*(4), 369–373.

Park, E. S., & Lord, D. (2007). Multivariate Poisson-Lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record: Journal of The Transportation Research Board, 2019*(1), 1–6.

Pettit, L. I. (1990). The conditional predictive ordinate for the normal distribution. *Journal of the Royal Statistical Society: Series B (Methodological), 52*(1), 175–184.

Plurad, D., Demetriades, D., Gruzinski, G., Preston, C., Chan, L., Gaspard, D., et al. (2006). Pedestrian injuries: The Association of Alcohol Consumption with the Type and Severity of Injuries and Outcomes. *Journal of the American College of Surgeons, 202*(6), 919–927.

Pulugurtha, S. S., & Sambhara, V. R. (2011). Pedestrian crash estimation models for signalized intersections. *Accident Analysis & Prevention, 43*(1), 439–446.

Rasciute, S., & Downward, P. (2010). Health or happiness? What is the impact of physical activity on the individual?. Kyklos, 63(2), 256-270.

Retting, R. A., Ferguson, S. A., & McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. *American Journal of Public Health, 93*(9), 1456–1463.

Schroeder, B. J., & Rouphail, N. M. (2011). Event-based modeling of driver yielding behavior at unsignalized crosswalks. *Journal of Transportation Engineering, 137* (7), 455–465.

Shi, J., Chen, Y., Ren, F., & Rong, J. (2007). Research on pedestrian behavior and traffic characteristics at unsignalized midblock crosswalk: Case study in Beijing. *Transportation Research Record, 2038*(1), 23–33.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS user manual.

Tabibi, Z., Pfeffer, K., & Sharif, J. T. (2012). The influence of demographic factors, processing speed and short-term memory on Iranian children's pedestrian skills. *Accident Analysis & Prevention, 47*, 87–93.

Taylor, B. M., & Diggle, P. J. (2014). INLA or MCMC? A tutorial and comparative evaluation for spatial prediction in log-Gaussian Cox processes. *Journal of Statistical Computation and Simulation, 84*(10), 2266–2284.

Thomas, B., & DeRobertis, M. (2013). The safety of urban cycle tracks: A review of the literature. *Accident Analysis & Prevention, 52*, 219–227.

Tulu, G. S., Washington, S., Haque, M. M., & King, M. J. (2015). Investigation of pedestrian crashes on two-way two-lane rural roads in Ethiopia. *Accident Analysis & Prevention, 78*, 118–126.

Ukkusuri, S., Hasan, S., & Aziz, H. M. A. (2011). Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. *Transportation Research Record, 2237*(1), 98–106.
Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications. *Accident Analysis and Prevention, 41*(1), 137–145.
Wu, J., Xu, H., Zheng, Y., & Tian, Z. (2018). A novel method of vehicle-pedestrian near-crash identification with roadside LiDAR data. *Accident Analysis & Prevention, 121*, 238–249.
Xu, J., Ge, Y., Qu, W., Sun, X., & Zhang, K. (2018). The mediating effect of traffic safety climate between pedestrian inconvenience and pedestrian behavior. *Accident Analysis & Prevention, 119*, 155–161.
Yue, L., Abdel-Aty, M., Wu, Y., Zheng, O. u., & Yuan, J. (2020). In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention. *Journal of Safety Research, 73*, 119–132.
Zegeer, C. V., & Bushell, M. (2012). Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention, 44* (1), 3–11.
Zheng, L., & Sayed, T. (2019). A full Bayes approach for traffic conflict-based before–after safety evaluation using extreme value theory. *Accident Analysis & Prevention, 131*, 308–315.

Mankirat Singh is a research associate in the Department of Civil Engineering at California State Polytechnic University, Pomona. His research focuses on highway safety and statistical modeling.

Wen Cheng Ph.D., P.E., PTOE, is a professor in the Department of Civil Engineering at California State Polytechnic University, Pomona. He received a M.S. from University of Arizona, and a Ph.D. from Arizona State University. His research focuses on highway safety, statistical modeling, traffic signal design and traffic simulation.

Dean Samuelson is the Chief of Traffic Safety Investigations Branch at Caltrans. He has extensive experience in traffic safety data analysis.

Jerry Kwong is one project manager of Division of Research, Innovation and System Information in Caltrans. He has extensive experience in traffic safety and infrastructure design.

Bengang Li M.S.C.E., is a research assistant in the Department of Civil Engineering at California State Polytechnic University, Pomona. His research focuses on highway safety and deep learning.

Menglu Cao M.S.C.E., is a research assistant in the Department of Civil Engineering at California State Polytechnic University, Pomona. Her research focuses on highway safety and data mining.

Yihua Li Ph.D., is an associate professor in Department of Logistics Engineering, Logistics and Traffic College, at Central South University of Forestry and Technology, Hunan, China . His research focuses on traffic safety, logistics, and big data.

Journal of Safety Research 78 (2021) 170–179

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Does driver seatbelt use increase usage among front seat passengers? An exploratory analysis

Tracy Buchman

Department of Occupational & Environmental Safety & Health University of Wisconsin-Whitewater, 800 W Main Street, Whitewater, WI 53190, United States

ARTICLE INFO

Article history:
Received 16 October 2020
Received in revised form 20 January 2021
Accepted 10 May 2021
Available online 26 May 2021

Keywords:
Seatbelt use
Front seat passengers
Traffic safety
Observational survey
Demographic predictors

ABSTRACT

Introduction and Method: Observational data collected during the Wisconsin 2017, 2018, and 2019 National Occupant Protection Use Survey (NOPUS) were analyzed for this study to explore the influence of drivers' seatbelt use on front seat passengers' usage in the same vehicle. The analyses include comparing seatbelt usage rates for drivers and front passenger(s) based on their gender and based on geographical area as well as analyses of the aggregated data. *Results:* The descriptive analyses strongly suggest that seatbelt usage rates of passengers differ considerably depending on whether the driver uses the seatbelt. When female drivers wear seatbelts, seatbelt usage rates for female front seat passengers for the three years 2017, 2018, and 2019 are 97.8%, 96.3%, and 97.1% respectively, with corresponding usage rates for male passengers being 95.5%, 93.0%, and 96.0% respectively. When male drivers wear seatbelts, the seatbelt usage rates for male front seat passengers for the three years 2017, 2018, and 2019, are 93.4%, 95.5%, and 94.3%, respectively, with the corresponding usage rates for female passengers being 97.7%, 96.0%, and 97.7%, respectively. The evidence suggests that drivers' use of seatbelts significantly improves the seatbelt usage of front seat passengers. Seatbelt usage rates of male passengers as well as female passengers are higher while traveling with female drivers who use seatbelts than while traveling with male drivers who use seatbelts. *Conclusions and Practical Applications:* Future seatbelt use campaigns should target males.

1. Introduction

In 2017, properly fastened seatbelts saved an estimated 14,955 lives of passenger vehicle occupants, and 2,549 additional lives could have been saved if all unrestrained passenger vehicle occupants involved in fatal crashes had worn their seatbelts (National Center for Statistics and Analysis, 2019a). Among passenger vehicle occupants killed in 2018, who had known restraint use, almost half (47%) were unrestrained (National Center for Statistics and Analysis, 2019c). Seatbelts, when used, are estimated to reduce the risk of fatal injury to front-seat passenger car occupants by 45% and the risk of moderate-to-critical injury by 50% (National Center for Statistics and Analysis, 2019b; National Highway Traffic Safety Administration, 2019).

Numerous studies have documented the benefits of seatbelt use including reduced injury risk, injury severity and fatalities (Blincoe, Miller, Zaloshnja, & Lawrence, 2015; Goetzke & Islam, 2015; Høye, 2016; Sunshine, Dwyer-Lindgren, Chen, & Mokdad, 2017) and discussed the role that gender, age, and level of education influence in seatbelt use (Birru, Rudisill, Fabio, & Zhu, 2016; Demirer, Durat, & Haşimoğlu, 2012; Enriquez & Pickrell, 2019; Richard et al., 2019; Wells, Williams, & Farmer, 2002). Birru et al. (2016) compared self-reported seatbelt use in the Appalachian and non-Appalachian counties of the United States and found that regardless of sex, age, or rurality, respondents who reside in Appalachian counties were less likely to consistently wear their seatbelt, females typically wear seatbelts more than males, and use increases with age. Demirer et al. (2012) investigated, through questionnaires distributed to 1,000 participants in four different education levels and determined that increased levels of education are correlated with increased seatbelt usage, fewer crashes, and reduced crash severities. Richard et al. (2019) determined through exploratory research that occasional seatbelt users were more likely to be older and male. Wells et al. (2002) conducted belt use observations at gas stations in Boston, Chicago, Houston, and New York City, in addition to short interviews with drivers, and found belt use was higher in primary enforcement cities, among women, and among those with at least a college degree. Finally, the 2018 National Occupant Protection Use Survey (NOPUS) found that seatbelt use in front seats continued to be higher for females (92.0%) than that of males (87.7%; Enriquez & Pickrell, 2019).

E-mail address: buchmant@uww.edu

https://doi.org/10.1016/j.jsr.2021.05.005

An extensive body of research has also shown that traffic safety countermeasures, including focus on engineering, law enforcement, and public education, are proven strategies for increasing seatbelt use and reducing traffic fatalities (Carter, Flannagan, Bingham, Cunningham, & Rupp, 2014; Dinh-Zarr et al., 2001; Goodwin, 2015; Harper, Strumpf, Burris, Smith, & Lynch, 2014; Shults et al., 2019; Shults & Beck, 2012; Strine et al., 2010; Sung, Mizenko, & Coleman, 2017).

Goodwin (2015) highlighted that the most effective strategy for achieving and maintaining restraint use at acceptable levels is a well-publicized high visibility enforcement of strong occupant restraint use laws. The strategy includes three components: laws, enforcement, and publicity, and if one of the components is weak or missing, the effectiveness decreases. Harper et al. (2014) investigated the differential effect of mandatory seatbelt laws on seatbelt use among socioeconomic subgroups and identified the differential effect of legislation across higher versus lower education individuals using a difference-in-differences model based on state variations in the timing of the passage of laws. They found positive effects of mandatory seatbelt laws for all education groups but stronger effects for those with less education. Shults et al. (2019) and Shults and Beck (2012) studied self-reported seatbelt use and showed an increase in states with primary enforcement laws and enhanced enforcement.

Carter et al. (2014) evaluated the potential impact of rescinding seatbelt laws on annual crash related fatalities, nonfatal injuries, and associated economic costs. Their results showed that primary seatbelt laws are most effective among those who are least likely to wear seatbelts, including adolescents and those who are less-educated, lower-income, male, or live in rural areas. Dinh-Zarr et al. (2001) used systematic reviews to evaluate the effectiveness of three interventions to increase seatbelt use, and found strong evidence for the effectiveness of seatbelt laws and for the incremental effectiveness of primary seatbelt laws. Additionally, they found evidence for the effectiveness of enhanced enforcement programs for seatbelt laws.

Several studies also demonstrate how advances in technology increase seatbelt use (Høye, 2016; Jermakian & Weast, 2018; Kidd, McCartt, & Oesch, 2014; Kidd, Singer, Huey, & Kerfoot, 2018; Kidd & Singer, 2019; Van Houten et al., 2010; Van Houten, Hilton, Schulman, & Reagan, 2011).

The State of Wisconsin enacted a primary safety belt enforcement law in July 2009, and the Wisconsin Department of Transportation has conducted annual statewide observation surveys of safety belt use since March 1987. Results of these surveys consistently displayed a gender difference of approximately 10% or more in safety belt use favoring women (Field observation of safety belt use in Wisconsin (July 2016)—Wisconsin Digital Archives—Wisconsin Digital Archives, n.d.).

Other studies have evaluated the influence that seatbelt usage by a driver has on seatbelt use by other front seat occupants (Boakye, Shults, & Everett, 2019; Nambisan & Vasudevan, 2007). They concluded that passengers are much more likely to use seatbelts if their respective drivers used seatbelts. Nambisan and Vasudevan (2007) found no behavioral differences for male drivers compared to female drivers, or for male passengers compared to female passengers and determined there was no difference regarding the combination of the genders of the drivers and passenger (i.e., both of one gender, or of opposite genders).

In 2019, seatbelt use by adult front seat occupants was estimated to be 90.7%, based on the National Occupant Protection Use Survey (NOPUS), representing an increasing trend over a 15-year period (National Center for Statistics and Analysis, 2019d). Despite recent movement towards higher national seatbelt use rates, almost 10% of front seat occupants still only use seatbelts part-time. The purpose of this study was to continue to explore the influence of drivers' seatbelt use on other front seat passengers' usage of the same vehicle, essentially replicating the Nambisan and Vasudevan (2007) study for the State of Wisconsin over the years 2017–2019. After several years of investing to increase seatbelt usage, this study may help policy makers focus future campaign and enforcement efforts.

2. Materials and methods

For front seat occupants, including drivers and passengers, usage of seatbelt based on field observations were adopted to determine the rates. The rates were calculated for three consecutive years, 2017, 2018, and 2019. Multiple sites were selected adhering to the National Highway Traffic Safety Administration (NHTSA) and Department of Transportation, Criteria (2011). This weighted survey methodology permits selection of observation sites that are representative of road segments in the state of Wisconsin.

Data from each observation site collected in a year are saved in large data sets. All sites throughout Wisconsin are divided into four strata. Milwaukee metropolitan area and surrounding counties consist of stratum 1. The rest of the state provided representative counties to form the sample sites. In the analysis of seatbelt use rates, descriptive statistics were obtained first, and then statistical tests were performed on the descriptive statistics between comparative groups. Statistical significance results were reported on those tests. The analyses were done in three levels. First overall statewide results are calculated. The same analyses were broken down to stratum 1 versus strata 2, 3, and 4 combined.

The main descriptive statistics on the average rate of seatbelt use were computed based on the observation of drivers and front seat passengers who use seatbelts and all drivers and front seat passengers (regardless of the seatbelt usage). The statistical tests for significance compared these rates between groups with defining conditions. One such defining condition is the rate of the male passenger wearing a seatbelt while a female driver is wearing a seatbelt versus male passenger wearing a seatbelt while a male driver is wearing a seatbelt.

2.1. Statistical analyses

Standard procedures to test differences of two proportions were used throughout this research to determine any significant difference between comparison groups. The null hypothesis is stated as the rate of seatbelt usage by front seat passengers as influenced by gender of driver; the rate of seatbelt usage by front passengers is not influenced by drivers' seatbelt use. The alternative hypothesis states that the rate of seatbelt usage by front seat passengers is higher if drivers use seatbelts.

The following terminology will be followed in this research (closely adopted from the equations used in Nambisan and Vasudevan (2007)):

p_1 = proportion of overall seatbelt usage (i.e., for all front seat occupants without regard to seatbelt usage by the driver).
p_2 = proportion of front seat occupants using seatbelts when their drivers use the seatbelt.

Then, a hypothesis of equality of proportions is to be tested against the 1-sided alternative that $p_1 < p_2$. Formally:

Null Hypothesis: H_0: $p_1 - p_2 = 0$;
Alternative Hypothesis: H_1: $p_1 - p_2 < 0$.

The z-statistic based test for the null hypothesis is appropriate for this test. The test statistic is defined as:

$$Z_{OBS} = \frac{(\widehat{p}_2 - \widehat{p}_1) - (p_2 - p_1)}{\sigma_{(\widehat{p}_2 - \widehat{p}_1)}} \quad (1)$$

where

$\widehat{p}_1$ = the sample proportion for overall seatbelt usage (calculated for all front seat occupants without regard to seatbelt usage by the driver).

$$= X_1/n_1 \quad (2)$$

$\widehat{p}_2$ = the sample proportion of front seat occupants using seatbelt when drivers use seatbelt.

$$= X_2/n_2 \quad (3)$$

$p_2 - p_1$ = hypothetical difference for the test
= 0, since the test is for equal proportions

$$\sigma_{(\widehat{p}_2 - \widehat{p}_1)} = \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}} \quad (4)$$

where

X_1 = the number of front seat occupants wearing seatbelts (all occupants without regard to seatbelt usage by the driver).
X_2 = the number of front seat occupants wearing seatbelts (when driver uses seatbelt).
n_1 = the total number of front seat occupants in the vehicles with belts (all occupants regardless to seatbelt usage by the driver), and
n_2 = the total number of front seat occupants in the vehicle (when driver uses seatbelt).

To test the null hypothesis against the alternative hypothesis, the observed Z values (Z_{OBS}) were compared with $Z_{critical}$ values with desired level of confidence (here 95%). From the standard normal distribution table, it was found that the 95% confidence level corresponds to the $Z_{critical}$ value of 1.645. If $Z_{OBS} > Z_{critical}$, the null hypothesis is rejected. Otherwise, the null hypothesis is not rejected. By rejecting the null hypothesis and hence, by accepting the alternative hypothesis, it can be concluded with 95% confidence that the alternative hypothesis is correct.

2.2. Data collection

Data collected during the Wisconsin 2017, 2018, and 2019 National Occupant Protection Use Survey (NOPUS) were analyzed for this study. Data represents field observations at 240 locations across the state of Wisconsin within 48 counties accounting for 52% of the primary road segments in the state each of the three years. All seatbelt use observations were conducted during weekdays and weekends between 7:00 a.m. and 6:00 p.m. The schedule included rush hour (before 9:30 a.m. and after 3:30 p.m.) and non-rush hour observations. Data collection was conducted for 60 minutes by one observer at each site utilizing iPads to document observations. Start times were staggered to ensure that a representative number of weekday/weekend/rush hour/non-rush hour sites were included each day. All passenger vehicles, including commercial vehicles weighing less than 10,000 pounds, were eligible for observation.

2.3. Site selection

The sites for data collection include 240 locations across the state of Wisconsin. Locations were selected using The Uniform Criteria for State Observational Surveys of Seatbelt Use, used by NHTSA since 2011 (Criteria, 2011). NHTSA survey guidelines (NHTSA, 2011) require a state to sample seatbelt use in counties that comprise at least 85% of the state's passenger vehicle occupant fatalities (excluding pedestrian, bicyclist, motorcyclist, moped-user, and commercial vehicle-related deaths). In addition to covering over 85% of the state's occupant fatalities, the selected counties also contain almost 91% of the state's population and 89% of its Vehicle Miles Traveled (VMT).

For this study, counties were divided into four strata regions representing three road types; interstate highways, other highways, comprised of state trunk highways, and local roads/streets, with the aim of sampling locations corresponding to where VMT in the state is occurring. Stratum 1 represents the Milwaukee Metropolitan Statistical Area as defined by the U.S. Census because of its disproportionately high concentration of population and VMT. This area is also of interest to Wisconsin DOT staff because it has consistently reported the lowest seatbelt usage of any other measured region in the state, mostly due to consistently low usage in the region's large African American community. The other three Strata are defined by dividing the collective VMT of the remaining counties into approximate thirds based on previous year's VMT. As noted, Stratum 1 contains the census-defined Milwaukee Metropolitan Area, Stratum 2 contains the counties having the next largest populations and amounts of VMT per county, and so forth with Stratum 3 and Stratum 4 having the next downward levels of populations and VMT per county in the state, respectively. While it was not possible to ensure that each stratum reported the exact same level of previous year's VMT—especially considering that the strata corresponding to the Milwaukee Metropolitan Area was predetermined--care was taken to ensure that the VMT of each stratum was as close as possible from year to year.

2.4. Types of analysis

Seatbelt usage for drivers and passengers are studied separately in this research. While calculating the rates of usage, gender specific strata are considered and compared. Rates for both female and male passengers were found for both female and male drivers.

2.5. Summary of the analyses

As done in the previous research of Nambisan and Vasudevan (2007), the analyses compared seatbelt usage rates for drivers and front passenger(s) based on their gender and based on geographical area (stratum 1 vs. others) as well as analyses of the aggregated data. The analyses are divided into two major sections, when drivers use the seatbelts and when drivers do not use seatbelts. Though the sample size for the latter group is substantially lower than for the former group, it is imperative to find the average rate differences between the groups.

Table 1 shows seatbelt usage rates for passengers in the front seat without regard to the drivers' use of the seatbelts for each of the three years considered in this study: 2017, 2018, and 2019. For each year the table presents information in row wise panels. The top panel presents seatbelt usage for front seat passengers only when any passenger was present. The bottom panel presents the similar rates for all front seat riders, including drivers.

Table 1
Average seatbelt usage rates for front seat passengers (without considering seatbelt usage by drivers).

Area	Category	Year 2017			Year 2018			Year 2019		
		Passengers with SB	Total # Passengers	% SB Usage	Passengers with SB	Total # Passengers	% SB Usage	Passengers with SB	Total # Passengers	% SB Usage
Statewide	All Male	1267	1502	84.30%	1771	2008	88.20%	1995	2235	89.26%
	All Female	2919	3143	92.90%	3760	4062	92.60%	4327	4601	94.04%
	All Passengers	4186	4645	90.12%	5531	6070	91.12%	6322	6836	92.48%
Stratum 1	All Male	177	209	84.69%	459	505	90.89%	590	644	91.61%
	All Female	407	421	96.67%	929	981	94.70%	1196	1249	95.76%
	All Passengers	584	630	92.70%	1388	1486	93.41%	1786	1893	94.35%
Strata 2, 3, & 4	All Male	1090	1293	84.30%	1312	1503	87.29%	1405	1591	88.31%
	All Female	2512	2722	92.29%	2831	3081	91.89%	3131	3352	93.41%
	All Passengers	3602	4015	89.71%	4143	4584	90.38%	4536	4943	91.77%

2.6. Seatbelt usage by passengers based on seatbelt usage by drivers

To test the null hypothesis in different segments of population across multiple strata, the rates of seatbelt usage of passengers were estimated based on whether drivers were wearing seatbelts or not. The data are analyzed based on statewide, stratum 1 and strata 2–4 geographical area types and by gender.

2.6.1. Statewide

Table 2 presents seatbelt usage rates of passengers only, contrasted on different scenarios. The gender of the driver and their seatbelt situations were contrasted for all three years. For example, when the driver is a male and wearing a seatbelt, the seatbelt usage rates for passengers may be different depending on the gender of the passenger. The table also has two sets of columns: the first set of columns is for drivers using seatbelts, and the second set of columns is for drivers who do not use seatbelts. Each of these sets of columns includes columns with data for the number of passengers using seatbelts, the total number of passengers observed, and the corresponding percentage using seatbelts.

The influence of drivers' seatbelt usage on passengers' usage is evident from Table 2 following similar techniques used in Nambisan and Vasudevan (2007). When male drivers wear seatbelts, the seatbelt usage rates for male front seat passengers for the three years 2017, 2018, and 2019 are 93.4%, 95.5%, and 94.3%, respectively, with the corresponding usage rates for female passengers being 97.7%, 96.0%, and 97.7%, respectively. When female drivers wear seatbelts, the seatbelt usage rates for female front seat passengers for the three years 2017, 2018, and 2019 are 97.8%, 96.3%, and 97.1%, respectively, with corresponding usage rates for male passengers being 95.5%, 93.0%, and 96.0%, respectively. Likewise, for those three years, the overall usage rates for all passengers with male drivers use seatbelts are 96.8%, 95.9%, and 96.9%, respectively, and for female drivers 96.7%, 94.8%, and

Table 2
Seatbelt usage rates of passengers based on drivers' seatbelt usage rates (statewide).

Year	Passenger Category / Driver Category	Passengers with SB	Total # Passengers	% Usage	Passengers with SB	Total # Passengers	% Usage
2017		Male Drivers with SB			Male Drivers without SB		
	All Male	584	625	93.44%	19	131	14.50%
	All Female	2045	2092	97.75%	54	170	31.76%
	All Passengers	2629	2717	96.76%	73	301	24.25%
		Female Drivers with SB			Female Drivers without SB		
	All Male	636	666	95.50%	10	58	17.24%
	All Female	788	806	97.77%	14	55	25.45%
	All Passengers	1424	1472	96.74%	24	113	21.24%
2018		Male Drivers with SB			Male Drivers without SB		
	All Male	846	886	95.49%	27	112	24.11%
	All Female	2599	2707	96.01%	53	163	32.52%
	All Passengers	3445	3593	95.88%	80	275	29.09%
		Female Drivers with SB			Female Drivers without SB		
	All Male	846	910	92.97%	26	66	39.39%
	All Female	1027	1066	96.34%	24	61	39.34%
	All Passengers	1873	1976	94.79%	50	127	39.37%
2019		Male Drivers with SB			Male Drivers without SB		
	All Male	940	997	94.28%	33	126	26.19%
	All Female	3114	3186	97.74%	59	188	31.38%
	All Passengers	4054	4183	96.92%	92	314	29.30%
		Female Drivers with SB			Female Drivers without SB		
	All Male	979	1025	95.51%	26	67	38.81%
	All Female	1117	1150	97.13%	18	53	33.96%
	All Passengers	2096	2175	96.37%	44	120	36.67%

96.4%, respectively. These rates are notably higher than the corresponding values for all passengers when considered without regard to the drivers' use of seatbelts (as shown in Table 1). The seatbelt usage rates of male passengers are much lower (14.5%, 24.1%, and 26.2% respectively for the three years) when male drivers do not use seatbelts. Table 2 also provides similarly detailed information regarding passengers' seatbelt usage rates depending on female drivers' seatbelt usage. It is interesting to note that seatbelt use rates for passengers when female drivers do not use seatbelts is slightly better than that when male drivers do not use seatbelts. A comparison of the information summarized in Tables 1 and 2, shows that seatbelt usage rates for all categories are higher when either male or female drivers wear seatbelts compared to the average seatbelt usage rates shown in Table 1.

2.6.2. Stratum 1

Table 3 shows details of seatbelt usage rates of passengers when drivers use and do not use seatbelts for stratum 1 for the years 2017–2019. Its format is identical to that of Table 2. Data presented in Table 3 show trends very similar to the statewide usage rates shown in Table 2. For male drivers wearing seatbelts, seatbelt usage rates for male front seat passengers for the three years 2017, 2018, and 2019 are 96.1%, 97.9%, and 94.6%, respectively. For female passengers the rates are 99.6%, 98.7%, and 98.8%, respectively, for the three years of study. The overall usage rates for all passengers, when male drivers use seatbelts, are 98.9%, 98.5%, and 97.7%, respectively, for these years. As in the statewide analysis, these rates are notably higher than the corresponding values listed in Table 1 (for all passengers when considered without regard to drivers' usage of seatbelts). Seatbelt usage rates of male passengers are much lower (7.1%, 20.8%, and 19.2%, respectively, for the three years) when the male drivers do not use seatbelts. Similar results are also evident in Table 3 regarding passenger seatbelt usage rates depending on female drivers' seatbelt usage.

2.6.3. Strata 2, 3, & 4

Table 4 shows details of seatbelt usage rates of passengers when drivers use and do not use seatbelts for combined strata 2, 3, and 4 for the years 2017–2019. Its format is identical to that of Tables 2 and 3. Data presented in Table 4 show trends very similar to the statewide usage rates shown in Tables 2 and 3. For male drivers wearing seatbelts, the seatbelt usage rates for male front seat passengers for the three years 2017, 2018, and 2019 are 93.0%, 94.6%, and 94.1%, respectively. And for female passengers the rates are 97.5%, 95.2%, and 97.3%, respectively, for the three years of study. The overall usage rates for all passengers when male drivers use seatbelts are 96.4%, 95.1%, and 96.6%, respectively for these years. As in the statewide analysis, these rates are notably higher than the corresponding values listed in Table 1 (for all passengers when considered without regard to drivers' usage of seatbelts). Seatbelt usage rates of male passengers are much lower (15.4%, 25.0%, and 28.0%, respectively, for the three years) when the male drivers do not use seatbelts. Similar results are also evident in Table 4 regarding passengers' seatbelt usage rates depending on female drivers' seatbelt usage habits.

2.7. Statistical analysis

The descriptive analyses suggest that the seatbelt usage rates of passengers differ depending on whether the driver uses a seatbelt. The statistical significance of the differences in seatbelt use rates is summarized in this section. The methods used for analyses were described in the methodology section. For Eqs. (1) and (2), values of X_1 and n_1 are obtained from Table 1. Values for X_2 and n_2 are obtained from Tables 2–4, as appropriate. The Z values are used to test the statistical significance at the 95% confidence level (Nambisan & Vasudevan, 2007).

Tables 5–7 summarize results from the statistical analysis of comparisons of seatbelt usage rates of passengers when drivers use seatbelts to seatbelt usage rates of passengers without regard to driver's use of the seatbelt for statewide data, stratum 1 and

Table 3
Seatbelt usage rates of passengers based on drivers' seatbelt usage rates (stratum 1).

Year	Passenger Category	Passengers with SB	Total # Passengers	% Usage	Passengers with SB	Total # Passengers	% Usage
	Driver Category						
2017		Male Drivers with SB			Male Drivers without SB		
	All Male	75	78	96.15%	1	14	7.14%
	All Female	279	280	99.64%	2	9	22.22%
	All Passengers	354	358	98.88%	3	23	13.04%
		Female Drivers with SB			Female Drivers without SB		
	All Male	99	104	95.19%	.	.	.
	All Female	124	125	99.20%	1	6	16.67%
	All Passengers	223	229	97.38%	1	17	5.88%
2018		Male Drivers with SB			Male Drivers without SB		
	All Male	234	239	97.91%	5	24	20.83%
	All Female	609	617	98.70%	7	34	20.59%
	All Passengers	843	856	98.48%	12	58	20.69%
		Female Drivers with SB			Female Drivers without SB		
	All Male	215	227	94.71%	5	15	33.33%
	All Female	305	313	97.44%	5	14	35.71%
	All Passengers	520	540	96.30%	10	29	34.48%
2019		Male Drivers with SB			Male Drivers without SB		
	All Male	283	299	94.65%	5	26	19.23%
	All Female	885	896	98.77%	12	33	36.36%
	All Passengers	1168	1195	97.74%	17	59	28.81%
		Female Drivers with SB			Female Drivers without SB		
	All Male	291	297	97.98%	6	17	35.29%
	All Female	295	303	97.36%	2	15	13.33%
	All Passengers	586	600	97.67%	8	32	25.00%

Table 4
Seatbelt usage rates of passengers based on drivers' seatbelt usage rates (strata 2, 3, & 4).

Year	Passenger Category	Passengers with SB	Total # Passengers	% Usage	Passengers with SB	Total # Passengers	% Usage
	Driver Category						
2017		Male Drivers with SB			Male Drivers without SB		
	All Male	509	547	93.05%	18	117	15.39%
	All Female	1766	1812	97.46%	52	161	32.30%
	All Passengers	2275	2359	96.44%	70	278	25.18%
		Female Drivers with SB			Female Drivers without SB		
	All Male	537	562	95.55%	10	47	21.28%
	All Female	664	681	97.50%	13	49	26.53%
	All Passengers	1201	1243	96.62%	23	96	23.96%
2018		Male Drivers with SB			Male Drivers without SB		
	All Male	612	647	94.59%	22	88	25.00%
	All Female	1990	2090	95.22%	46	129	35.66%
	All Passengers	2602	2737	95.07%	68	217	31.34%
		Female Drivers with SB			Female Drivers without SB		
	All Male	631	683	92.39%	21	51	41.18%
	All Female	722	753	95.88%	19	47	40.43%
	All Passengers	1353	1436	94.22%	40	98	40.82%
2019		Male Drivers with SB			Male Drivers without SB		
	All Male	657	698	94.13%	28	100	28.00%
	All Female	2229	2290	97.34%	47	155	30.32%
	All Passengers	2886	2988	96.59%	75	255	29.41%
		Female Drivers with SB			Female Drivers without SB		
	All Male	688	728	94.51%	20	50	40.00%
	All Female	822	847	97.05%	16	38	42.11%
	All Passengers	1510	1575	95.87%	36	88	40.91%

Table 5
Seatbelt usage rates of passengers when drivers use seatbelts (statewide).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male with SB	All Male	584	625	93.44%	7.83	Yes
		All Female	2045	2092	97.75%	7.87	Yes
		All Passengers	2629	2717	96.76%	11.18	Yes
	Female with SB	All Male	636	666	95.50%	4.45	Yes
		All Female	788	806	97.77%	4.58	Yes
		All Passengers	1424	1472	96.74%	6.36	Yes
2018	Male with SB	All Male	846	886	95.49%	6.46	Yes
		All Female	2599	2707	96.01%	6.01	Yes
		All Passengers	3445	3593	95.88%	8.63	Yes
	Female with SB	All Male	846	910	92.97%	2.97	Yes
		All Female	1027	1066	96.34%	3.31	Yes
		All Passengers	1873	1976	94.79%	4.41	Yes
2019	Male with SB	All Male	940	997	94.28%	6.22	Yes
		All Female	3114	3186	97.74%	7.78	Yes
		All Passengers	4054	4183	96.92%	10.02	Yes
	Female with SB	All Male	979	1025	95.51%	3.29	Yes
		All Female	1117	1150	97.13%	3.43	Yes
		All Passengers	2096	2175	96.37%	4.75	Yes

strata 2–4 respectively. Each of these tables has two sets of rows for each year: the first set of rows is for the male drivers and the second set of rows for the female drivers. Within each set of rows, the information is provided based on the gender of the passengers.

3. Results and discussion

Tables 5–10 show that the seatbelt usage rates of passengers in the front seat change significantly with driver wearing of a seatbelt. The increase in rates are prevalent among both male and female passengers when driver's use the seatbelt as compared to the case when the driver's use of seatbelt is not considered. These results are consistent over all three study years (Tables 5–7). Likewise, seatbelt usage rates of passengers are significantly lower when the drivers do not use seatbelts (Tables 8–10). These results are consistent across the years, as well as whether the analyses are aggregated statewide or shown segregated for the strata categories.

Data from Tables 2–4 show that the seatbelt usage rates are, in general, higher for female passengers than for male passengers for all combination of drivers (male/female, with/without seatbelts). The tables also illustrate that seatbelt usage rates of male passengers as well as female passengers are higher while traveling with female drivers who use seatbelts compared to traveling with male drivers who use seatbelts. For drivers who do not use seatbelts, in general, male passengers showed a lower rate of seatbelt usage when compared to female passengers.

4. Conclusions and practical applications

The findings presented here for Wisconsin suggest females are more likely to wear seatbelts than males. That result is consistent

Table 6
Seatbelt usage rates of passengers when drivers use seatbelts (stratum 1).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male with SB	All Male	75	78	96.15%	3.47	Yes
		All Female	279	280	99.64%	3.15	Yes
		All Passengers	354	358	98.88%	5.26	Yes
	Female with SB	All Male	99	104	95.19%	3.23	Yes
		All Female	124	125	99.20%	2.14	Yes
		All Passengers	223	229	97.38%	3.16	Yes
2018	Male with SB	All Male	234	239	97.91%	4.44	Yes
		All Female	609	617	98.70%	4.72	Yes
		All Passengers	843	856	98.48%	6.61	Yes
	Female with SB	All Male	215	227	94.71%	1.95	Yes
		All Female	305	313	97.44%	2.40	Yes
		All Passengers	520	540	96.30%	2.79	Yes
2019	Male with SB	All Male	283	299	94.65%	1.79	Yes
		All Female	885	896	98.77%	4.44	Yes
		All Passengers	1168	1195	97.74%	4.97	Yes
	Female with SB	All Male	291	297	97.98%	4.67	Yes
		All Female	295	303	97.36%	1.48	No
		All Passengers	586	600	97.67%	4.08	Yes

Table 7
Seatbelt usage rates of passengers when drivers use seatbelts (strata 2, 3, & 4).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male with SB	All Male	509	547	93.05%	5.89	Yes
		All Female	1766	1812	97.46%	8.20	Yes
		All Passengers	2275	2359	96.44%	10.98	Yes
	Female with SB	All Male	537	562	95.55%	8.43	Yes
		All Female	664	681	97.50%	6.63	Yes
		All Passengers	1201	1243	96.62%	9.84	Yes
2018	Male with SB	All Male	612	647	94.59%	5.90	Yes
		All Female	1990	2090	95.22%	4.91	Yes
		All Passengers	2602	2737	95.07%	7.80	Yes
	Female with SB	All Male	631	683	92.39%	3.83	Yes
		All Female	722	753	95.88%	4.57	Yes
		All Passengers	1353	1436	94.22%	5.09	Yes
2019	Male with SB	All Male	657	698	94.13%	4.85	Yes
		All Female	2229	2290	97.34%	7.21	Yes
		All Passengers	2886	2988	96.59%	9.40	Yes
	Female with SB	All Male	688	728	94.51%	5.31	Yes
		All Female	822	847	97.05%	5.04	Yes
		All Passengers	1510	1575	95.87%	6.46	Yes

Table 8
Seatbelt usage rates of passengers when drivers do not use seatbelts (statewide).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male without SB	All Male	19	131	14.50%	−19.15	Yes
		All Female	54	170	31.76%	−16.89	Yes
		All Passengers	73	301	24.25%	−25.75	Yes
	Female without SB	All Male	10	58	17.24%	−14.14	Yes
		All Female	14	55	25.45%	−11.41	Yes
		All Passengers	24	113	21.24%	−17.93	Yes
2018	Male without SB	All Male	27	112	24.11%	−15.14	Yes
		All Female	53	163	32.52%	−16.14	Yes
		All Passengers	80	275	29.09%	−22.29	Yes
	Female without SB	All Male	26	66	39.39%	−8.15	Yes
		All Female	24	61	39.34%	−8.55	Yes
		All Passengers	50	127	39.37%	−11.86	Yes
2019	Male without SB	All Male	33	126	26.19%	−14.90	Yes
		All Female	59	188	31.38%	−18.35	Yes
		All Passengers	92	314	29.30%	−24.15	Yes
	Female without SB	All Male	26	67	38.81%	−8.87	Yes
		All Female	18	53	33.96%	−9.23	Yes
		All Passengers	44	120	36.67%	−12.77	Yes

Table 9
Seatbelt usage rates of passengers when drivers do not use seatbelts (stratum 1).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male without SB	All Male	1	14	7.14%	−10.59	Yes
		All Female	2	9	22.22%	−5.36	Yes
		All Passengers	3	23	13.04%	−11.22	Yes
	Female without SB	All Male	.	.			
		All Female	1	6	16.67%	−5.25	Yes
		All Passengers	1	17	5.88%	−14.97	Yes
2018	Male without SB	All Male	5	24	20.83%	−8.35	Yes
		All Female	7	34	20.59%	−10.63	Yes
		All Passengers	12	58	20.69%	−13.57	Yes
	Female without SB	All Male	5	15	33.33%	−4.70	Yes
		All Female	5	14	35.71%	−4.60	Yes
		All Passengers	10	29	34.48%	−6.66	Yes
2019	Male without SB	All Male	5	26	19.23%	−9.27	Yes
		All Female	12	33	36.36%	−7.08	Yes
		All Passengers	17	59	28.81%	−11.07	Yes
	Female without SB	All Male	6	17	35.29%	−4.84	Yes
		All Female	2	15	13.33%	−9.37	Yes
		All Passengers	8	32	25.00%	−9.04	Yes

Table 10
Seatbelt usage rates of passengers when drivers do not use seatbelts (strata 2, 3, & 4).

Year	Driver Category	Passenger Category	Passenger with SB	Total # Passengers	% Usage	Z-Value	Statistically Significant?
2017	Male without SB	All Male	18	117	15.38%	−19.77	Yes
		All Female	52	161	32.30%	−16.12	Yes
		All Passengers	70	278	25.18%	−24.38	Yes
	Female without SB	All Male	10	47	21.28%	−10.41	Yes
		All Female	13	49	26.53%	−10.39	Yes
		All Passengers	23	96	23.96%	−15.00	Yes
2018	Male without SB	All Male	22	88	25.00%	−13.27	Yes
		All Female	46	129	35.66%	−13.24	Yes
		All Passengers	68	217	31.34%	−18.57	Yes
	Female without SB	All Male	21	51	41.18%	−6.64	Yes
		All Female	19	47	40.43%	−7.17	Yes
		All Passengers	40	98	40.82%	−9.94	Yes
2019	Male without SB	All Male	28	100	28.00%	−13.22	Yes
		All Female	47	155	30.32%	−16.97	Yes
		All Passengers	75	255	29.41%	−21.65	Yes
	Female without SB	All Male	20	50	40.00%	−6.93	Yes
		All Female	16	38	42.11%	−6.40	Yes
		All Passengers	36	88	40.91%	−9.68	Yes

with prior research (Carter et al., 2014; Enriquez & Pickrell, 2019; Richard et al., 2019; Wells et al., 2002).

Two prior studies analyzed the question of gender and the effects of driver seatbelt usage on passenger usage. Both Boakye et al. (2019) and Nambisan and Vasudevan (2007) found passengers were significantly more likely to were seatbelts if drivers did so, a result replicated here.

However, Nambisan and Vasudevan (2007) found no gender differences in terms of whether the driver and passenger were both male, both female, or of mixed gender, while Boakye et al. (2019) found passenger seatbelt use was significantly lower for same-sex drivers. The results presented here suggest that gender plays a stronger role. First, when the driver wears a seatbelt, female passengers are more likely than males to wear a seatbelt regardless of driver gender (see Tables 5–7). Second, when drivers do not wear seatbelts, male passengers are more likely to wear seatbelts when the driver is a female (see Tables 8–10).

Several studies indicate that effective and well-planned media and high-visibility enforcement (HVE) campaigns such as Click it or Ticket (CIOT) can have a significant impact on increasing seatbelt usage rates (Goodwin, 2015; Manlove, Stanley, & Peck, 2015; National Highway Traffic Safety Administration, 2009; Nichols, Chaffe, Solomon, & Tison, 2016; Solomon, Compton, & Preusser, 2004; Thomas, Cook, & Olson, 2011; Vasudevan, Nambisan, Singh, & Pearl, 2009; Williams, Wells, McCartt, & Preusser, 2000). Online social media platforms provide an effective means to target populations with the use of participatory social media using influencers to challenge followers toward modifying behavior (Drake, Zhang, Applewhite, Fowler, & Holcomb, 2017). Hezaveh and Cherry (2019) and Noar (2006) also demonstrated successful use of safety campaigns design to reach targeted geographic areas and groups with lower seatbelt usage rates.

An implication of this research is that future HVE campaigns could be targeted towards males. The reasons for doing so are two-fold. First, males are less likely to wear seatbelts generally, which justifies that approach. Second, males as passengers are less likely to wear seatbelts in the presence of male drivers, and targeted campaigns could reduce this disparity.

The Wisconsin Strategic Highway Safety Plan (Ross & Pabst, 2017) has identified increased seatbelt use as a high priority issue.

The findings in this study will allow for Wisconsin statutorily required Traffic Safety Commissions to utilize a data driven approach for future resource allocation dedicated to improve public outreach and awareness through well-planned media and HVE campaigns such as CIOT with the goal of increased seatbelt use.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Abhik Bhattacharya, PhD.

References

Birru, H., Rudisill, T. M., Fabio, A., & Zhu, M. (2016). A comparison of self-reported seatbelt usage among the Appalachian and non-Appalachian United States. *Annals of Epidemiology, 26*(3), 227–230. https://doi.org/10.1016/j.annepidem.2016.02.001.

Blincoe, L. J., Miller, T. R., Zaloshnja, E., & Lawrence, B. A. (2015). The economic and societal impact of motor vehicle crashes, 2010 (Revised)1. *Annals of Emergency Medicine, 66*(2), 194–196. https://doi.org/10.1016/j.annemergmed.2015.06.011.

Boakye, K. F., Shults, R. A., & Everett, J. D. (2019). Nighttime seatbelt use among front seat passengers: Does the driver's belt use matter? *Journal of Safety Research, 70*, 13–17. https://doi.org/10.1016/j.jsr.2019.04.004.

Carter, P. M., Flannagan, C. A. C., Bingham, C. R., Cunningham, R. M., & Rupp, J. D. (2014). Modeling the impact of rescinding Michigan's primary and secondary seatbelt laws on death and injury from passenger vehicle crashes. *Traffic Injury Prevention, 15*(7), 701–705. https://doi.org/10.1080/15389588.2013.865167.

Demirer, A., Durat, M., & Haşimoğlu, C. (2012). Investigation of seatbelt use among the drivers of different education levels. *Safety Science, 50*(4), 1005–1008. https://doi.org/10.1016/j.ssci.2011.12.013.

Dinh-Zarr, T. B., Sleet, D. A., Shults, R. A., Zaza, S., Elder, R. W., Nichols, J. L., et al. (2001). Reviews of evidence regarding interventions to increase the use of safety belts. *American Journal of Preventive Medicine, 21*(4, Supplement 1), 48–65. https://doi.org/10.1016/S0749-3797(01)00378-6.

Drake, S. A., Zhang, N., Applewhite, C., Fowler, K., & Holcomb, J. B. (2017). A social media program to increase adolescent seatbelt use. *Public Health Nursing, 34*(5), 500–504. https://doi.org/10.1111/phn.12342.

Enriquez, J., & Pickrell, T. M. (2019). *Seatbelt use in 2018—Overall results. (Traffic Safety Facts Report No. DOT HS 812 662).* National Highway Traffic Safety Administration.

Field observation of safety belt use in Wisconsin (July 2016)—Wisconsin Digital Archives—Wisconsin Digital Archives. (n.d.). Retrieved November 21, 2019, from <https://www.wistatedocuments.org/digital/collection/p267601coll4/id/13880/rec/11>.

Goetzke, F., & Islam, S. (2015). Determinants of seatbelt use: A regression analysis with FARS data corrected for self-selection. *Journal of Safety Research, 55*, 7–12. https://doi.org/10.1016/j.jsr.2015.07.004.

Goodwin, A. (2015). Countermeasures That Work: NHTSA, (pp. 437).

Harper, S., Strumpf, E. C., Burris, S., Smith, G. D., & Lynch, J. (2014). The effect of mandatory seatbelt laws on seatbelt use by socioeconomic position. *Journal of Policy Analysis & Management, 33*(1), 141–161. https://doi.org/10.1002/pam.21735.

Hezaveh, A. M., & Cherry, C. R. (2019). Neighborhood-level factors affecting seatbelt use. *Accident Analysis & Prevention, 122*, 153–161. https://doi.org/10.1016/j.aap.2018.10.005.

Høye, A. (2016). How would increasing seatbelt use affect the number of killed or seriously injured light vehicle occupants? *Accident Analysis & Prevention, 88*, 175–186. https://doi.org/10.1016/j.aap.2015.12.022.

Jermakian, J. S., & Weast, R. A. (2018). Passenger use of and attitudes toward rear seatbelts. *Journal of Safety Research, 64*, 113–119. https://doi.org/10.1016/j.jsr.2017.12.006.

Kidd, D. G., McCartt, A. T., & Oesch, N. J. (2014). Attitudes toward seatbelt use and in-vehicle technologies for encouraging belt use. *Traffic Injury Prevention, 15*(1), 10–17. https://doi.org/10.1080/15389588.2013.792111.

Kidd, D. G., & Singer, J. (2019). The effects of persistent audible seatbelt reminders and a speed-limiting interlock on the seatbelt use of drivers who do not always use a seatbelt. *Journal of Safety Research, 71*, 13–24. https://doi.org/10.1016/j.jsr.2019.09.005.

Kidd, D. G., Singer, J., Huey, R., & Kerfoot, L. (2018). The effect of a gearshift interlock on seatbelt use by drivers who do not always use a belt and its acceptance among those who do. *Journal of Safety Research, 65*, 39–51. https://doi.org/10.1016/j.jsr.2018.03.005.

Manlove, K., Stanley, L., & Peck, A. (2015). A quantitative approach to assessing the efficacy of occupant protection programs: A case study from Montana. *Accident Analysis & Prevention, 83*, 214–221. https://doi.org/10.1016/j.aap.2015.06.009.

Nambisan, S. S., & Vasudevan, V. (2007). Is seatbelt usage by front seat passengers related to seatbelt usage by their drivers? *Journal of Safety Research, 38*(5), 545–555. https://doi.org/10.1016/j.jsr.2007.06.002.

National Center for Statistics and Analysis. (2019a). Lives Saved in 2017 by Restraint Use and Minimum-Drinking-Age Laws (Traffic Safety Facts Report No. DOT HS 812 683). National Highway Traffic Safety Administration.

National Center for Statistics and Analysis. (2019b). Passenger vehicles: 2017 data. NHTSA. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812805>.

National Center for Statistics and Analysis. (2019c). 2018 fatal motor vehicle crashes: Overview. (Traffic Safety Facts Report No. DOT HS 812 826). National Highway Traffic Safety Administration. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812826>.

National Center for Statistics and Analysis. (2019d). Seat belt use in 2019 – Overall Results [Traffic Safety Facts Research Note. Report No. DOT HS 812 875]. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812875.

National Highway Traffic Safety Administration. (2009). Evaluation of the Buckle Up in Your Truck Programs.

National Highway Traffic Safety Administration. (2019, April). Occupant protection in passenger vehicles: 2017 data. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812691>.

Nichols, J., Chaffe, R., Solomon, M., & Tison, J. (2016). The Click It or Ticket Evaluation, 2013. *Traffic Safety Facts, 5*.

Noar, SethM (2006). A 10-Year retrospective of research in health mass media campaigns: Where do we go from here? *Journal of Health Communication, 11*(1), 21–42. https://doi.org/10.1080/10810730500461059.

Richard, C. R., Thomas, F. D., Blomberg, R. D., Brown, J. L., Wright, T., Graham, L., et al. (2019). *Characteristics and predictors of occasional seatbelt use using Strategic Highway Research Program 2 data (Report No. DOT HS 812 840).* National Highway Traffic Safety Administration.

Ross, D., & Pabst, D. (2017). Wisconsin Strategic Highway Safety Plan 2017-2020. 66.

Shults, R. A., & Beck, L. F. (2012). Self-reported seatbelt use, United States, 2002–2010: Does prevalence vary by state and type of seatbelt law? *Journal of Safety Research, 43*(5), 417–420. https://doi.org/10.1016/j.jsr.2012.10.010.

Shults, R. A., Bergen, G., Smith, T. J., Cook, L., Kindelberger, J., & West, B. (2019). Characteristics of single vehicle crashes with a teen driver in South Carolina, 2005–2008. *Accident Analysis & Prevention, 122*, 325–331. https://doi.org/10.1016/j.aap.2017.08.002.

Solomon, M. G., Compton, R. P., & Preusser, D. F. (2004). Taking the Click It or Ticket model nationwide. *Journal of Safety Research, 35*(2), 197–201. https://doi.org/10.1016/j.jsr.2004.03.003.

Strine, T. W., Beck, L. F., Bolen, J., Okoro, C., Dhingra, S., & Balluz, L. (2010). Geographic and sociodemographic variation in self-reported seatbelt use in the United States. *Accident Analysis & Prevention, 42*(4), 1066–1071. https://doi.org/10.1016/j.aap.2009.12.014.

Sung, J., Mizenko, K., & Coleman, H. (2017). *A Comparative Analysis of State Traffic Safety Countermeasures and Implications for Progress "Toward Zero Deaths" in the United States. (Report No. DOT HS 812 392).* National Highway Traffic Safety Administration.

Sunshine, J., Dwyer-Lindgren, L., Chen, A., & Mokdad, A. H. (2017). Seat-belt use In US counties: Limited progress toward healthy people 2020 objectives. *Health Affairs; Chevy Chase, 36*(4), 636–639.

Thomas, A. M., Cook, L. J., & Olson, L. M. (2011). Evaluation of the Click It or Ticket intervention in Utah. *Accident Analysis & Prevention, 43*(1), 272–275. https://doi.org/10.1016/j.aap.2010.08.022.

Uniform Criteria for State Observational Surveys of Seatbelt Use. (2011, April 1). Federal Register. <https://www.federalregister.gov/documents/2011/04/01/2011-7632/uniform-criteria-for-state-observational-surveys-of-seat-belt-use>.

Van Houten, R., Hilton, B., Schulman, R., & Reagan, I. (2011). Using Accelerator Pedal force to increase seatbelt use of service vehicle drivers. *Journal of Applied Behavior Analysis, 44*(1), 41–49. https://doi.org/10.1901/jaba.2011.44-41.

Van Houten, R., Malenfant, J. E. L., Reagan, I., Sifrit, K., Compton, R., & Tenenbaum, J. (2010). Increasing seatbelt use in service vehicle drivers with a gearshift delay. *Journal of Applied Behavior Analysis, 43*(3), 369–380. https://doi.org/10.1901/jaba.2010.43-369.

Vasudevan, V., Nambisan, S. S., Singh, A. K., & Pearl, T. (2009). Effectiveness of media and enforcement campaigns in increasing seatbelt usage rates in a state with a secondary seatbelt law. *Traffic Injury Prevention, 10*(4), 330–339. https://doi.org/10.1080/15389580902995190.

Wells, J. K., Williams, A. F., & Farmer, C. M. (2002). Seatbelt use among African Americans, Hispanics, and Whites. *Accident Analysis & Prevention, 34*(4), 523–529. https://doi.org/10.1016/S0001-4575(01)00050-1.

Williams, A. F., Wells, J. K., McCartt, A. T., & Preusser, D. F. (2000). "Buckle Up NOW!" an enforcement program to achieve high belt use. *Journal of Safety Research, 31*(4), 195–201. https://doi.org/10.1016/S0022-4375(00)00036-0.

Dr. Tracy Buchman is an Assistant Professor and Program Coordinator of Safety Internships in the Occupational and Environmental Safety and Health Department at the University of Wisconsin Whitewater where she teaches undergraduate and

graduate courses in occupational safety and health, healthcare safety, and emergency management. She has almost thirty years of professional practice, research, and teaching in the field of Safety and is an advisory board member of the University of Illinois at Chicago Occupational and Environmental Health and Safety Education and Research Center. Her areas of interest include traffic safety, risk reduction and sustainability.

JOURNAL OF SAFETY RESEARCH

Editor-in-Chief: **Thomas W. Planek,** National Safety Council, USA
Associate Editors: **Kenneth P. Kolosh,** National Safety Council, USA
Sergey Sinelnikov, National Safety Council, USA
Jonathan Thomas, National Safety Council, USA
Managing Editor: **Kathleen Porretta,** National Safety Council, USA

Journal of Safety Research 78 (2021) 221–228

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Effectiveness of the CDC HEADS UP online training on healthcare providers' mTBI knowledge and self-efficacy

Kelly Sarmiento *, Jill Daugherty, Dana Waltzman

Centers for Disease Control and Prevention (CDC), National Center for Injury Prevention and Control (NCIPC), Division of Injury Prevention, Atlanta, GA, United States

A R T I C L E I N F O

Article history:
Received 10 December 2020
Received in revised form 25 January 2021
Accepted 23 April 2021
Available online 7 May 2021

Keywords:
Concussion
Guideline
Physician
Nurse
Education
Training

A B S T R A C T

Background: Many healthcare providers do not consistently implement recommendations contained in clinical guidelines on mild traumatic brain injury (mTBI). As such, the Centers for Disease Control and Prevention (CDC) created the HEADS UP to Healthcare Providers online training to promote uptake of five key recommendations in the CDC Pediatric mTBI Guideline. *Methods:* Using data from modules in the CDC HEADS UP to Healthcare Providers online training, healthcare providers' self-reported knowledge and self-efficacy prior to and immediately following completion of the training was analyzed. *Results:* Improvements for 8 out of the 10 knowledge questions had a high level of practical significance. The knowledge question with the highest level of practical significance pre- to post-test improvement was for the key guideline recommendation on neuroimaging (pre-test correct: 70.2%; post-test correct: 87.8%; ($p < 0.0001$, Cohen's $g = 0.39$). Four out of the six questions had a self-efficacy level increase of a high level of practical significance ($r > 0.50$) between the pre- and post-tests. The self-efficacy question with pre- to post-test improvement with the highest level of practical significance was "I am confident in my ability to manage the return to sports progression for my patients" ($p < 0.001$; $r = 0.54$). *Conclusions:* The HEADS UP to Healthcare Providers online training led to significant improvements in knowledge and self-efficacy related to mTBI diagnosis and management. Expanded use of this training among healthcare providers who commonly provide care for pediatric patients with mTBI may be beneficial. *Practical Applications:* This study highlights several factors guideline developers may take into consideration when creating an implementation tool, such as using health behavior theories, working with partners and key stakeholders, and focusing on digital-based tools.

1. Introduction

To provide comprehensive guidance to healthcare providers who care for pediatric patients with mild traumatic brain injury (mTBI), the Centers for Disease Control and Prevention (CDC) published an evidence-based guideline on mTBI diagnosis, prognosis, management, and treatment in 2018 (Lumba-Brown et al., 2018). As demonstrated in the guideline, the science and clinical recommendations regarding the diagnosis and management of pediatric patients with mTBI have evolved substantially over the last two decades. Despite this progress, previous studies have found that many healthcare providers do not consistently implement recommendations contained in clinical guidelines on mTBI (Carl & Kinsella, 2014; Greene, Kernic, Vavilala, & Rivara, 2014; Melnick et al., 2012; Stache, Howell, & Meehan, 2016).

Challenges with implementation of evidence-based guidelines into clinical practice are not unique to mTBI. While a plethora of clinical guidelines for a variety of health topics are available to healthcare providers, many patients receive treatment that is not based on scientific evidence (Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice, 2011). Several barriers to guideline implementation by healthcare providers have been identified (Fischer, Lange, Klose, Greiner, & Kraemer, 2016; Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice, 2011). Many of these barriers are believed to stem from an interaction between individual characteristics of the guideline (e.g., clarity, specificity, strength of the evidence), perceptions of healthcare providers (e.g., self-efficacy, perceived importance of the recommendations, relevance to practice) and practice environment or context-related characteristics (e.g., inpatient, ambulatory, long-term care setting) (Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice, 2011).

* Corresponding author at: 4770 Buford Hwy NE, Atlanta, GA 30341, United States.

E-mail address: KSarmiento@cdc.gov (K. Sarmiento).

https://doi.org/10.1016/j.jsr.2021.04.004
0022-4375/

Reaching healthcare providers with information on mTBI is an ongoing, yet critical challenge. CDC created a set of implementation tools (e.g., a checklist for healthcare providers and discharge instructions for patients and their families) to promote uptake of its recent clinical practice guideline on pediatric mTBI. An example of one such implementation tool is the HEADS UP to Healthcare Providers online training. Developed in partnership with the American Academy of Pediatrics (AAP), the training provides an overview of the evidence-based recommendations in the CDC Pediatric mTBI Guideline, as well as practical strategies to integrate these recommendations into clinical practice. The training provides a special emphasis on five key recommendations in the CDC guideline: when to use neuroimaging for pediatric patients with mTBI; use of validated, age-appropriate symptom scales to diagnose mTBI; the role of certain risk factors on prolonged recovery; instructing patients on return to activity customized to their symptoms; and how best to manage a patient's return to non-sports activities soon after the injury (Table 1).

Development of the HEADS UP to Healthcare Providers training was guided by constructs of the Health Belief Model (HBM) (Becker, 1974). First developed in the 1950s, the main assumption of the HBM is that an individual's beliefs (e.g., perceived risk and susceptibility to a particular condition), coupled with their perception of the benefits of a specific action, serve as drivers for behavior adoption or change. In the 1970s, the HBM was expanded to include a construct on self-efficacy, which is a person's perceived ability (efficacy expectations) to perform an action or task (Bandura, 1977). The use of self-efficacy as a valuable predictor of behavior change in the HBM, and other health theories, is well-documented in the literature (Leventhal, Meyer, & Nerenz, 1980; So, 2013). Moreover, self-efficacy is linked to and considered to be a precursor for behavior change related to guideline implementation (Fischer et al., 2016).

Thus, the purpose of this study is to assess the effectiveness of the HEADS UP to Healthcare Providers online training on healthcare providers' mTBI knowledge and self-efficacy related to the five key recommendations in the CDC Pediatric mTBI Guideline. Findings can be used to inform the development of additional implementation tools for the CDC Pediatric mTBI Guideline, as well as tools for guidelines that cover other health topics.

2. Methods

This study analyzed data obtained from pre- and post-test modules contained in the HEADS UP to Healthcare Providers training. First released in 2011, the training was revised and re-launched in 2018 to include a specific focus on the recommendations in the CDC Pediatric mTBI Guideline. The primary audience of the training is healthcare providers who care for pediatric patients with mTBI (e.g., pediatricians, family practice providers); however, the training is accessible and available to anyone at no cost from the CDC website. The pre-test module contains 11 questions on participant demographics and experience (e.g., provider type, practice location [based on zip code], and use of mTBI assessment tools), 10 knowledge questions, and 6 questions focused on self-efficacy related to diagnosis and management of pediatric mTBI. In the one-year time period that we analyzed, 19,208 individuals took the training. However, 5,493 people were dropped from the analysis due missing provider type, which was a key part of our inclusion criteria (i.e., desire to ensure that these were actual healthcare providers) and an additional 3,815 people were dropped due to failure to take any part of the post-test. This left us with 9,900 individuals with complete data. Individuals who did not self-identify in the pre-test module as a physician, nurse practitioner, or physician assistant were then excluded from the analysis (n = 7,795). The final sample included 2,105 healthcare providers: 781 physicians, 1,078 nurse practitioners, and 246 physician assistants. These groups were selected for the analysis as they are the key audiences for the CDC Pediatric mTBI Guideline.

Questions in the post-test module were identical to those in the pre-test; the sole exception being that questions on demographics were not included in the post-test module. The knowledge questions in the test modules were derived from and aligned to the five key recommendations in the CDC Pediatric mTBI Guideline (Table 1). Responses to the knowledge questions were re-coded for analysis such that correct responses = 1 and incorrect responses = 0. The self-efficacy questions were created to assess healthcare providers' confidence related to diagnosis and management of pediatric mTBI. The self-efficacy questions were each measured on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." Respondents received 1 to 5 points for each item, with higher scores indicating a greater level of agreement with each statement. Responses were condensed for presentation into three categories: disagree/strongly disagree, neither agree nor disagree, and strongly agree/agree. In addition, healthcare providers' knowledge and self-efficacy were analyzed using a scale from 10 pre-test knowledge questions and 6 pre-test self-efficacy questions. CDC determined that data collection was not subject to Institutional Review as the data were collected as part of the regular function of the training and designed for training improvement and evaluation.

2.1. Data analysis

Descriptive statistics for study variables were computed using the sample of healthcare providers who completed both the pre- and post-test between January 1 and December 31, 2019. McNemar's tests were computed to detect statistically significant differences between responses to the pre- and post-test knowledge questions, while Wilcoxon signed rank tests for paired observations were used to compare pre- and post-test self-efficacy questions. Medians were reported for the self-efficacy questions given the ordinal nature of the data. SAS version 9.4 (http://www.sas.com) and IBM SPSS were used to compute all statistics.

Effect sizes were computed for each McNemar's and Wilcoxon signed rank test. Cohen's g was computed for each McNemar's test; a g of less than 0.15 is considered to have a small effect size, a g between 0.15 and 0.25 has a medium effect size, and a g of 0.25 or greater has a large effect size (Cohen, 1988). Effect sizes (r) were also computed for each Wilcoxon signed rank test using the Z-score and interpreted in accordance with Cohen (1988). An r of 0.1 represents a small effect size, an r of 0.3 represents a medium effect size and an r of 0.5 represents a large effect size (Cohen, 1988). Medium and large effects were considered to indicate a practical or substantive change ("practical significance") between the pre-test and post-test. The number of missing data was negligible for the knowledge questions; between 0 and 16 (0.0–0.8%). The number of missing data was higher for the self-efficacy questions, ranging from 0 to 165 (0.0–7.8%). An internal analysis demonstrated that those with missing data were not significantly different than those with complete data.

3. Results

Most healthcare providers (65.4%) had worked 5 years or fewer in their profession, and roughly half (55.6%), evaluated a pediatric patient for a suspected mTBI in the 12 months preceding administration of the survey (Table 2). When asked how often they adhere to current evidence-based recommendations on mTBI related to decision and assessment tools and discharge instructions, 41.8%

Table 1
Knowledge questions contained in the Centers for Disease Control and Prevention's HEADS UP to Healthcare Providers online training pre- and post-test modules and their alignment with the five key recommendations in the CDC Pediatric Mild Traumatic Brain Injury (mTBI) Guideline.

Five Key Recommendations in the CDC Pediatric mTBI Guideline	Knowledge questions contained in the HEADS UP to Healthcare Providers pre- and post-test modules
1. Do not routinely image patients to diagnose mTBI.	• Which indications should prompt admission to a hospital for a patient with mTBI? A. Any signs of intracranial injury that require monitoring and repeat neurological exams. B. Fluctuating or deteriorating neurological or cognitive symptoms. C. Patient complains of trouble concentrating and feeling fatigued. D. **Both A and B.** E. All of the above. • A healthcare provider should order a head CT scan for patients with suspected mTBI: A. In all cases. B. **If indicated by a validated decision tool.** C. When requested by the patient's parents. D. Both B and C. E. Never.
2. Use validated, age-appropriate symptom scales to diagnose mTBI.	• A 14-year old patient hit her head while playing soccer. She presents with a headache and says she "just doesn't feel right." What is the appropriate course of action? A. Conduct a symptom assessment, and if she receives an acceptable score, provide approval for return to sports. B. Order a CT scan to evaluate for intracranial injury. C. **Assess her symptoms and concussion history and require that she be observed at home for 24 to 48 hours for signs of deteriorating neurological function.** D. Set a date no sooner than 14 days from today for her to return to play if her symptoms have lasted more than 30 minutes; 7 days if they lasted less than 30 minutes. • Which of the following is TRUE regarding validated mTBI symptom rating scales: A. They can be used as the sole diagnostic criteria. B. Only computer-based assessment scales are validated. C. **They should assess changes from a patient's usual or baseline symptom presentation.** D. Both B and C are true. E. All of the above are true.
3. Assess evidence-based risk factors for prolonged recovery.	• Which of the following factors are associated with a prolonged recovery from an mTBI? A. Neurological or psychiatric disorder. B. Higher cognitive ability. C. Older age (older children/adolescents). D. **Both A and C.** E. All of the above. • A healthcare provider should refer a patient for further evaluation by a specialist when: A. Problems with attention, memory and learning, response speed, and other cognitive impairments interfere with school. B. The patient experiences ongoing headaches 2 to 3 days after the injury. C. Sleep problems emerge or continue despite the patient engaging in appropriate sleep hygiene measures. D. **Both A and C.** E. All the above.
4. Provide patients with instructions on return to activity customized to their symptoms.	• What happens when an athlete's symptoms return after they've initiated a step-wise return to play progression? A. The athlete should progress to the next level so long as they do not lose consciousness, vomit or have problems with balance. B. **The athlete should drop back to the previous level at which they were asymptomatic and try to move forward only after a 24-hour period of rest has passed and they are again asymptomatic.** C. If they are on Step 1–3, the athlete should repeat the step-wise process, starting at Step 1. If they are on Step 4 – 5, they should stay on the same step another day. D. The athlete should rest for 10 to 14 days. • A 6-year-old boy diagnosed with mTBI complains of continuing headaches one week after the injury, but no other neurological symptoms. What is the appropriate action to manage his headaches? A. Refer him to neurology for baseline neurocognitive testing. B. **Recommend over-the-counter analgesics.** C. Recommend complete ("strict") physical and cognitive rest until he is asymptomatic. D. Both B and C. E. None of the above.
5. Counsel patients to return gradually to non-sports activities after no more than 2–3 days of rest.	• Prior to discharging a patient from the ED or your office, it is important to counsel patients and their parents that: A. Children are allowed to return to sports activities before school if they feel well enough. B. Most children will experience a prolonged recovery. C. Within a few weeks, the patient can begin non-strenuous activities that do not exacerbate symptoms. D. **They should monitor for signs of deteriorating neurological function.** E. All of the above. • When discharging a patient with mTBI, you should counsel patients and their parents to: A. Rest cognitively and physically for one to two weeks until they no longer experience symptoms. B. Give them approval to return to sports if their symptoms aren't severe. C. **Following one to two days of rest, gradually begin return to non-sports activity as long as symptoms do not worsen.** D. None of the above.

reported using decisions tools, 42.3% reported using standardized assessments, and 62.4% reported providing written discharge instructions "often or very often." To learn about clinical practice recommendations, most healthcare providers preferred seeking information from websites (e.g., UpToDate and Medscape) (78.6%), viewing presentations from experts (51.5%), and/or attending medical conferences (50.7%).

The percentage of correct responses for each of the knowledge questions increased significantly between the pre- and post-tests (Table 3). Improvements for 8 out of the 10 knowledge questions had a high level of practical significance. Of these, the questions with pre- to post-test improvement with the highest level of practical significance was for the key guideline recommendation on neuroimaging ("A healthcare provider should order a head computerized tomography (CT) scan for patients with suspected mTBI") (pre-test correct: 70.2%; post-test correct: 87.8%; $p < 0.0001$, Cohen's $g = 0.39$). Other knowledge questions that demonstrated improvements with a high level of practical significance included: "A 14-year old patient hit her head while playing soccer. She presents with a headache and says she 'just doesn't feel right.' What is the appropriate course of action?" (pre-test correct: 84.0%; post-test correct: 94.8%; $p < 0.0001$, Cohen's $g = 0.36$); "What happens when an athlete's symptoms return after they've initiated a step-wise return to play progression?" (pre-test correct: 83.5%; post-test correct: 95.3%; $p < 0.0001$, Cohen's $g = 0.36$); and, "A 6-year-old boy diagnosed with mTBI complains of continuing headaches one week after the injury, but no other neurological symptoms. What is the appropriate action to manage his headaches?" (pre-test correct: 30.5%; post-test correct: 50.0%; $p < 0.0001$, Cohen's $g = 0.34$). The two questions with improvements in knowledge with low and medium levels of significance included: "Which of the following is TRUE regarding validated mTBI symptom rating scales?"; pre-test correct: 74.9%; post-test correct: 77.8% ($p < 0.01$, Cohen's $g = 0.07$) and "Which indications should prompt admission to a hospital for a patient with mTBI?"; pre-test correct: 76.0%; post-test correct: 83.4% ($p < 0.0001$, Cohen's $g = 0.19$).

The level of self-efficacy measured for each of the six questions in the post-test demonstrated statistically significant improvements as compared to the pre-test (Table 4). All six questions had a self-efficacy level increase of a high level of practical significance ($r > 0.50$) between the pre- and post-tests. These questions were: "I am confident in my ability to diagnose an mTBI" ($p < 0.001$; $r = 0.62$), "I am confident in my ability to treat mTBI symptoms" ($p < 0.001$; $r = 0.63$), "I am confident in my ability to manage the return to sports progression for my patients" ($p < 0.001$; $r = 0.63$), "I am confident in my ability to manage return to school for my patients" ($p < 0.001$; $r = 0.63$)."I am confident in my ability to identify patients who should be referred for evaluation by an mTBI specialist" ($p < 0.001$; $r = 0.59$), and "I am confident in my ability to communicate with patients about mTBI prevention strategies" ($p < 0.001$; $r = 0.54$).

4. Discussion

This study examined the effectiveness of the HEADS UP to Healthcare Provider online training on improving healthcare providers' mTBI knowledge and self-efficacy related to the five key recommendations in the CDC Pediatric mTBI Guideline. Healthcare providers who completed the HEADS UP to Healthcare Providers online training not only demonstrated significant improvements in knowledge but also reported improved self-efficacy related to mTBI diagnosis and management. These findings suggest that the HEADS UP to Healthcare Providers online training may be an effective tool to support implementation of the CDC Pediatric mTBI Guideline. Expanded use among healthcare providers who care for pediatric patients with mTBI may be beneficial.

More than 19,000 people accessed the HEADS UP to Healthcare Providers online training during the study period. High use of this online training may be attributed to mTBI and concussion training requirements instituted by some health organizations, schools, and states, as well as the inclusion of continuing education credits available through AAP upon its completion (Fischer et al., 2016). Prior research suggests that continuing education opportunities and online training for healthcare providers on concussion is associated with improvements in clinical practice (Babul, Turcotte, Lambert, Hadly, & Sadler, 2020; Broshek, Samples, Beard, & Goodkin, 2014). An online training approach was used for the HEADS UP to Healthcare Providers training as it is a cost-effective approach that allowed organizations that require the training to disseminate it widely and provide flexibility for where (e.g., home or place of work) and when (during or outside of practice hours) a healthcare provider could complete their training requirement. As the use of online trainings to educate healthcare providers by CDC and other organizations have increased in popularity, evaluation of the effectiveness of this educational approach is critical. At least two systematic reviews concluded that training healthcare providers through online trainings in place of non-computer-based trainings (e.g., in-person presentations, lectures, and workshops) is equally effective in improving healthcare providers' knowledge and clinical behaviors (Cook et al., 2008; Richmond, Copsey, Hall, Davies, & Lamb, 2017). These findings suggest that health educators may consider the development of an online training as one component of a comprehensive approach for guideline implementation.

Healthcare providers with knowledge of and a high self-efficacy related to clinical recommendations may be more likely to adopt and adhere to guidelines (Fischer et al., 2016). Consistent with increases in self-efficacy related to mTBI diagnosis and management, some of the largest improvements in knowledge between pre- and post-tests were observed for the questions aligned with the CDC Pediatric mTBI Guideline recommendations on diagnostic use of neuroimaging and managing a patient's return to activity. Changing healthcare provider behaviors around CT scans for mTBI is an ongoing challenge (Halaweish, Riebe-Rodgers, Randall, & Ehrlich, 2018). Decreasing routine use of CT scans for patients with mTBI may help reduce adverse health outcomes related to radiation exposure (Mannix, Meehan, Monuteaux, & Bachur, 2012; Stanley et al., 2014); up to 35% of CT scans conducted in the emergency department for patients with mTBI may not be warranted based on clinical guidance (Melnick et al., 2012). While some clinical recommendations have been consistent for numerous years (such as that on neuroimaging), the CDC Pediatric mTBI Guideline recommendation of a gradual return to non-sports activities represents a shift in clinical care (Lumba-Brown et al., 2018). Previous guidance recommended a longer rest period; however, healthcare providers are now advised to instruct pediatric patients with mTBI to return to their regular non-sports activities within 2–3 days. As compared to prescribing "strict rest," this change in guidance is associated with a shorter recovery and a lower symptom burden (Thomas, 2015). Findings from this study indicate that training may show promise in furthering adoption of neuroimaging and return to activity recommendations that can improve patient health outcomes. However, additional studies are needed to assess the sustainability of these improvements and their translation into clinical practice.

Previous studies suggest that medical students and residents may not receive adequate training on mTBI diagnosis and management and that more educational opportunities on this topic are needed (Donaworth, Grandhi, Logan, Gubanich, & Myer, 2016; Haider et al., 2017). Donaworth and colleagues (2016) found most

Table 2
Background characteristics of respondents (n = 2,105) who completed the Centers for Disease Control and Prevention HEADS UP to Healthcare Providers online training, 2019.

	Frequency	Percent
Healthcare provider type		
Physician	781	37.1
Nurse practitioner	1,078	51.2
Physician assistant	246	11.7
Total	**2,105**	**100.0**
Number of years in practice		
0–5	1,377	65.4
6–10	207	9.8
11–20	284	13.5
21–30	166	7.9
31+	71	3.4
Total	**2,105**	**100.0**
Percentage of practice that is pediatric		
0–25%	1,145	54.5
26–50%	407	19.4
51–75%	38	1.8
76+%	511	24.3
Total	**2,101**	**100.0**
Have you evaluated a patient for a suspected mild traumatic brain injury (mTBI) in the previous 12 months		
Yes	1,166	55.6
No	848	40.5
Unsure	82	3.9
Total	**2,096**	**100.0**
Healthcare provider uses decision tools to evaluate for mTBI in their practice		
Very often	418	19.9
Often	458	21.9
Sometimes	613	29.3
Never	607	29.0
Total	**2,096**	**100.0**
Healthcare provider uses standardized assessments of concussion in their practice		
Very often	466	22.2
Often	422	20.1
Sometimes	607	29.0
Never	601	28.7
Total	**2,096**	**100.0**
Healthcare provider provides written discharge instructions for patients with mTBI		
Very often	865	41.3
Often	443	21.1
Sometimes	354	16.9
Never	434	20.7
Total	**2,096**	**100.0**
How healthcare provider prefers to learn about clinical practice recommendations[a]		
Websites (like UpToDate and Medscape)	1,654	78.6
Presentations from experts (such as Grand Rounds)	1,085	51.5
Medical conferences	1,067	50.7
Scientific publications	998	47.4
Medical organizations	888	42.2
Blogs and social media	275	13.1

[a] Respondents were permitted to select multiple responses to this question, therefore the total adds up to over 100%

U.S. medical school curriculums do not include lectures on concussion, and that the majority of medical students do not gain clinical experience with diagnosis and management of concussion during their medical school training. Interestingly, approximately two-thirds of healthcare providers who completed the CDC training reported working five years or fewer in their profession. A desire to learn about concussion, widespread use of digital or mobile-based tools, and use of social media to promote the training, may be some of reasons why the training was accessed more frequently by newer healthcare providers (Donaworth et al., 2016; Ventola, 2014). Taken together, this points to the potential of the HEADS UP to Healthcare Provider training to help fill a current information gap for healthcare providers new to their profession.

This study only measured changes in knowledge and self-efficacy based on constructs of the HBM prior to and immediately following the HEADS UP to Healthcare Providers training. As such, environmental and individual provider characteristics (a limitation of the HBM) were not taken into consideration (Janz & Becker, 1984). Moreover, as noted above, the long-term effectiveness on healthcare providers' knowledge and self-efficacy was not measured. Ensuring successful and sustained improvements among healthcare providers following use of the training may benefit further from a multi-pronged approach that is inclusive of system-based changes (e.g., use of electronic health records (EHR)) and support from decision-makers (Campanella et al., 2016). Previous studies suggest that integrating an online training into a comprehensive implementation effort may lead to improved patient-healthcare provider communication and symptom-based assessments (Arbogast et al., 2017). Arbogast and colleagues found that implementation of a concussion-specific EHR-based decision support tool, along with use of the HEADS UP to Healthcare Provider online training, substantially increased documentation of healthcare provider-patient discussions about recovery (e.g., return to school and sports; Arbogast et al., 2017). This is consistent with other studies that found that EHR-based systems may strengthen guideline adherence among healthcare providers (Campanella et al., 2016).

4.1. Practical applications

Guideline efforts may be inclusive of development, dissemination, and implementation planning (Fischer et al., 2016). Yet, approximately one-third of guidelines published between 2010 and 2017 did not offer guideline implementation tools (Liang, Abi Safi, & Gagliardi, 2017). Including health educators' and other public health professionals' participation in guideline development may help to ensure that implementation strategies are considered while recommendations are drafted. This may include ensuring recommendations are written with patient-centered language and practical strategies (Fischer et al., 2016).

Guideline implementation tools may improve healthcare providers' adherence to guideline recommendations (Liang, Abi Safi, et al., 2017). This study highlights several factors guideline developers may take into consideration when creating such tools. First, guideline implementation tools developed using a theoretical framework, such as the online training examined in this paper, are considered to be most effective (Liang, Bernhardsson, et al., 2017). Guideline developers can ensure that relevant implementation tools (customized for both patients and healthcare providers) are designed using health behavior theories and tested to assess their ability to support evidence-based patient care (Fischer et al., 2016; Liang, Abi Safi, et al., 2017). The HEADS UP to Healthcare Providers training took advantage of constructs of adult learning theory and components of the HBM by integrating interactive knowledge checks and content aimed at building healthcare provider's self-efficacy and promoting positive learning outcomes (Cook, Levinson, & Garside, 2010). Second, working with partners and key stakeholders to promote and disseminate online trainings, such as this one, can help improve training use among the target audiences and has benefitted other aspects of the HEAD UP campaign. Finally, the widespread usage of digital based tools and preferences for seeking out clinical information from websites suggests that a shift from print materials to online, tablet, and mobile-based formats may improve uptake and adherence to guidelines (Gagliardi & Alhabib, 2015; Ventola, 2014). Online training, which can be a cost-effective approach to reach a large audience, may be one important way to achieve this.

Table 3
Comparison of pre- and post-test concussion-related knowledge questions contained in the Centers for Disease Control and Prevention HEADS UP to Healthcare Providers online training, 2019 (n = 2,105).

	Pre-test	Post-test	Difference		
	Percent Correct[b]	Percent Correct	s	p-value	Cohen's g
Guideline recommendation 1: Do not routinely image patients to diagnose mTBI					
Which indications should prompt admission to a hospital for a patient with mTBI?			59.00	<0.0001	0.19
Correct	76.0	83.4			
Incorrect	24.0	16.6			
A healthcare provider should order a head CT scan for patients with suspected mTBI:			288.14	<0.0001	0.39
Correct	70.2	87.8			
Incorrect	29.8	12.2			
Guideline recommendation 2: Use validated, age-appropriate symptom scales to diagnose mTBI					
A 14-year old patient hit her head while playing soccer. She presents with a headache and says she "just doesn't feel right." What is the appropriate course of action?			163.71	<0.0001	0.36
Correct	84.0	94.8			
Incorrect	16.0	5.2			
Which of the following is TRUE regarding validated mTBI symptom rating scales?			7.47	0.0063	0.07
Correct	74.9	77.8			
Incorrect	25.1	22.2			
Guideline recommendation 3: Assess evidence-based risk factors for prolonged recovery					
Which of the following factors are associated with a prolonged recovery from an mTBI?			159.05	<0.0001	0.25
Correct	59.7	75.2			
Incorrect	40.3	24.9			
A healthcare provider should refer a patient for further evaluation by a specialist when:			281.03	<0.0001	0.31
Correct	48.5	69.8			
Incorrect	51.5	30.2			
Guideline recommendation 4: Provide patients with instructions on return to activity customized to their symptoms.					
What happens when an athlete's symptoms return after they've initiated a step-wise return to play progression?			179.97	<0.0001	0.36
Correct	83.5	95.3			
Incorrect	16.6	4.8			
A 6-year-old boy diagnosed with mTBI complains of continuing headaches one week after the injury, but no other neurological symptoms. What is the appropriate action to manage his headaches?			279.34	<0.0001	0.34
Correct	30.5	50.0			
Incorrect	69.5	50.0			
Guideline recommendation 5: Counsel patients to return gradually to non-sports activities after no more than 2–3 days of rest.					
Prior to discharging a patient from the ED or your office, it is important to counsel patients and their parents that:			103.35	<0.0001	0.27
Correct	80.4	89.4			
Incorrect	19.6	10.6			
When discharging a patient with mTBI, you should counsel patients and their parents to:			189.72	<0.0001	0.29
Correct	68.5	84.3			
Incorrect	31.5	15.7			

The questions are displayed as the percentage of respondents who answered the question correctly (i.e., answered a true question as true or a false question as false or selected the correct response for the multiple-choice item).

4.2. Limitations

There are several limitations to this study. First, data from the pre- and post-test modules were obtained from a convenience sample of healthcare providers who completed the HEADS UP to Healthcare Providers training. Thus, the findings are not intended to be generalizable to a wider population. Second, this study did not have a control group of healthcare providers who did not complete the training with which the authors could compare the results. Thus, it is unclear if other sources of information identified as commonly used by healthcare providers in the study (e.g., websites and presentations from experts) are similarly effective. Future studies may explore this. Third, the pre- and post-test knowledge questions were composed of multiple choice and true/false questions. This may have led to an overestimate in the level of knowledge of the respondents, as respondents had a 25–50% chance of randomly guessing the correct response. Fourth, social desirability may play a role in the respondents' answers to questions related to self-efficacy. It is likely that the respondents know what the "correct" response is or what "should" be the answer, particularly with the self-efficacy items. This may inflate both the pre- and post-test agreement with these items; however, it may not impact changes in responses observed between the pre- and post-tests. Finally, most healthcare providers completed the post-test immediately after taking the training. Thus, it is not possible to determine whether changes in knowledge and self-efficacy will persist over time. Further, this study did not evaluate actual changes in patient care. Thus, it is not possible to determine whether an individual's gains in knowledge and self-efficacy will translate into changes in their clinical care practices. Future research that examines whether the training led to actual improvements in diagnosis and management decisions may be beneficial.

5. Conclusion

This study examined the effectiveness of the HEADS UP to Healthcare Providers online training on healthcare providers' mTBI

Table 4
Pre- and post-test concussion self-efficacy questions contained in the Centers for Disease Control and Prevention HEADS UP to Healthcare Providers online training, 2019 (n = 2,105).

	Pre-test				Post-test						
	Frequency	Percent	Median	Median IQR	Frequency	Percent	Median	Median IQR	*Z*-score	*p*-value	*R*
I am confident in my ability to diagnose an mTBI											
Strongly agree/agree	1,259	64.5	4.0	3–4	I am confident in my ability to diagnose an	95.8	4.0	4–5	−27.3	<0.001	−0.62
Neither agree nor disagree	446	22.9			80	3.8					
Disagree/strongly disagree	247	12.7			8	0.4					
I am confident in my ability to treat mTBI symptoms											
Strongly agree/agree	1,178	60.5	4.0	3–4	1,975	93.8	4.0	4–5	−27.6	<0.001	−0.63
Neither agree nor disagree	528	27.1			122	5.8					
Disagree/strongly disagree	240	12.3			8	0.4					
I am confident in my ability to manage the return to sports progression for my patients											
Strongly agree/agree	1,155	59.2	4.0	3–4	1,981	94.1	4.0	4–5	−27.8	<0.001	−0.63
Neither agree nor disagree	506	25.9			112	5.3					
Disagree/strongly disagree	291	14.9			12	0.6					
I am confident in my ability to manage return to school for my patients											
Strongly agree/agree	1,256	64.5	4.0	3–4	2,012	95.6	4.0	4–5	−27.7	<0.001	−0.63
Neither agree nor disagree	454	23.3			88	4.2					
Disagree/strongly disagree	236	12.1			5	0.2					
I am confident in my ability to identify patients who should be referred for evaluation by an mTBI specialist											
Strongly agree/agree	1,360	69.7	4.0	3–4	2,033	96.6	4.0	4–5	−26.2	<0.001	−0.59
Neither agree nor disagree	408	20.9			68	3.2					
Disagree/strongly disagree	184	9.4			4	0.2					
I am confident in my ability to communicate with patients about mTBI prevention strategies											
Strongly agree/agree	1,514	78.0	4.0	4–5	2,045	97.2	5.0	4–5	−23.8	<0.001	−0.54
Neither agree nor disagree	309	15.9			56	2.7					
Disagree/strongly disagree	117	6.0			4	0.2					

For all attitude items, "strongly agree"=5, "agree"=4, "neither agree nor disagree"=3, "disagree"=2, and "strongly disagree"=1.
IQR = Inter-quartile range.

knowledge and self-efficacy related to the five key recommendations in the CDC Pediatric mTBI Guideline. Findings suggest that upon completion of the training, healthcare providers demonstrated significant improvements in knowledge related to the five key recommendations in the guideline, as well as improvements in self-efficacy related to mTBI diagnosis and management. Expanding use of this training may be an effective way to reach a large number of healthcare providers and improve use of recommendations in the CDC Pediatric mTBI Guideline.

6. Financial disclosure

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors have indicated they have no potential conflicts of interest to disclose.

8. Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

Arbogast, K. B., Curry, A. E., Metzger, K. B., Kessler, R. S., Bell, J. M., Haarbauer-Krupa, J., ... Master, C. L. (2017). Improving primary care provider practices in youth concussion management. *Clinical Pediatrics (Phila), 56*(9), 854–865. https://doi.org/10.1177/0009922817709555.

Babul, S., Turcotte, K., Lambert, M., Hadly, G., Sadler, K., & Swaine, I. L. (2020). Delivering evidence-based online concussion education to medical and healthcare professionals: The Concussion Awareness Training Tool (CATT). *Journal of Sports Medicine, 2020*, 1–10. https://doi.org/10.1155/2020/8896601.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037//0033-295x.84.2.191.

Becker, M. H. (1974). The health belief model and personal health behavior. *Health Education Monographs, 2*, 324–473.

Broshek, D. K., Samples, H., Beard, J., & Goodkin, H. P. (2014). Current practices of the child neurologist in managing sports concussion. *Journal of Child Neurology, 29*(1), 17–22. https://doi.org/10.1177/0883073812464525.

Campanella, P., Lovato, E., Marone, C., Fallacara, L., Mancuso, A., Ricciardi, W., & Specchia, M. L. (2016). The impact of electronic health records on healthcare quality: A systematic review and meta-analysis. *Eur J Public Health, 26*(1), 60–64. https://doi.org/10.1093/eurpub/ckv122.

Carl, R. L., & Kinsella, S. B. (2014). Pediatricians' knowledge of current sports concussion legislation and guidelines and comfort with sports concussion management: A cross-sectional study. *Clinical Pediatrics (Phila), 53*(7), 689–697. https://doi.org/10.1177/0009922814526979.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: NY Routledge.

Cook, D. A., Levinson, A. J., & Garside, S. (2010). Time and learning efficiency in Internet-based learning: A systematic review and meta-analysis. *Advances in Health Sciences Education, 15*(5), 755–770. https://doi.org/10.1007/s10459-010-9231-x.

Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2008). Internet-based learning in the health professions: A meta-analysis. *JAMA, 300*(10), 1181–1196. https://doi.org/10.1001/jama.300.10.1181.

Donaworth, M. A., Grandhi, R. K., Logan, K., Gubanich, P. J., & Myer, G. D. (2016). Is current medical education adequately preparing future physicians to manage concussion: An initial evaluation. *The Physician and Sportsmedicine, 44*(1), 1–7. https://doi.org/10.1080/00913847.2016.1135039.

Fischer, F., Lange, K., Klose, K., Greiner, W., & Kraemer, A. (2016). Barriers and strategies in guideline implementation-a scoping review. *Healthcare (Basel), 4* (3), 36. https://doi.org/10.3390/healthcare4030036.

Gagliardi, A. R., & Alhabib, S. (2015). Trends in guideline implementation: A scoping systematic review. *Implementation Science, 10*, 54. https://doi.org/10.1186/s13012-015-0247-8.

Greene, N. H., Kernic, M. A., Vavilala, M. S., & Rivara, F. P. (2014). Variation in pediatric traumatic brain injury outcomes in the United States. *Archives of Physical Medicine & Rehabilitation, 95*(6), 1148–1155.

Haider, M. N., Leddy, J. J., Baker, J. G., Kiel, J. M., Tiso, M., Ziermann, K. A., & Willer, B. S. (2017). Concussion management knowledge among residents and students and how to improve it. Concussion (London, England), 2(3), CNC40-CNC40. 10.2217/cnc-2017-0001.

Halaweish, I., Riebe-Rodgers, J., Randall, A., & Ehrlich, P. F. (2018). Compliance with evidence-based guidelines for computed tomography of children with head and abdominal trauma. *Journal of Pediatric Surgery, 53*(4), 748–751. https://doi.org/10.1016/j.jpedsurg.2017.07.008.

Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice, G. (2011). In R. Graham, M. Mancher, D. Miller Wolman, S. Greenfield, & E. Steinberg (Eds.), Clinical Practice Guidelines We Can Trust. Washington (DC): National Academies Press (US). Copyright 2011 by the National Academy of Sciences. All rights reserved.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly, 11*(1), 1–47. https://doi.org/10.1177/109019818401100101.

Leventhal, H., Meyer, D., & Nerenz, D. (1980). The common sense representation of illness danger. *Contributions to Medical Psychology, 2*, 7–30.

Liang, L., Abi Safi, J., & Gagliardi, A. R. (2017a). Number and type of guideline implementation tools varies by guideline, clinical condition, country of origin, and type of developer organization: Content analysis of guidelines. *Implementation Science, 12*(1), 136. https://doi.org/10.1186/s13012-017-0668-7.

Liang, L., Bernhardsson, S., Vernooij, R. W., Armstrong, M. J., Bussieres, A., Brouwers, M. C., & Gagliardi, A. R. (2017b). Use of theory to plan or evaluate guideline implementation among physicians: A scoping review. *Implementation Science, 12*(1), 26. https://doi.org/10.1186/s13012-017-0557-0.

Lumba-Brown, A., Yeates, K. O., Sarmiento, K., Breiding, M. J., Haegerich, T. M., Gioia, G. A., ... Timmons, S. D. (2018). Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury among children. *JAMA Pediatrics, 172*(11), e182853. https://doi.org/10.1001/jamapediatrics.2018.2853.

Mannix, R., Meehan, W. P., Monuteaux, M. C., & Bachur, R. G. (2012). Computed tomography for minor head injury: Variation and trends in major United States pediatric emergency departments. *Journal of Pediatrics, 160*(1), 136–139.e131. https://doi.org/10.1016/j.jpeds.2011.06.024.

Melnick, E. R., Szlezak, C. M., Bentley, S. K., Dziura, J. D., Kotlyar, S., & Post, L. A. (2012). CT overuse for mild traumatic brain injury. *The Joint Commission Journal on Quality and Patient Safety, 38*(11), 483–489. https://doi.org/10.1016/s1553-7250(12)38064-1.

Richmond, H., Copsey, B., Hall, A. M., Davies, D., & Lamb, S. E. (2017). A systematic review and meta-analysis of online versus alternative methods for training licensed health care professionals to deliver clinical interventions. BMC Medical Education, 17(1), 227-227. 10.1186/s12909-017-1047-4

So, J. (2013). A further extension of the Extended Parallel Process Model (E-EPPM): Implications of cognitive appraisal theory of emotion and dispositional coping style. *Health Communication, 28*(1), 72–83. https://doi.org/10.1080/10410236.2012.708633.

Stache, S., Howell, D., & Meehan, W. P. 3rd., (2016). Concussion management practice patterns among sports medicine physicians. *Clinical Journal of Sport Medicine, 26*(5), 381–385. https://doi.org/10.1097/jsm.0000000000000270.

Stanley, R. M., Hoyle, J. D., Dayan, P. S., Atabaki, S., Lee, L., Lillis, K., ... Kuppermann, N. (2014). Emergency department practice variation in computed tomography use for children with minor blunt head trauma. *Journal of Pediatrics, 165*(6), 1201–1206.e2. https://doi.org/10.1016/j.jpeds.2014.08.008.

Thomas, D. G., Apps, J. N., Hoffmann, R. G., McCrea, M., & Hammeke, T. (2015). Benefits of strict rest after acute concussion: a randomized controlled trial. Pediatrics, 135(2), 213–223. https://doi.org/10.1542/peds.2014-0966

Ventola, C. L. (2014). Mobile devices and apps for health care professionals: Uses and benefits. *P&T, 39*(5), 356–364.

Kelly Sarmiento, MPH is a public health advisor at the Centers for Disease Control and Prevention's (CDC) Injury Center in Atlanta, Georgia. Over the last 15 year at CDC's Injury Center, Kelly has worked on multiple national and award-winning educational campaigns and initiatives that focus on raising awareness and improving prevention, recognition, and management of traumatic brain injuries, including concussion. Kelly has a Master of Public Health from Yale University School of Epidemiology and Public Health, and a BA in Anthropology and a BA in Spanish from the University of California, Santa Barbara.

Jill Daugherty, PhD, MPH is an epidemiologist at the Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Division of Injury Prevention. She earned her doctorate in sociology from Emory University in Atlanta where she focused on social inequality. Her work at the CDC focuses on disparities in traumatic brain injury and concussion knowledge and prevention. Previously she was employed at CDC's National Center for Health Statistics in Maryland.

Dana Waltzman, PhD, is a behavioral scientist on the Traumatic Brain Injury (TBI) Team in the Division of Injury Prevention (DIP) at the Injury Center. Her work focuses on understanding the public health burden, prevention, and long-term effects of TBI. Dr. Waltzman received a bachelor of science degree in psychology, a bachelor of arts degree in cognitive science from the University of Georgia, and a doctoral degree in cognitive neuroscience at the University of California, Los Angeles. She completed two post-doctoral fellowships at the Memory and Aging Center at the University of California, San Francisco, and a joint fellowship at Stanford University and the Department of Veterans Affairs, Palo Alto.

Journal of Safety Research 78 (2021) 270–275

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Evaluating 24/7 Sobriety Program participant reoffense risk

Kimberly Vachal[a,*], Shantanu Awasthi[b], Ihsan Ullah Khan[a], Yun Zhou[a], Bong-Jin Choi[c], Seguy Tchakounte-Wakem[a]

[a] Department of Transportation and Logistics, North Dakota State University, Fargo, ND, United States
[b] Department of Mathematics, North Dakota State University, Fargo, ND, United States
[c] Department of Statistics, North Dakota State University, Fargo, ND, United States

ARTICLE INFO

Article history:
Received 10 January 2021
Received in revised form 19 March 2021
Accepted 21 June 2021
Available online 17 July 2021

Keywords:
Survival analysis
Alcohol-impaired driving
Stratified Cox regression model

ABSTRACT

Objective: Our study investigated risk factors in survival among a subpopulation of drivers in North Dakota's 24/7 Sobriety Program. Participants mandated for a second driving-under-the-influence of alcohol (DUI) arrest were studied for a three-year interval that commenced with the start date for a 360-day enrollment. *Method:* A Stratified Cox regression model was developed to compute the hazard ratios for survival. A subsequent DUI-related offense as event of interest. Relation to the explanatory variable array that could be construed from administrative records were investigated. *Results:* Older drivers were 6.31 times more likely to reoffend than the younger driver cohort of 18–35-years. The survival curve slope showed the fastest decline in the 361-day to 730-day interval. Neither gender nor residence region was a significant predictor in DUI reoffense over the three-year monitoring interval. Preliminary work suggests reoffense was more likely if an individual had program history prior to this court mandated 360-day term in the 24/7 Sobriety Program for a second DUI. The program experience finding was unexpected but could not be studied in greater detail due to data and resource limitations. *Conclusions:* Administrative records access created a novel opportunity to explore an evolving impaired driving prevention strategy that has shown early promise. Individual driver survival in and after the 24/7 Sobriety Program was studied for three-years. Findings show age, post-program time interval, and possibly program history as areas to explore to improve survival rates. Driver DUI offense were most common shortly after program completion. Although limited to a single state, findings increase knowledge for refining strategies designed to impact driver subpopulations at higher risk for reoffense.

1. Introduction

Across the United States nearly 11,000 people died and more than 300,000 were injured in alcohol-impaired driving crashes in 2017 (NHTSA, 2019). Over one million driving under the influence of alcohol (DUI) arrests were made in 2019 (Federal Bureau of Investigation, xxxx). While many were first-time offenders, about one in three had a prior DUI offense. Therefore, an improved understanding of strategies to reduce reoffense propensity among this prevalent impaired-driver subpopulation would produce more effective resource and policy decisions in preventing DUI reoffense.

Generally, DUI prevention involves multiple facets including policy, enforcement, and education strategies. Policy deterrents such as raising the minimum legal drinking age and lowering the legal BAC thresholds have had positive impacts as universal population strategies (Carpenter et al., 2007; Miron & Tetelbaum, 2007; NHTSA, 2005; Tippetts et al., 2005; Wagenaar et al., 2007; Fell et al. 2008). Education programs and enforcement campaigns continue to be widely coupled in supporting these policies. Research shows, however, only about 1 in 1,000 result in arrest (Zaloshnja, Miller, & Blincoe, 2013). Technological interventions, such as ignition interlock system (IIS), have also been proven effective (Bjerre, 2005; Elder et al., 2011; Marques et al., 2010; McCartt et al., 2013; McGinty et al., 2017, Rauch et al., 2002) but have not been adopted in several states (National Conference of State Legislatures, xxxx). Innovative solutions are continually sought and vetted but DUI-related offenses remain a chronic public safety issue (Ferguson, 2012; Wagenaar et al., 2007).

A novel prevention approach couples policy and technology to elicit driver self-responsibility is the 24/7 Sobriety Program After a DUI arrest, drivers that commit to refrain from alcohol while in the program may obtain a temporary restricted driver license (Kleiman et al., 2008; Loudenburg et al., 2010). Participant compliance is monitored with methods such as preliminary breath tests and continuous transdermal monitoring for alcohol presence. Indi-

* Corresponding author.
E-mail address: kimberly.vachal@ndsu.edu (K. Vachal).

https://doi.org/10.1016/j.jsr.2021.06.012
0022-4375/

viduals found noncompliant are taken into custody and may face an immediate incarceration. The program is typically focused on preventing DUI-related violations among repeat offenders. Early studies have shown effectiveness in reduced DUI recidivism among participants (Loudenburg et al., 2010; Midgette & Kilmer, 2015). The U.S. Department of Justice labeled the 24/7 Sobriety Program as promising (Midgette & Kilmer, 2015).

As programs designed to reduce DUI events among driver subpopulation evolve, increasing knowledge about reoffense propensity is key. The goal here was to investigate independent factor influences in reoffense among 24/7 Sobriety Program participants in North Dakota. Interest was focused on drivers mandated to the program after a second DUI arrest. State administrative records for the 24/7 Sobriety Program, driver license, and traffic crash were codified to define the survival interest event and independent factors for the analysis.

2. Methodology and data

Survival analysis was used to study participant program success in terms of DUI reoffense events. Survival analysis is an econometric technique used in analysis of time-dependent phenomena. The method is commonly employed to analyzing data where the outcome variable is a time interval defined by a common origin time point and occurrence time point for an event of interest. It has been widely used in engineering, medicine, and economics fields (Zhao et al., 2015). In the transportation field, models have been applied in various aspects including pedestrian risk exposure, motor and non-motor vehicles transaction analysis, incident duration analysis, and crash analysis (Bagloee & Asadi, 2016; Li, 2015; Tiwari et al., 2007; Zhao et al., 2015).

2.1. Method

General objectives of survival analysis include examination of event pattern times, the comparison of survival time distribution among subgroups, and examining how factors affect the risk for an event of interest. Survival time is determined by defining two time points: the time of origin and the time of failure. Survival analysis can be performed using three main methods: fully parametric, semi-parametric, and non-parametric. In parametric survival analysis, the model is constructed by performing regression analysis on the assumption that the outcome variables follow a well-known distribution. Non-parametric and semi-parametric models are more suitable when little information about the underlying distribution is available due to the lack of specific distribution or small size of the sample (Washington et al., 2010).

Survival analysis possesses an important characteristic that distinguishes it from other statistical techniques as data are usually incomplete or censored. Censoring refers to incomplete information about the survival time of some participants. This occurs in situations such as the study terminating before the participant experiences an event of interest or the participant leaving the study. Three important functions are the survival, hazard, and cumulative hazard may be considered (Liu, 2012; Machin et al., 2006; Washington et al., 2010). Survival function is a key term along with censoring and event. It is defined as the probability of the outcome event not occurring up to a specific point in time, including the time point of observation *(t)*, and is denoted by $S(t)$, which gives the probability that the random variable T exceeds the specified time t. The survival (or survivor) function can be written as:

$$S(t) = P(T > t) = \int_{t}^{\infty} f(x)dx \tag{1}$$

where $S(t)$ is a non-increasing function with a value of 0 when time tends to be infinity and 1 when time is zero.

The hazard function is the instantaneous rate at which events occur for individuals which are surviving at time t,

$$\boldsymbol{h}(t) = \lim_{\delta t \to 0^{+}} \frac{P(t \leq T < t + \delta t \geq t | T \geq t)}{\delta t} = \frac{f(t)}{S(t)} = -\frac{d}{dt} S(t) \tag{2}$$

where $\boldsymbol{h}$ is the hazard function and the cumulative hazard function is given as:

$$\boldsymbol{H}(t) = \int_{0}^{t} h(u)du. \tag{3}$$

where $\boldsymbol{H}$ is the cumulative hazard function. The cumulative hazard function is related to the survival function as follows:

$$S(t) = exp[-\int_{0}^{t} h(u)du] = exp^{-H(t)}. \tag{4}$$

Thus, the higher the hazard, the lower the survival.

Independent variable strata were analyzed for proportionality hazards (PH) assumption. The inhomogeneity hazard functions considered subsets of a group with the same characteristic. It is detailed in the results section.

2.2. Data

Data were collected from state agencies' administrative records related to the 24/7 Sobriety Program, driver license, and crash record systems from 2008 to 2018. The records were obtained under limited-use agreements with the respective agencies. These disparate datasets were linked and merged into pseudo driver histories for 24/7 Sobriety Program participants. The final dataset was parsed to retain de-identified records for drivers entering the program for a second DUI conviction.

The 24/7 Sobriety Program participant population database held 18,697 records in a complete program history covering January 2008 to December 2018.[1] During a three-year program entry period between 2013 and 2015, 1,287 participants, ages 18-years or older that started the program for a second DUI offense, were selected for the study. This participant group was defined to minimize effects of legislated program changes that took effect during 2013. It also provided a sufficient duration for a three-year follow-up period. The final program participant dataset was left-merged with the driver and crash records in a deterministic matching procedure, using records from 2008 to 2018. Few crash records were matched to the 24/7 Sobriety Program participants. The count was determined insufficient and crash records were dropped from the analysis.

Administrative record information was recoded to identify critical events for the survival analysis. The first DUI-related citation date in the driver record, after a participant's 24/7 Sobriety Program entry date, was defined a DUI reoffense event of interest. The events were collected for each driver during in the three-year monitoring interval following the start date. Participants who had no subsequent DUI-related offense were censored at the three-year point. Participant study-design censoring occurs because the event dates are analyzed to a specific point in time, at which point some individuals have not experienced the event of interest. Here the point in time was three years after the program start date, with censoring in cases where no DUI-related citation event occurred.

[1] The 24/7 Sobriety Program was introduced in 2008 as a regional pilot program. It was implemented statewide by legislation in 2010. The last substantial change to program administrative rules occurred in August 2012.

Table 1
Study variable definitions and source.

Variable	Description	Source
Event of interest	The first subsequent DUI citation in driver record after entering into the 24/7 Sobriety Program event date.	Driver license
Gender	Driver gender at study entry (0 = Female, 1 = Male).	Driver license
Age group	Driver age group study entry (0 = 35 years and older, 1 = 18–34 years).	Driver license
Region	Driver license address location in state at entry (0 = East, 1 = West).	Driver license
Alcohol-related program history	Alcohol-related traffic conviction as reason for prior program participation, transformed to dichotomous event (0 = No, 1 = Yes).	24/7 Sobriety Program records
Drug-related program history	Drug-related conviction as reason for program prior participation, transformed to dichotomous event (0 = No, 1 = Yes).	24/7 Sobriety Program records
Non-alcohol/drug related program history	Other conviction (e.g. domestic violence) as reason for prior program participation, transformed to dichotomous event (0 = No, 1 = Yes).	24/7 Sobriety Program records
24/7 Sobriety Program history	Previous participation in the program prior to the enrollment as a second-time DUI offender, transformed to dichotomous event (1 = No, 0 = Yes).	Driver license

Participant demographic variables were constructed from driver license records. The driver address, at the time of study entry, was used to create east and west region driver cohorts. Annual statewide driver surveys have shown significant regional differences in drivers' attitudes and practices with regard to alcohol-impaired driving (Vachal et al., 2019). Similarly, gender and age groups were created for the participants based on risk reflected in the same statewide annual survey of drivers. Two age group cohorts (18–34 years and 35 years and older) were used to capture the role of experience and maturity in driving decisions. Younger adults have greater risk propensity in driving practices and perceptions. Gender was also included as males survey responses have had significantly higher-risk tendencies relative to the female cohort.

During the study, researchers decided to explore a potential relation of program history to survival among the second-time DUI program participant subpopulation (Table 1). Study records unveiled that a participant may have a start date in the 24/7 Sobriety Program prior to the legally mandated term for a second DUI arrest. In North Dakota, the 24/7 Sobriety Program is administered by the state driver licensing agency and managed by the state's judicial arm. Legislated DUI sanctions require second-time offenders to complete 360 days in the program. Program history was a binary variable defined as a driver entry prior to the start date of the second DUI offense. The entry was determined to be unique from the second-time offense based on a 30-day period between start dates, as recommended in discussions with program experts. The 30-day lapse allows time to complete state driver license and court in regard to initial event records.

Previous participation may have been court mandated with another conviction type or a voluntarily first-time DUI offender. The volunteer aspect was introduced within the past two years with a program administration change. The program history could not be studied in detail due to insufficient detail. It was, however, posited as a negative predictor in participant survival since it was viewed as a pre-study reoffense event.

The Institutional Review Board at North Dakota State University reviewed this study and determined it exempt. While the human subject research review process is continuous, it was determined that no informed consent and protection of subjects was required. The identity of the human subjects cannot readily be ascertained, directly or through identifiers, as recorded by the investigator.

3. Results

Among 1,287 study participants, 1,178 individuals have no subsequent DUI citations within three years of their start in the 24/7 Sobriety Program. The survival rate of second-time DUI offenders aged 18 or older and who started the program between 2013 and 2015 was 91.5%. The remaining 109 participants had at least one DUI-related offense within the three years after starting the program, so the reoffense rate was 8.5%.

3.1. Hazards function homogeneity

Gender, age, and region were considered in relation to survival outcomes with regard to independent variable assumptions. The hazard functions of each stratum were different. In order to check the PH assumption, Kaplan-Meier curves were specified in log(-log (survival functions)) versus log(survival time). The PH assumption is satisfied when the graph of the survival function versus survival time produce a graph with parallel curves. Observation showed the PH assumption was not valid since curves were intersecting for each group (Fig. 1). Therefore, a Stratified Cox regression model was used since it permits covariates with non-proportional hazard rates (Lee & Wang, 2003).

Survival probabilities were estimated based on the defined strata. The Stratified Cox regression model, however, possesses a limitation of automatically excluding the strata variable from the set of explanatory variables. Thus, no inferences can be made on the strata variable. To overcome this issue, both gender and region were considered as the strata variables since they violated PH assumption. On the contrary, age group was removed as a strata variable since it is time dependent. A new variable 'TD age' was created, which is the time dependent version of age variable generated by multiplying the age with log (survival function).

3.2. Survival findings

The results of Stratified Cox regression model are presented in the Table 1. The effect of age group and TD age on recidivism can be observed. A hazard ratio of 6.31 for the age group 35 years and older implies that the risk of experiencing the event of interest (i.e., recidivate) was 6.31 times higher in the age group 35 years and older as compared to the age group of 18–34 years. This finding was surprising since the younger population is usually identified as higher-risk.

Similarly, more accurate interpretation for age can be obtained from the TD age variable since it incorporates time dependency of age. The hazard ratio of 0.99 can be attributed to the higher probability of suffering the event of interest for one-year increase in TD age. A total of 109 observations were used in the model development. The model fit criterion of −2 Log Likelihood (−2 Log L) was used for the model selection. According to this criterion, the lower the statistical value, the better the model. The model results indicate that the model with covariates is better than the intercept only model.

The frequency of offender history types described earlier are presented in Table 2. Almost 91% of the offenders had no 24/7 Sobriety Program history, while 7% had alcohol-related program history. Initially, it was anticipated that the program history type may be influential in survival but counts were severely limited. Only 10 offenders had previous 24/7 Sobriety Program history, with one offender having a drug-related and other 24/7 Sobriety

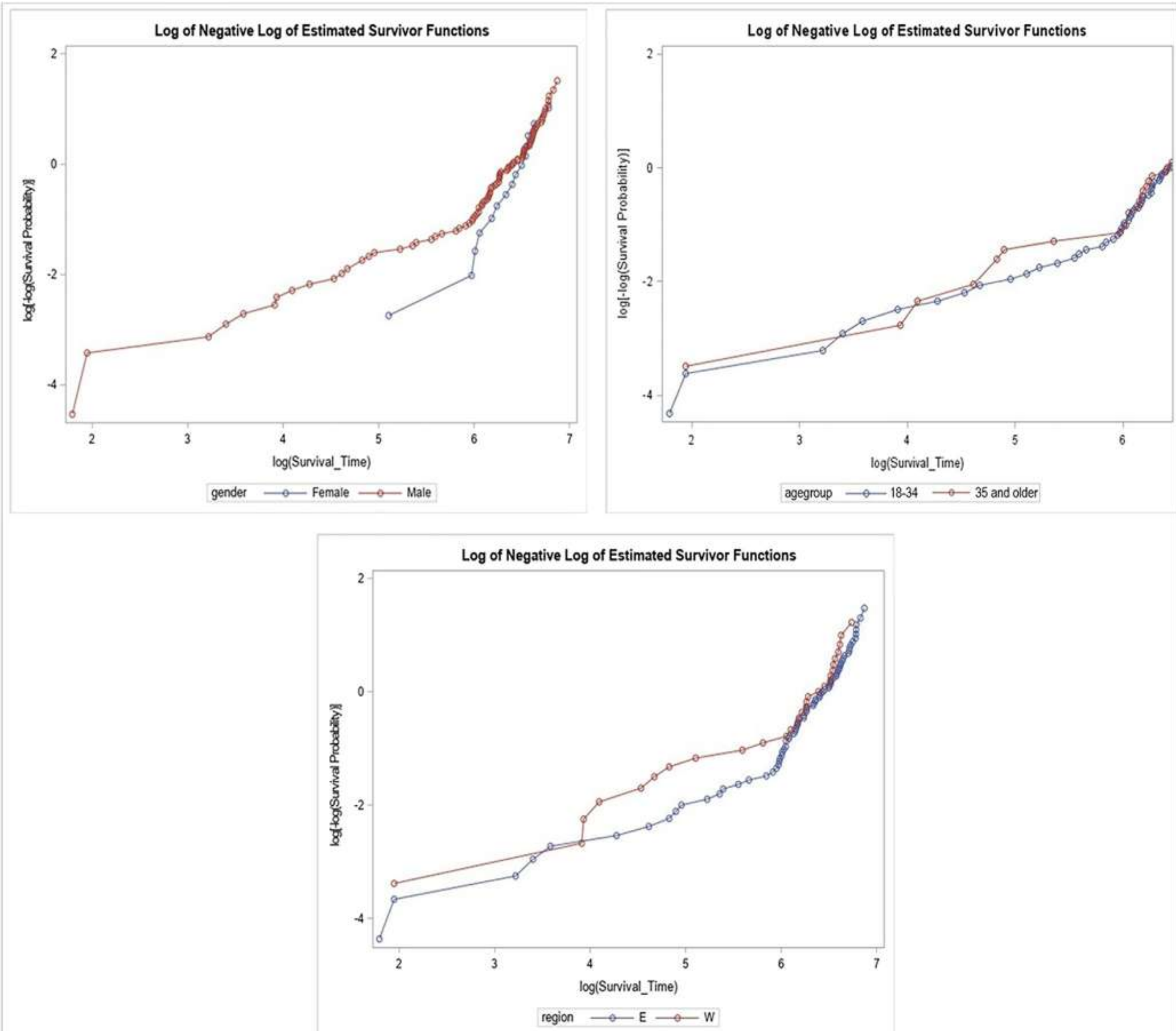

Fig. 1. Log of Negative Log of Estimated survival functions for gender, age group and region.

Table 2
Stratified Cox regression model results about DUI recidivism for the participants in all four program history types.

Parameter	Parameter estimate	Standard error	Hazard ratio
Age group (35 years and older)	1.84	0.35	6.31
TD age	−0.01	0.00	0.99
−2 Log L (without covariates)	681.718		
−2 Log L (with covariates)	654.149		
Number of observations	109		

Program history. The largest group, including 99 offenders, had no previous 24/7 Sobriety Program history. The other three offender history types are very rare, hence, may provide biased survival curves. Therefore, further analysis focuses on the participants who have no previous 24/7 Sobriety Program history before entering the program as a second-time offender.

Fig. 2 shows the survival rate change within the participants who have no previous 24/7 Sobriety Program history and have subsequent DUI citations. Fig. 2 shows three trends. In the first tendency, ranging approximately from 0 days to 360 days, the survival curve declines at a relatively slow rate. Then from approximately 360 days to 750 days, the slope of the survival curve declines at a substantially greater rate. Finally, from approximately 750 days onward, the survival curve decline slows.

Table 3 shows the number of offenders who have a subsequent DUI offense and no previous 24/7 Sobriety Program history. Table 3 shows trends similar to Fig. 2. The recidivism rate started with a slower increase from a period of 3 to 12 months, then significantly increased from month 12 to month 24, finally slowing between months 24 and 36. The interval 12 –24 months after enrolling in the program is a high-risk period for the participants without previous 24/7 program participation. These participants were more likely to have another DUI-related citation 12–24 months after program entry than in the first 12 months or 25–36 months after entry. The cumulative reoffense rate within three years was 8.18% for second-time DUI offenders who did not have previous 24/7 Sobri-

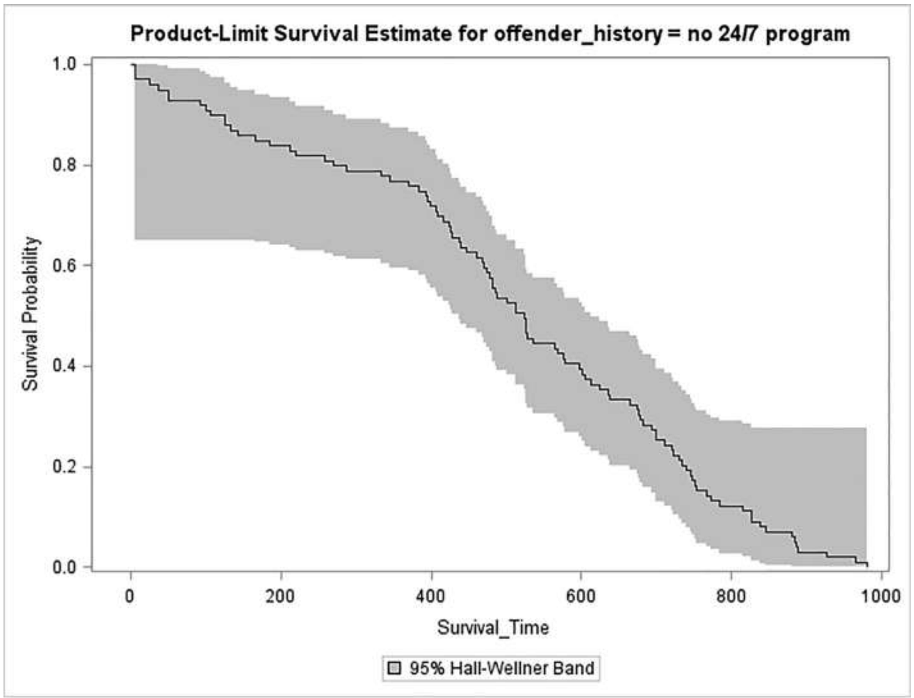

Fig. 2. Survival rate change among 24/7 Sobriety Program study participants with no program history, considering DUI reoffense event dates.

Table 3
The 24/7 Sobriety Program history type for the participants regarding DUI subsequent recidivism from 2013 to 2015.

Offender history	Number of participants without subsequent DUI	Number of participants with subsequent DUI	Total
Alcohol-related	24	8	32
Drug-related	11	1	12
Non-alcohol/drug related	32	1	33
No 24/7 Sobriety Program history	1111	99	1210
Total	1178	109	1287

Table 4
DUI Reoffense among 24/7 Sobriety Program with no program history, rates based on all participants in study cohort.

Time interval after program start date	Number of drivers with DUI reoffense	Reoffense rate (Percentage)
3 months	7	0.58
6 months	8	0.66
9 months	5	0.41
12 months	3	0.25
15 months	15	1.24
18 months	17	1.40
21 months	11	0.91
24 months	12	0.99
27 months	10	0.83
30 months	8	0.66
33 months	3	0.25
36 months	0	0
Total	99	8.18

ety Program history. Due to the limited number of observations with program history, comparisons were not drawn (Table 4).

4. Conclusion

Repeat DUI offenders remain a problem for many states in reducing alcohol-impaired driving incidents. While universal strategies such as higher legal drinking ages and lowering legal BAC thresholds have had positive outcomes in deterring alcohol-impaired driving, other programs are needed to address this higher-risk driver subpopulation where reoffense remains too common. One promising strategy is the 24/7 Sobriety Program. The aim is to elicit driver self-responsibility to prevent alcohol-impaired driving. Survival analysis was used to study participant success during and two years following their first mandated program term, following a second DUI arrest. Administrative records were linked and recoded to complete the study. Survival analysis, including Stratified Cox regression, were developed to investigate explanatory variable relations in DUI reoffense.

While survival rates were high, results may be useful in targeting strategies to older drivers and those with previous program experience. Results showed that older participants were more likely to have a DUI reoffense than the younger cohort. Reoffense was most likely during the 12 month-period after completing the mandated 360-day program term. Thus, monitoring DUI-related reoffense events that occur shortly after the program term ends may be especially beneficial in prevention. An unexpected finding was the inverse relation between age and longer-term survival. Investigating this along with potential addiction issues associated with alcohol use may be informative.

Although the study was limited to North Dakota drivers, findings increase knowledge for refining strategies designed to impact driver subpopulations at higher risk for reoffense. The administrative records available for this study contained limited participant detail. It may be possible to expand the independent variable array in states with stronger administrative record linkages or more extensive record systems. For instance, comprehensive criminal history may strengthen future evaluations of the 24/7 Sobriety Program, or similar driver-centric efforts, in terms of an individual's larger criminal behavior context. Considering external factors, such as alternative transport availability and enforcement activity levels, may further discern program influences in impaired driving prevention.

Acknowledgments

The authors offer appreciation to the North Dakota Department of Transportation and the North Dakota Secretary of State for supporting this research. Acknowledgement is also given to the Mountain-Plains Consortium which is sponsored by the U.S. Department of Transportation through its University Transportation Centers program.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Vachal, Khan, Zhou, Awasthi, Tchakounte-Wakem and Choi; data collection: Zhou and Awasthi; analysis and interpretation of results: Zhou, Khan, Awasthi, Vachal, Choi and Tchakounte-Wakem; draft manuscript preparation: Khan, Zhou, Awasthi, Vachal, and Tchakounte-Wakem. All authors reviewed the results and approved the final version of the manuscript.

References

Bagloee, S. A., & Asadi, M. (2016). Crash analysis at intersections in the CBD: A survival analysis model. *Transportation Research Part A: Policy and Practice, 94*, 558–572. https://doi.org/10.1016/J.TRA.2016.10.019.

Bjerre, B. (2005). Primary and secondary prevention of drink driving by the use of alcolock device and program: Swedish experiences. *Accident Analysis and Prevention, 37*(6), 1145–1152. https://doi.org/10.1016/j.aap.2005.06.020.

Carpenter, C. S., Kloska, D. D., O'Malley, P., & Johnston, L. (2007). Alcohol control policies and youth alcohol consumption: Evidence from 28 years of monitoring the future. *The B.E. Journal of Economic Analysis & Policy, 7*(1). https://doi.org/10.2202/1935-1682.1637.

Elder, R. W., Voas, R., Beirness, D., Shults, R. A., Sleet, D. A., Nichols, J. L., & Compton, R. (2011, March 1). Effectiveness of ignition interlocks for preventing alcohol-impaired driving and alcohol-related crashes: A community guide systematic review. American Journal of Preventive Medicine. Elsevier Inc. 10.1016/j.amepre.2010.11.012.

Federal Bureau of Investigation, Crime in the U.S. (https://ucr.fbi.gov/crime-in-the-u.s).

Fell, J. C., Tippetts, A. S., & Levy, M. (2008). Evaluation of seven publicized enforcement demonstration programs to reduce impaired driving: Georgia, Louisiana, Pennsylvania, Tennessee, Texas, Indiana, and Michigan. *Annals of Advances in Automotive Medicine/Annual Scientific Conference, 52*, 23.

Ferguson, S. A. (2012). Alcohol-Impaired Driving in the United States: Contributors to the Problem and Effective Countermeasures. *Traffic Injury Prevention, 13*(5), 427–441. https://doi.org/10.1080/15389588.2012.656858.

Kleiman, M.A.R., Hawken, A. (2008). Fixing the Parole System | Issues in Science and Technology. Issues in Science and Technology 24, No. 4. Retrieved from https://issues.org/kleiman/

Lee, E. T., & Wang, J. W. (2003). Statistical methods for survival data analysis.

Liu, X. (2012). *Survival analysis: Models and applications*. Wiley.

Loudenburg, R., Drube, G., & Leonardson, G. (2010). South Dakota 24/7 Sobriety Program Evaluation Findings Report. *Retrieved from* https://atg.sd.gov/docs/AnalysisSD24.pdf.

Machin, D., Cheung, Y. B., Parmar, M. K. B., & Parmar, M. K. B. (2006). *Survival analysis: A practical approach*. Wiley.

Marques, P. R., Voas, R. B., Roth, R., & Tippetts, A. S. (2010). Evaluation of the New Mexico Ignition Interlock Program. *Retrieved from*. www.ntis.gov.

McCartt, A. T., Leaf, W. A., Farmer, C. M., & Eichelberger, A. H. (2013). Washington state's alcohol ignition interlock law: Effects on recidivism among first-time DUI offenders. *Traffic Injury Prevention, 14*(3), 215–229. https://doi.org/10.1080/15389588.2012.708885.

McGinty, E. E., Tung, G., Shulman-Laniel, J., Hardy, R., Rutkow, L., Frattaroli, S., & Vernick, J. S. (2017). Ignition interlock laws: Effects on fatal motor vehicle crashes, 1982–2013. *American Journal of Preventive Medicine, 52*(4), 417–423. https://doi.org/10.1016/j.amepre.2016.10.043.

Midgette, G., & Kilmer, B. (2015). The effect of Montana's 24/7 Sobriety Program on DUI re-arrest: Insights from a natural experiment with limited administrative data. *RAND Corporation*. https://doi.org/10.7249/WR1083.

Miron, J., & Tetelbaum, E. (2007). Does the minimum legal drinking age save lives? Cambridge, MA. 10.3386/w13257.

National Conference of State Legislatures, State Ignition Interlock Laws (https://www.ncsl.org/research/transportation/state-ignition-interlock-laws.aspx).

NHTSA. (2005). Calculating Lives Saved Due to Minimum Drinking Age Laws. NHTSA's National Center for Statistics and Analysis (Vol. 809). Washington, DC, U.S., Department of Transportation.

Rauch, W. J., Zador, P. L., Ahlin, E. M., Baum, H. M., Duncan, D., Beck, K. H., ... Gretsinger, N. (n.d.). A longitudinal survival analysis of drivers with multiple alcohol-related traffic offenses: Fifth year follow-up of a randomized ignition interlock license restriction trial in Maryland.

Tippetts, A. S., Voas, R. B., Fell, J. C., & Nichols, J. L. (2005). A meta-analysis of.08 BAC laws in 19 jurisdictions in the United States. *Accident Analysis and Prevention, 37*(1), 149–161. https://doi.org/10.1016/j.aap.2004.02.006.

Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). Survival analysis: Pedestrian risk exposure at signalized intersections. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(2), 77–89. https://doi.org/10.1016/J.TRF.2006.06.002.

Vachal, K., Benson, L., & Kubas, A. (2019). North Dakota Statewide Traffic Safety Survey, 2019: Traffic safety performance measures for state and federal agencies. North Dakota State University, Fargo: Upper Great Plains Transportation Institute. Retrieved from https://www.ugpti.org/resources/reports/details.php?id=974.

Wagenaar, A. C., Maldonado-Molina, M. M., Erickson, D. J., Ma, L., Tobler, A. L., & Komro, K. A. (2007). General deterrence effects of U.S. statutory DUI fine and jail penalties: Long-term follow-up in 32 states. *Accident Analysis and Prevention, 39*(5), 982–994. https://doi.org/10.1016/j.aap.2007.01.003.

Wagenaar, A. C., Maldonado-Molina, M. M., Ma, L., Tobler, A. L., & Komro, K. A. (2007). Effects of Legal BAC Limits on Fatal Crash Involvement: Analyses of 28 States from 1976 through 2002. *Journal of Safety Research, 38*(5), 493–499. https://doi.org/10.1016/j.jsr.2007.06.001.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). Statistical and econometric methods for transportation data analysis, Second edition. Statistical and Econometric Methods for Transportation Data Analysis, Second Edition. CRC Press. 10.1198/tech.2004.s238.

Zhao, F., Guo, H., Wang, W., & Jiang, X. (2015). Modeling lateral interferences between motor vehicles and non-motor vehicles: A survival analysis based approach. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/J.TRF.2015.07.005.

Zaloshnja, E., Miller, T., & Blincoe, L. (2013). Costs of alcohol-involved crashes, United States, 2010. *Annals of Advances in Automotive Medicine, 57*, 3–12.

Dr. Kimberly Vachal is the lead investigator for the Rural Transportation Safety and Security Center at North Dakota State University's Upper Great Plains Transportation Institute. She conducts research and outreach to promote a safe and secure transportation, with focus on rural regions. She has over 25 years of experience in transportation research focused on public policy and program assessment. She has written over 50 research papers and authored several journal articles. She earned her PhD in Public Policy at George Mason University.

Shantanu Awasthi is a PhD Candidate in the Mathematics Department at the North Dakota State University with a teaching assistant appointment. He has been active as a graduate research assistant with the Upper Great Plains Transportation Institute at NDSU since 2019. He received his Department Research Award for Graduate Students in 2021.

Ihsan Ullah Khan is a PhD Candidate in the Department of Transportation, Logistics and Finance at the North Dakota State University. He has over four years of experience in statistical data analysis. His research interests include traffic safety, statistical modelling, and transportation planning.

Yun Zhou is a PhD Candidate in the Department of Transportation, Logistics and Finance at North Dakota State University (NDSU). She has been a graduate research assistant with the Upper Great Plains Transportation Institute at NDSU since 2018. Her current research is about reducing drunk driving and improving safety of teen drivers. She has experience in statistical data analysis and machine learning. Her research interests include traffic safety, transportation data science and supply chain modeling.

Dr. Bong-Jin Choi is an Assistant Professor in the Departments of Statistics and Public Health at North Dakota State University. He has authored numerous publications in related to public health and statistical methods. Dr. Choi was awarded his PhD in Statistics from the University of South Florida.

Seguy Tchakounte-Wakem is a Research Specialist at North Dakota State University's Upper Great Plains Transportation Institute. He has experience in statistical data analysis. His research interests include traffic safety, transportation data science and supply chain logistics.

Journal of Safety Research 78 (2021) 36–46

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Evaluating sleep deprivation and time-of-day influences on crash avoidance maneuvers of young motorcyclists using a dynamic simulator

Clément Bougard [a,b,*], Damien Davenne [b], Sébastien Moussay [b], Stéphane Espié [c]

[a] Groupe PSA, Centre technique de Vélizy, Vélizy-Villacoublay, Cedex, France
[b] Normandie University, Unicaen, INSERM, COMETE, CHU de Caen, Cyceron, Caen, France
[c] Université Gustave Eiffel/TS2/SATIE/MOSS, Orsay, France

ARTICLE INFO

Article history:
Received 15 May 2020
Received in revised form 15 February 2021
Accepted 19 May 2021
Available online 3 June 2021

Keywords:
Time-of-day
Sleep deprivation
Motorcycle dynamic simulator
Emergency braking
Swerving

ABSTRACT

Introduction: Motorcyclists are particularly at risk of being injured when involved in a road traffic accident. To avoid such crashes, emergency braking and/or swerving maneuvers are frequently performed. The recent development of dynamic motorcycle simulators may allow to study the influences of various disturbance factors such as sleep deprivation (SD) and time-of-day (TOD) in safe conditions. *Methods:* Twelve young healthy males took part in 8 tests sessions at 06:00 h, 10:00 h, 14:00 h, 18:00 h after a night with or without sleep, in a random order. Participants had to perform an emergency braking and a swerving maneuver, both realized at 20 and 40 kph on a motorcycle dynamic simulator. For each task, the total distance/time necessary to perform the maneuver was recorded. Additional analysis was conducted on reaction and execution distance/time (considered as explanatory variables). *Results:* Both crash avoidance maneuvers (emergency braking and swerving) were affected by increased speed, resulting in longer time and distance at 40 kph than at 20 kph. Emergency braking was mainly influenced by sleep deprivation, which significantly increased the total distance necessary to stop at 40 kph (+1.57 m; + 20%; $p < 0.01$). These impaired performances can be linked to an increase in reaction time (+21%; $p < 0.01$). Considering the swerving maneuver, TOD and SD influences remained limited. TOD only influenced the reaction time/distance measured at 40 kph with poorer performance in the early morning (+30% at 06:00 h vs 18:00 h; $p < 0.05$). *Discussion:* Our results confirm that crash avoidance capabilities of young motorcyclists were influenced by the lack of sleep, mainly because of increased reaction times. More complex tasks (swerving maneuver) remained mostly unchanged in this paradigm. *Practical Applications:* Prevention campaigns should focus on the dangers of motorcycling while sleepy. Motorcycling simulators can be used to sensitize safely with sleep deprivation and time-of-day influences.

1. Introduction

Powered Two-Wheeler (PTW) riders are particularly at risk of being seriously injured or killed when involved in a road traffic accident. Although PTWs make up a small proportion of circulating vehicles in most countries (3–15%), PTW riders contribute to nearly 15–20% of all road deaths worldwide (European Commission, 2018; ONISR, 2019; NHTSA, 2019). The fatality rate (number of fatal accidents when accounting for per vehicle mile traveled) is significantly higher for PTW riders than for car drivers. This related increased risk is 27-fold in the United States (National Highway Traffic Safety Administration, 2019), 20-fold in Europe (ONISR, 2019), and 29-fold in Australia (Haworth & Mulvihill, 2005). Even if PTW riding was considered a recreational activity in the last decades, Day et al. (2013) claimed that PTWs would be an integral part of the global transport future. The current trends confirm that PTWs are increasingly used for commuting (in particular to avoid traffic congestion in dense urban areas) and even for work purposes (e.g., taxi, delivery; de Rome, Brown, Baldock, & Fitzharris, 2018; Wu, Hours, & Martin, 2018; Möller et al., 2020). This may raise attention to their crash contributing factors.

Various studies based on epidemiological data have been conducted to identify associated risk factors in PTW safety. Regarding behavioral aspects, age (21–30 years), gender (male), experience (poor), absence of helmet and/or protective clothes, speeding, and blood alcohol concentration are highly related to crash occurrence and severity (de Rome et al., 2018; Vlahogianni, Yannis, &

* Corresponding author at: Groupe PSA, Centre technique de Velizy, Route de Gisy, 78140 Vélizy-Villacoublay, France.

E-mail address: clement.bougard@live.fr (C. Bougard).

https://doi.org/10.1016/j.jsr.2021.05.006
0022-4375/

Golias, 2012; Wali et al., 2018, 2019; Wu et al., 2018). However, the results of these retrospective studies remained limited to the a posteriori available data. Other potential risk factors (such as fatigue and sleepiness) may thus be under-estimated since they are rarely taken into account in road accident reports, in particular regarding PTW crashes (Wu et al., 2018). For example, despite a lower traffic density at night, PTW riders are more at risk of being seriously injured or killed while riding in the dark (de Lapparent, 2006; Möller et al., 2020; Quddus, Noland, & Chin, 2002). The severity of the accidents (number of fatalities/100 victims) is 1.5-fold higher at night than during the day (ONISR, 2019). Therefore, in France, 2005 road safety statistics reported that 35% of fatal accidents involving PTW riders and 26% of those with a serious injury happened at night (ONISR, 2007 Similarly, despite the fact that the riding experience itself is often considered as energizing, which would help offset feelings of fatigue, recent findings indicate the opposite. As previously observed for car drivers, international observations reported that fatigue and sleepiness are also associated with a higher risk for PTW riders of getting involved into a crash (Chen & Chen, 2016; Wali, Khattak, & Khattak, 2018; Lam et al., 2019; Santos, Cêlho, Santos, & Ceballos, 2019; Zheng, Ma, Guo, Cheng, & Zhang, 2019). More precisely, Chen and Chen (2016) showed that motorcyclists who had less than 6 hours sleep were more likely to reveal more risky riding behaviors, especially for speed. Wali et al. (2018) also indicated that the risk of getting involved in a crash increased by approximately 200% if the rider had 5 h or less sleep (surrogate of drowsy-riding and/or fatigue).

Precrash maneuvers are associated with injury severity sustained by PTW riders involved in single-vehicle crashes (Wang et al., 2016). When a precrash maneuver is performed, the two most common maneuvers observed are braking and/or swerving (ACEM, 2009; Wang et al., 2016). Furthermore, emergency braking and swerving maneuvers performed on the left or on the right are also learned in PTW training programs and motorcycle license practice tests (Haworth et al., 1997), contributing to alleviate the crash risks and improve PTW safety (Wali et al., 2018). Emergency braking and/or swerving on a PTW require a high-level of perceptive and manipulative capabilities (vanElslande, 2012). However, vigilance impairment, which can be induced by sleep deprivation (SD) and/or time-of-day (TOD), is recognized to negatively affect the different components of executive functioning implied in motorcycling (Bougard et al., 2016a, 2016b). Previous works from our laboratory have shown that motorcycling performance evaluated through a broad set of basic riding skills on a test track is sensitive to the effects of TOD, with an improvement of riding performance throughout the day following a normal night of sleep (Bougard et al., 2008, 2012). These riding performances evolved concomitantly to the rhythm of body temperature, considered as a good marker of the biological clock (Colquhoun, 1971). Considering SD effects, performance impairment depends on the ability under consideration (reaction time, motor coordination, balance, flexibility). More precisely regarding precrash avoidance maneuvers, Bougard et al. (2012) have shown that emergency braking was impaired in real motorcycling conditions by both disturbance factors, TOD and SD, mainly due to an impairment in the reaction time phase. Swerving capabilities were far less impacted. Nonetheless, for safety reasons due to the real motorcycling conditions of our experiment, the duration of the tests was limited (15 min) and the speed was reduced (20–40 kph), which may have limited the impacts of disturbance factors such as sleepiness. Nonetheless, to improve PTW safety, crash avoidance maneuvers deserve more research attention (Kuang et al., 2015) and evaluating the influence of fatigue and sleepiness on their execution is of particular interest.

Recently, the development of PTW simulators allowed for the analysis of PTW rider behaviors in various situations. The advantage of such material is to propose different critical riding situations without inducing any risk of injuries (de Winter, Van Leeuwen, & Happee, 2012). Some studies used a fixed platform in a static vertical position in order to evaluate the influence of the level of expertise on hazard detection (Cheng, Ng, & Lee, 2011; Hosking, Liu, & Bayly, 2010), or the effects of alcohol consumption on riding performance (Centola et al., 2020; Filtness, Rudin-Brown, Mulvihill, & Lenné, 2013). Others used a dynamic configuration that offers greater validity for PTW riding tests (Benedetto et al., 2014). In particular, this allows evaluating PTW rider behaviors in more realistic situations, such as lane position while negotiating a curve (Crundall, van Loon, Stedmon, & Crundall, 2012), or gaze locations in bends (Lobjois & Mars, 2020). Lastly, Kovácsová et al. (2020) studied emergency braking at intersections using a motion-base simulator. They reported that the more dangerous the situation, the more likely riders were to initiate braking. Moreover, even if riders braked in an impending situation, they were often unsuccessful in avoiding a collision. For swerving maneuvers, although motion cueing algorithms are increasingly efficient regarding PTW dynamics, reproducing these maneuvers on a dynamic simulator still remains challenging. Nonetheless, and despite such sophisticated tools presenting numerous advantages, none has yet evaluated the influences of TOD and SD on PTW riding behavior during precrash maneuvers using a moving platform.

The aim of this study was to evaluate the influences of time-of-day and sleep deprivation on two different crash avoidance maneuvers, namely emergency braking and swerving, performed on a dynamic motorcycle simulator. To do so, and to allow for possible comparisons, we choose to reproduce the experimental procedure used by Bougard et al. (2012), who were, to the best of our knowledge, the only ones to evaluate these effects in real motorcycling conditions. We hypothesized that similarly to real conditions, emergency braking performance on the simulator, which is a quite simple task, would improve during the day after a normal night of sleep (according to the diurnal fluctuations of body temperature) and that this improvement would be affected by SD. We also hypothesized that modifications of braking performance would essentially be paralleled by modifications in reaction time. In contrast, swerving (being a more complex maneuver) should remain constant throughout the day, regardless of the sleep condition.

2. Materials and methods

The study protocol complied with the tenets of the Declaration of Helsinki and was approved by the local ethics committee (Comité de Protection des Personnes Nord-Ouest III, France, n° 2007-A00581-52).

2.1. Participants

Twelve healthy young males (age: 22.9 ± 1.9 years old; height: 177.6 ± 7.5 cm; weight: 79.2 ± 13.2 kg) voluntarily signed an informed consent before being included in this study. To guarantee the homogeneity of the sample, particular attention was paid to motorcycling experience (participants held a motorcycling license for 3.8 ± 2.6 years) and to the participant's chronotype (person's circadian typology, reflecting morning and evening preferences) according to their answers to the Horne and Ostberg (1976) questionnaire. All the participants were 'intermediate type' (score from 42 to 58), reflecting a mean bedtime between 22:30 h and 00:00 h; and a wake-up time between 07:00 h and 09:00 h. Moreover, all the participants had a score <11 on the Epworth sleepiness scale (Johns, 1991) in order to avoid excessive diurnal sleepiness.

2.2. Study design

As illustrated in Fig. 1, each participant was evaluated during four test sessions, set up at 06:00 h, 10:00 h, 14:00 h, 18:00 h, following a night with or without sleep (corresponding to 24, 28, 32 and 36 waking hours in the SD condition) in a random order. The two days of testing were separated by a period of 1 week to allow for recovery from the night of sleep deprivation. For each of the two sleep conditions, the tests described in the following section were carried out once by each of the 12 participants at each time-of-day.

The day before the tests, participants were asked to wake up at 07:00 h (controlled by actimetry) before coming to the laboratory (which had room temperature of 21.6 ± 0.8 °C) at 19:00 h, and ate a standardized meal at 20:00 h. When the participants were evaluated following a night of normal sleep, they were asked to go to bed at 22:30 h in order to guarantee a minimum of 6 h in bed. An experimenter woke them up at 05:00 h. Under these conditions, sleep duration conformed to the participants' usual sleeping habits, since participants reported during the inclusion visit that their mean sleep duration was 7 h. For the night of SD, the participants remained in the presence of an experimenter and were not allowed to lie down. During this night of SD, the participants were only allowed to take part in activities not involving any physical load or excitement (such as reading, watching a movie, and playing cards), and energizing and stimulant drinks (coffee, tea, etc.) were not allowed (Reilly & Bambaeichi, 2003). All these precautions were also applied during the day, in between the test sessions.

A standardized breakfast was provided at 08:30 h, after the experimental session at 06:00 h (Baxter & Reilly, 1983), in order to limit inter-individual variability of the results (Bougard, Moussay, Gauthier, Espié, & Davenne, 2009). A standardized meal was also provided for lunch at 12:30 h, and a light snack at 16:30 h. To avoid any performance improvement during the experiment due to a practice effect (Millar, 1992), all the participants were trained during a pre-experimental session set up a week before at 13:00 h, to obtain stabilization of their performances. More precisely, they simulator sickness symptoms. After a 30 min break, they had to complete the same riding tests as for the experimental test sessions during two training blocks of 15 min each, separated by 30 min.

2.3. Measurements

2.3.1. Oral temperature

Before each riding session on the PTW dynamic simulator, the participants were asked to lie down on a bed for 15 min (Souissi et al., 2007). Then, oral temperature was measured by an experimenter using a digital clinical thermometer (Omron, accuracy: 0.05 °C) inserted sublingually for at least 3 min.

2.3.2. Simulated riding tests

As proposed in Bougard et al. (2012), who studied the influences of time-of-day and sleep deprivation on real motorcycling capabilities, participants had to perform an emergency braking and a swerving maneuver, both realized at 20 and 40 kph in a random order, on a motorcycle dynamic simulator (Fig. 2A). The same environment, an underground private parking area (Fig. 2B), was reproduced on the simulator (Fig. 2C). This simulator is based on the framework of a real motorcycle (125 YBR, Yamaha®), fixed on a mobile platform including 5DDL (roll, yaw, pitch, handlebar rotation, force feedback on the handlebar). It allows the reproduction of intensive braking, acceleration, and taking curves by tilting the PTW at ±10°. All the commands from a real PTW were conserved and encoded via various sensors, allowing to ride in the virtual environment. A speedometer with LEDs on each side was reproduced on the lower screen so that the rider did not need to switch his gaze and attention between the real and simulated environment. In order to cover a wide field of view (60°) a video projection screen (height: 128 cm; width: 177 cm) was placed at 120 cm from the simulator with the horizon line at the participant's eye level.

The test course comprised two different zones: a preparation zone (40 m) in which PTW riders had to achieve the required speed; an exercise zone (40 m) in which riders had to perform the required maneuver. At each extremity, an empty zone of 10 m was respected for visual comfort and fidelity with the real underground parking area.

The emergency braking test was carried out at either 20 kph (12.43 mph) or 40 kph (24.86 mph) with the brake lever and pedal in a "ready-to-brake" position (foot just above the rear wheel brake pedal and fingers beyond the front wheel brake lever). Participants had to achieve the required speed before entering the exercise zone. From the entrance of the exercise zone, one LED was switched on in a random way. At this signal, the participant had to brake in order to stop the PTW as soon as possible (Fig. 3). To better analyze the effects of TOD and SD on emergency braking, the combination of the different encoders allowed for the calculation of the following measures, retained as performance indices:

- Stopping time/distance: these measures describe the time/distance necessary for the rider to stop the PTW, once the LED was turned on.
- Reaction time/distance: these measures describe the time/distance between the LED turn-on and the rider pressed the brakes.
- Braking time/distance: these measures represent the time/distance during which the brakes were activated.

As for the swerving maneuver, the expected speed also had to be attained before entering the exercise zone. At the beginning of that zone, participants had to adjust the path of the PTW to the median line. Then, one of the two LEDs was switched on and the participant had to adjust the path of the PTW to the line on the same side as the LED, as soon as possible. The order in which participants had to perform the maneuver on the left or the right were randomly chosen. The following measures were retained as performance indices:

- Total time/distance for the avoidance: these measures correspond to the time/distance separating the LED turn-on and the stabilization of the PTW path on the designated line accord-

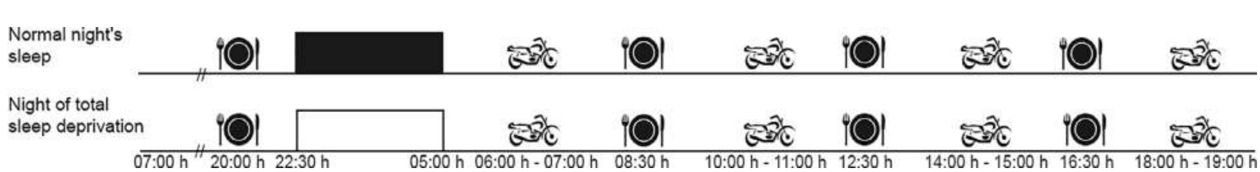

Fig. 1. Experimental protocol. Each participant participated in two days of test sessions, organized after either a normal night's sleep or a night of SD. During these days, standardized meals were provided in between test sessions.

Fig. 2. A. Participant on the dynamic simulator during the riding test; B. Real underground private parking area (Bougard et al., 2012); C. Simulated underground private parking area.

ing to the LED side (left/right) (Fig. 4). To determine when the path was stabilized on the line, the recorded data were analyzed backward from the end of the trial. Once the handlebar position exceeded three standard deviations from the mean value calculated in the preparation zone, the path of the PTW was considered as stabilized.

- Reaction time/distance: these measures represent the time/distance separating the LED turn-on from the initiation of the turn. To determine the initiation of the turn, the mean position of the handlebar in the preparation zone was calculated. Then, once in the exercise zone, the first time the handlebar position exceeded three standard deviations from the mean value previously calculated was considered as the initiation of the turn.
- Swerving time/distance: these measures correspond to the time/distance necessary for the rider to stabilize the path of the PTW on the appropriated line once he had initiated the turn.

Emergency braking and swerving are difficult and dangerous maneuvers on a motorcycle. In real-life conditions, riders can easily drop the bike or lose control. To improve realism and better capture the real-life experience of riding a motorcycle, a 'virtual' fall occurred (reproduced by the simulation model in the visual scene and also by little vibrations in the motorcycle seat) when a rider failed to perform safely (braked too hard or overturned the bike).

We analyzed first the effects of TOD and SD on the time/distance necessary to perform the whole maneuver. Then, we further analyzed the reaction and execution part (braking/swerving) of the maneuver, considered as explanatory variables.

2.4. Statistical analysis

Since, in this experiment, participants took part in several test sessions throughout the day following both a normal night's sleep and a night of sleep deprivation, they were considered as their own control. The repeated-measures ANOVA allow for comparing the effects of each disturbance factor separately (sleep condition or time-of-day) and also their possible interactions. In addition, with the motorcycling tests being performed at different speeds (20 kph and 40 kph) and according to instructions regarding the side of the maneuver (swerving maneuver), these factors were considered as categorical factors to identify their respective influence on participants performance.

The data recorded in temperature measurements during the eight test sessions were analyzed by a 2 (sleep condition: normal night; sleep deprivation) × 4 (time-of-day: 06:00 h, 10:00 h, 14:00 h and 18:00 h) repeated-measures analysis of variance (ANOVA). For the motorcycling tests, both maneuvers were evaluated at two different speeds. Consequently, emergency braking data were analyzed by a 2 (sleep condition: normal night; sleep deprivation) × 4 (time-of-day: 06:00 h, 10:00 h, 14:00 h, 18:00 h) repeated-measures analysis of variance (ANOVA) with a categorical factor (speed: 20 and 40 kph). For the swerving test, riders also had to perform the maneuver on both sides. The data was thus analyzed by a 2 (sleep condition: normal night; sleep deprivation) × 4 (time-of-day: 06:00 h, 10:00 h, 14:00 h, 18:00 h) repeated-measures analysis of variance (ANOVA) with two categorical factors (speed: 20 and 40 kph; side: left and right). To maintain clarity and a full understanding of the results, only the main effects and those of interaction limited to 2-levels are presented further.

For all the collected data, the condition of sphericity was tested (Mauchly's test). The *p*-value levels were corrected for possible deviations from sphericity by means of the Huynh–Feldt epsilon (ε). We report the uncorrected degrees of freedom, the ε value, and the *p*-value according to the corrected degrees of freedom. When significant differences were observed, a post hoc analysis was then performed with a Bonferroni test.

All differences were considered as significant for p-value <0.05. For each significant effect, we estimated the size effect using the

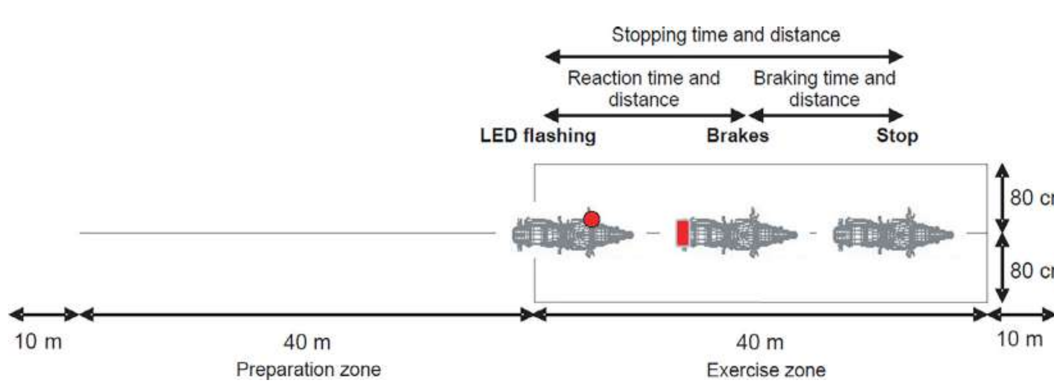

Fig. 3. Emergency braking test and recorded measurements.

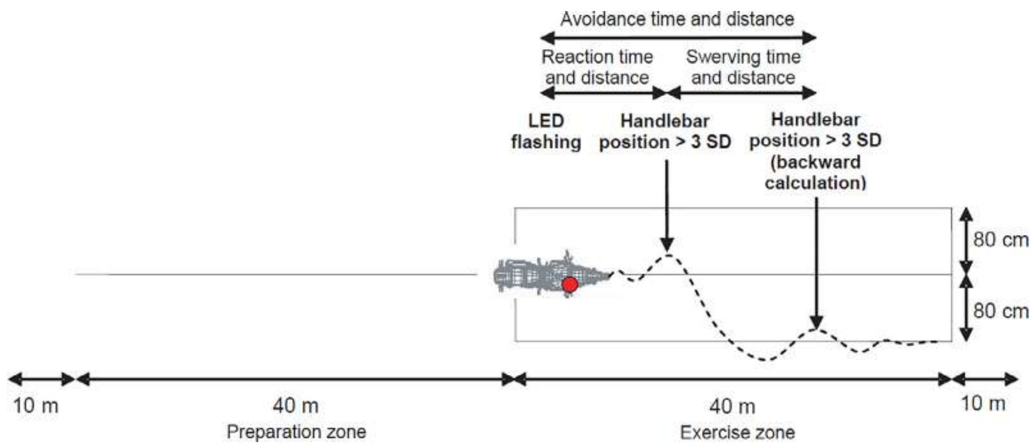

Fig. 4. Swerving maneuver and recorded measurements.

partial eta squared (partial η^2). All statistical analysis results are presented in detail in the Appendices (Table1 for Emergency braking and Table 2 for Swerving).

3. Results

3.1. Oral temperature

A significant interaction effect between 'sleep condition' and 'time-of-day' ($F_{(3,33)} = 5.72$; $\varepsilon = 1.00$; $p < 0.01$; partial $\eta^2 = 0.34$) was observed on oral temperature measurements. The post-hoc analysis indicated lower levels of temperature at 06:00 h after the normal night's sleep than after SD (35.81 ± 0.28 °C *vs* 35.99 ± 0.24 °C). In contrast, measurements realized at 10:00 h, 14:00 h and at 18:00 h did not depend on the previous sleep condition [(36.16 ± 0.31 °C *vs* 36.23 ± 0.12 °C), (36.44 ± 0.25 °C *vs* 36.36 ± 0.17 °C), (36.55 ± 0.24 °C *vs* 36.48 ± 0.20 °C), respectively].

3.2. Emergency braking

Before any analysis of emergency braking performance, it was necessary to check whether the participants rode at the correct speed. They respected speed instructions while riding at 20.86 ± 1.11 kph and 39.85 ± 3.81 kph at the beginning of the exercise zone ($F_{(1,22)} = 1582.38$; $p < 0.001$; partial $\eta^2 = 0.99$). Regarding the whole maneuver, a significant interaction effect between 'speed' and 'sleep condition' was observed on both the stopping distance ($F_{(1,22)} = 8.54$; $\varepsilon = 1.00$; $p < 0.01$; partial $\eta^2 = 0.28$) and time ($F_{(1,22)} = 7.76$; $\varepsilon = 1.00$; $p < 0.05$; partial $\eta^2 = 0.26$). At 20 kph, the participants stopped in a shorter distance (3.28 ± 0.26 m) and faster (0.63 ± 0.03) than at 40 kph (Fig. 5A). At 40 kph, the distance necessary to stop the PTW (normal night: 7.58 ± 0.20 m *vs* SD: 9.16 ± 0.39 m; +21%), and also the stopping time significantly increased with SD (normal night: 0.85 ± 0.02 s *vs* 1.01 ± 0.22 s after SD; +19%).

To have further information, it would be of interest to determine whether the effects observed on the stopping distance/time were induced by changes in the 'reaction' and/or 'braking' part of the maneuver.

As for the reaction part of the maneuver, a significant interaction effect between 'speed' and 'sleep condition' was observed on both the distance ($F_{(1,22)} = 9.96$; $\varepsilon = 1.00$; $p < 0.01$; partial $\eta^2 = 0.31$) and time ($F_{(1,22)} = 8.08$; $\varepsilon = 1.00$; $p < 0.01$; partial $\eta^2 = 0.27$). At 20 kph, participants reacted in a shorter distance (2.30 ± 0.25 m) and faster (0.39 ± 0.02 s) than at 40 kph. At 40 kph, the reaction distance (normal night: 6.53 ± 0.20 m *vs* SD: 8.05 ± 0.37 m; +23%) and time significantly increased with SD (normal night: 0.62 ± 0.02 s *vs* SD: 0.75 ± 0.03 s; +21%) (Fig. 5B).

Regarding the braking part of the maneuver, a significant effect of 'speed' was observed on the distance during which the brakes were activated ($F_{(1,22)} = 13.45$; $\varepsilon = 1.00$; $p < 0.001$; partial $\eta^2 = 0.38$). The participants pressed the brakes for shorter distances at 20 kph than at 40 kph (0.91 ± 0.20 m *vs* 1.08 ± 0.22 m; −16%). Surprisingly, the effect of 'speed' was not observed on braking times. In addition, an interaction effect between 'sleep condition' and 'time-of-day' was observed on braking distance ($F_{(3,66)} = 3.16$; $\varepsilon = 0.89$; $p < 0.05$; partial $\eta^2 = 0.13$) and time ($F_{(3,66)} = 3.21$; $\varepsilon = 0.83$; $p < 0.05$; partial $\eta^2 = 0.13$). The Bonferroni post-hoc analysis reported no significant differences regarding braking distances; while braking times measured after the normal night's sleep were significantly shorter at 06:00 h than at 18:00 h (0.19 ± 0.01 s *vs* 0.22 ± 0.01 s; −16%).

3.3. Swerving

The target speed recorded at the beginning of the exercise zone conformed to the instructions since participants rode at 21.32 ± 0.68 kph and 40.31 ± 1.24 kph ($F_{(1,45)} = 1796.40$; $p < 0.001$; partial $\eta^2 = 0.98$). 'Speed' significantly influenced the total distance ($F_{(1,45)} = 15.12$; $p < 0.001$; partial $\eta^2 = 0.25$) and time ($F_{(1,11)} = 64.57$; $p < 0.001$; partial $\eta^2 = 0.59$) necessary to perform the avoidance maneuver while swerving. The participants needed shorter distance (27.95 ± 0.86 m *vs* 32.67 ± 0.54 m; −14%) but took longer (4.88 ± 0.13 s *vs* 3.52 ± 0.12 s; + 38%) to change their path at 20 kph than at 40 kph.

To have further information, it would be of interest to determine whether the effects observed on the avoidance distance and time were induced by changes in reactions and/or changes in the movement part of the swerving maneuver.

Regarding the reaction part of the maneuver, a significant interaction effect between 'speed' and 'time-of-day' was observed on both the reaction distance ($F_{(3,135)} = 3.16$; $\varepsilon = 1.00$; $p < 0.05$; partial $\eta^2 = 0.07$) and time ($F_{(3,135)} = 2.79$; $\varepsilon = 1.00$; $p < 0.05$; partial $\eta^2 = 0.06$). At 20 kph, the participants travelled less distance before initiating their maneuver (1.87 ± 1.11 m) and reacted faster (0.32 ± 0.01 s) than at 40 kph. At 40 kph, reaction distance and time were increased at 06:00 h (4.44 ± 0.22 m; 0.42 ± 0.02 s, respectively) in comparison with measurements realized at 10:00 h (3.50 ± 0.20 m; 0.32 ± 0.02 s, respectively), 14:00 h (3.26 ± 0.20 m; 0.30 ± 0.02 s, respectively), and 18:00 h (3.40 ± 0.21 m; 0.30 ± 0.02 s, respectively).

Table 1
Repeated-measures analysis of variance (ANOVA) of Emergency braking. *Italic*: categorical factors. **Bold**: significant effect ($p < 0.05$). Speed (20; 40 kph), Sleep: sleep condition (normal night; sleep deprivation), TOD: time-of-day (06:00 h; 10:00 h; 14:00 h; 18:00 h).

	ddl	F	p-value	partial η^2		ddl	F	p-value	partial η^2		ddl	F	p-value	partial η^2
Speed instruction					Reaction distance					Braking distance				
Speed	(1,22)	1582.38	**0.000**	0.99	*Speed*	(1,22)	190.42	**0.000**	0.90	*Speed*	(1,22)	13.45	**0.001**	0.38
Sleep	(1,22)	0.00	0.951	0.00	Sleep	(1,22)	13.73	**0.001**	0.38	Sleep	(1,22)	0.25	0.620	0.01
Sleep × *Speed*	(1,22)	0.13	0.721	0.01	Sleep × *Speed*	(1,22)	9.96	**0.005**	0.31	Sleep × *Speed*	(1,22)	0.81	0.379	0.04
TOD	(3,66)	0.51	0.649	0.02	TOD	(3,66)	0.40	0.754	0.02	TOD	(3,66)	1.20	0.318	0.05
TOD × *Speed*	(3,66)	0.18	0.886	0.01	TOD × *Speed*	(3,66)	1.17	0.327	0.05	TOD × *Speed*	(3,66)	0.20	0.896	0.01
Sleep × TOD	(3,66)	2.30	0.086	0.09	Sleep × TOD	(3,66)	0.28	0.836	0.01	Sleep × TOD	(3,66)	3.16	**0.030**	0.13
Sleep × TOD × *Speed*	(3,66)	0.45	0.719	0.02	Sleep × TOD × *Speed*	(3,66)	0.92	0.437	0.04	Sleep × TOD × *Speed*	(3,66)	0.15	0.927	0.01
Stopping distance					Reaction time					Braking time				
Speed	(1,22)	195.43	**0.000**	0.90	*Speed*	(1,22)	190.42	**0.000**	0.90	*Speed*	(1,22)	0.16	0.691	0.01
Sleep	(1,22)	10.26	**0.004**	0.32	Sleep	(1,22)	12.65	**0.002**	0.37	Sleep	(1,22)	0.11	0.741	0.01
Sleep × *Speed*	(1,22)	8.54	**0.008**	0.28	Sleep × *Speed*	(1,22)	8.08	**0.009**	0.27	Sleep × *Speed*	(1,22)	0.00	0.976	0.00
TOD	(3,66)	0.34	0.796	0.02	TOD	(3,66)	0.08	0.923	0.00	TOD	(3,66)	1.54	0.212	0.07
TOD × *Speed*	(3,66)	0.37	0.775	0.02	TOD × *Speed*	(3,66)	0.79	0.465	0.03	TOD × *Speed*	(3,66)	0.50	0.686	0.02
Sleep × TOD	(3,66)	0.12	0.946	0.01	Sleep × TOD	(3,66)	0.19	0.880	0.01	Sleep × TOD	(3,66)	3.21	**0.038**	0.13
Sleep × TOD × *Speed*	(3,66)	0.83	0.480	0.04	Sleep × TOD × *Speed*	(3,66)	0.29	0.805	0.01	Sleep × TOD × *Speed*	(3,66)	1.95	0.142	0.08
Stopping time														
Speed	(1,22)	54.84	**0.000**	0.71										
Sleep	(1,22)	9.78	**0.005**	0.31										
Sleep × *Speed*	(1,22)	7.76	**0.011**	0.26										
TOD	(3,66)	0.29	0.805	0.01										
TOD × *Speed*	(3,66)	0.32	0.776	0.01										
Sleep × TOD	(3,66)	0.69	0.563	0.03										
Sleep × TOD × *Speed*	(3,66)	2.07	0.112	0.09										

Table 2
Repeated-measures analysis of variance (ANOVA) of Swerving. *Italic*: categorical factors. **Bold**: significant effect ($p < 0.05$). Speed (20; 40 kph), Side (left; right), Sleep: sleep condition (normal night; sleep deprivation), TOD: time-of-day (06:00 h; 10:00 h; 14:00 h; 18:00 h).

	ddl	F	p-value	partial η^2		ddl	F	p-value	partial η^2		ddl	F	p-value	partial η^2
Speed instruction					*Reaction distance*					*Adjustment distance*				
Speed	(1 45)	1796.40	**0.00**	0.98	*Speed*	(1,45)	113.00	**0.00**	0.72	*Speed*	(1,45)	7.07	**0.01**	0.14
Side	(1,45)	003	0.87	0.00	*Side*	(1,45)	0.28	0.60	0.01	*Side*	(1,45)	0.03	0.85	0.00
Sleep	(1,45)	1.24	0.27	0.03	Sleep	(1,45)	0.25	0.62	0.01	Sleep	(1,45)	0.62	0.43	0.01
Sleep × *Speed*	(1,45)	0.54	**0.04**	0.85	Sleep × *Speed*	(1,45)	0.02	0.89	0.00	Sleep × *Speed*	(1,45)	0.07	0.80	0.00
Sleep × *Side*	(1,45)	15.89	1.11	0.30	Sleep × *Side*	(1,45)	0,36	0.55	0.01	Sleep × *Side*	(1,45)	0.95	0.33	0.02
TOD	(3,135)	2.51	0.20	0.86	TOD	(3,135)	4.28	**0.01**	0.09	TOD	(3,135)	0.41	0.75	0.01
TOD × *Speed*	(3,135)	22.22	1.74	0.17	TOD × *Speed*	(3,135)	3.16	**0.03**	0.07	TOD × *Speed*	(3,135)	1.23	0.30	0.03
TOD × *Side*	(3,135)	23.36	1.82	0.16	TOD × *Side*	(3,135)	1.23	0.30	0.03	TOD × *Side*	(3,135)	0.02	0.99	0.00
Sleep × TOD	(3,135)	16.26	1.59	0.20	Sleep × TOD	(3,135)	0.68	0.57	0.01	Sleep × TOD	(3,135)	0.95	0.42	0.02
Sleep × TOD × *Speed*	(3,135)	6.77	0.66	0.57	Sleep × TOD × *Speed*	(3,135)	1.51	0.21	0.03	Sleep × TOD × *Speed*	(3,135)	0.67	0.57	0.01
Sleep × TOD × *Side*	(3,135)	34.59	3.38	0.02	Sleep × TOD × *Side*	(3,135)	0.41	0.75	0.01	Sleep × TOD × *Side*	(3,135)	0.51	0.68	0.01
Swerving distance					*Reaction time*					*Adjustment time*				
Speed	(1,45)	15.12	**0.00**	0.25	*Speed*	(1,45)	0.88	0.35	0.02	*Speed*	(1,45)	58.48	**0.00**	0.57
Side	(1,45)	0.06	0.81	0.00	*Side*	(1,45)	0.30	0.59	0.01	*Side*	(1,45)	0.00	0.96	0.00
Sleep	(1,45)	1.05	0.31	0.02	Sleep	(1,45)	0.02	0.89	0.00	Sleep	(1,45)	0.05	0.82	0.00
Sleep × *Speed*	(1,45)	0.15	0.70	0.00	Sleep × *Speed*	(1,45)	0.06	0.80	0.00	Sleep × *Speed*	(1,45)	0.64	0.43	0.01
Sleep × *Side*	(1,45)	0.41	0.52	0.01	Sleep × *Side*	(1,45)	0.41	0.52	0.01	Sleep × *Side*	(1,45)	1.87	0.18	0.04
TOD	(3,135)	0.34	0.80	0.01	TOD	(3,135)	4.08	**0.01**	0.08	TOD	(3,135)	0.03	0.99	0.00
TOD × *Speed*	(3,135)	1.03	0.38	0.02	TOD × *Speed*	(3,135)	2.79	**0.04**	0.06	TOD × *Speed*	(3,135)	0.76	0.50	0.02
TOD × *Side*	(3,135)	0.07	0.98	0.00	TOD × *Side*	(3,135)	1.38	0.25	0.03	TOD × *Side*	(3,135)	0.38	0.74	0.01
Sleep × TOD	(3,135)	1.28	0.28	0.03	Sleep × TOD	(3,135)	0.74	0.53	0.02	Sleep × TOD	(3,135)	1.11	0.34	0.02
Sleep × TOD × *Speed*	(3,135)	1.19	0.32	0.03	Sleep × TOD × *Speed*	(3,135)	1.70	0.17	0.04	Sleep × TOD × *Speed*	(3,135)	0.29	0.81	0.01
Sleep × TOD × *Side*	(3,135)	0.66	0.58	0.01	Sleep × TOD × *Side*	(3,135)	0.77	0.51	0.02	Sleep × TOD × *Side*	(3,135)	0.50	0.66	0.01
Swerving time														
Speed	(1,45)	64.57	**0.00**	0.59										
Side	(1,45)	0.04	0.84	0.00										
Sleep	(1,45)	0.02	0.88	0.00										
Sleep × *Speed*	(1,45)	0.60	0.44	0.01										
Sleep × *Side*	(1,45)	1.09	0.30	0.02										
TOD	(3,135)	0.02	0.99	0.00										
TOD × *Speed*	(3,135)	0.52	0.62	0.01										
TOD × *Side*	(3,135)	0.68	0.53	0.01										
Sleep × TOD	(3,135)	1.04	0.37	0.02										
Sleep × TOD × *Speed*	(3,135)	0.37	0.74	0.01										
Sleep × TOD × *Side*	(3,135)	0.59	0.59	0.01										

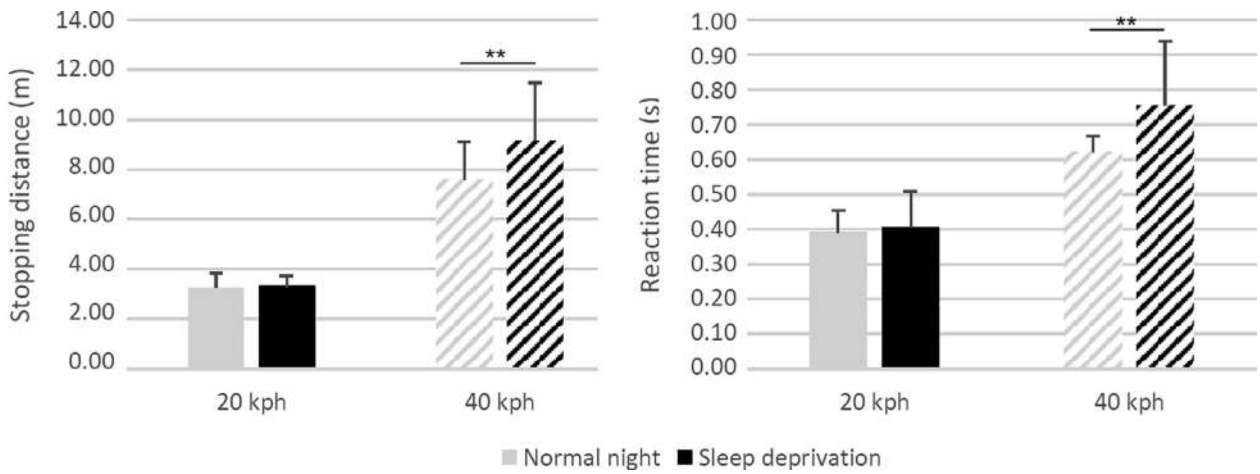

Fig. 5. Stopping distance (A) and reaction time (B) recorded at 20 kph (on the left) and 40 kph (on the right) after the normal night's sleep (grey) and sleep deprivation (black); $^{**}p < 0.01$.

As for the swerving part of the maneuver, a significant effect of 'speed' was observed on the distance ($F_{(1,45)} = 7.07$; $p < 0.05$; partial $\eta^2 = 0.14$) and time ($F_{(1,45)} = 58.48$; $p < 0.001$; partial $\eta^2 = 0.57$) necessary to adjust the trajectory of the PTW. More precisely, the participants needed shorter distance (27.14 ± 0.91 m vs 30.56 ± 1.44 m; −11%) but took longer to adjust their path (4.74 ± 0.14 s *vs* 3.36 ± 0.12 s; +41%) while riding at 20 kph than at 40 kph, respectively.

4. Discussion

This study aimed at determining whether motorcycling performances evaluated through crash avoidance maneuvers usually performed (namely emergency braking and swerving) evolved with sleep deprivation and time-of-day while evaluated on a dynamic simulator. Our major finding is that emergency braking performance was significantly affected by lack of sleep, more particularly at 40 kph. The increase observed in stopping time and distance when riding at 40 kph could be directly connected to the reaction part of the maneuver (+0.13 s, +21.4%). In contrast, it seems that swerving performance remained unchanged according to TOD or SD in our experiment. At 40 kph, only the reaction part of the maneuver evolved throughout the day with slower reactions in the early morning, while these influences had no consequence on the whole maneuver.

Oral temperature was measured in this study because it is considered as a major biomarker of circadian rhythmicity (Colquhoun, 1971). As classically observed in the literature, the measurements realized after the normal night's sleep fluctuated throughout the day, according to the circadian rhythm of core body temperature. The amplitude of this diurnal fluctuation (+0.7 °C) is classically reported in participants with intermediate chronotype and attests that the participants selected in this study presented a clear circadian rhythmicity (Reilly & Waterhouse, 2009). The diurnal fluctuation of temperature was preserved after the sleepless night with a decreased amplitude, induced by higher values recorded in the morning (Souissi, Sesboüé, Gauthier, Larue, & Davenne, 2003). This also confirmed that the test sessions were set up at appropriate time schedules to evaluate marked diurnal fluctuations in motorcycling performances.

The results observed in the emergency braking test confirmed that stopping distance increases with speed (Corno, Savaresi, Tanelli, & Fabbri, 2008). Our dynamic model induced a 5-m increase between maneuvers performed at 20 kph and 40 kph (+155%). Even if these results may seem intuitive, they reproduced previous measurements obtained in real riding conditions (Bougard et al., 2012; +9 m, +203%), confirming the pretty good generalizability of our results obtained on a dynamic simulator. While evolving at 20 kph faster, a 5 m difference in stopping distance may have serious consequences if a rider has to brake suddenly. Consequently, PTW riders have to respect speed limitations in dense areas. The stopping distance and time at 40 kph were increased after SD. These worsening performances (distance: +1.58 m, +21%; time: +0.16 s, +19%) indicated that participants were not able to stop as quickly as they did after the normal night's sleep, even at quite low speeds. This can have fatal consequences in real-life riding, particularly in a city center where speed limitations are fixed at 50 kph, with a lot of pedestrian crossings. Further analyses indicated that these impaired performances after the sleepless night should be mainly related to increased reaction times. After the normal night's sleep, reaction times measured at 20 kph (0.39 s) and 40 kph (0.62 s) were in agreement with several studies in real motorcycling condition, using artificial stimuli such as lights and road markings for safety reasons (Bougard et al., 2012; Davoodi, Hamid, Pazhouhanfar, & Muttart, 2012; Ecker, Wassermann, Ruspekhofer, Hauer, & Winkelbauer, 2001; Vavryn & Winkelbauer, 2004). At 40 kph, reaction times increased after SD, which is in agreement with previous findings obtained in laboratory conditions (Corsi-Cabrera, Arce, Del Rio-Portilla, Perez-Garci, & Guevara, 1999; Doran, vanDongen, & Dinges, 2001). Nonetheless, most of these studies also reported that reaction time still fluctuates throughout the day, independent of the lack of sleep (Arnal, Sauvet, Leger, van Beers, Bayon, Bougard, Rabat, Millet, & Chennaoui, 2015; Basner, Rao, Goel, & Dinges, 2013; Bougard, Davenne, Espie, Moussay, & Léger, 2016a). This was not the case on the simulator. It can be assumed that, due to monotony on the simulator in this relatively simple task, the effect of sleep deprivation was so important that the influence of diurnal fluctuations remained limited. Another interesting aspect is that reaction times were faster at 20 kph than at 40 kph. It has to be noticed that to ride at 20 kph, the participants only needed to select the third gear, while the engine was idling. They did not need to turn the throttle. In contrast, to ride at 40 kph the participants also needed to maintain their speed by controlling the throttle, similar to a real motorcycle, which is a bit more complex. Finally, the braking part of the maneuver remained mostly unchanged in our study. Only a significant effect of 'speed' was observed on the braking distance, but no effect of TOD nor SD. It is particularly difficult to reproduce the tire ground contact on a motorcycling simulator, even on a dynamic simulator. It can be inferred that the lack of rendering was a measuring weakness of this parameter. Nonetheless, in case of emergency, riders might press the brakes heavily in order to stop the PTW as quick as possible, in peculiar with an ABS system. These observations confirm those of Kovácsová et al. (2020), reporting no differences between braking styles observed with motion and no-motion simulator configurations. These results are also in agreement with those of Humphrey, Kramer, and Stanny (1994), reporting that the deleterious effects of SD on reaction time are mainly induced by an alteration of the first stages of information processing, the perception, but not from the motor execution part.

As for the swerving maneuver, the total distance and time necessary to perform the maneuver increased with speed. Riding 20 kph faster (from 20 to 40 kph), increased the avoidance distance by 5 meters approximately. These results are in agreement with previous observations in real motorcycling conditions (+7 meters between 20 and 40 kph; Bougard et al., 2012). This tends to indicate a good reproducibility of motorcycling behavior, but also of motorcycle dynamics on the simulator. In addition, our results confirm that up to 40 kph, and without any distraction, emergency braking requires shorter distance than swerving (Shuman, Husher, Varant, & Armstrong, 2006). Once again, in critical situations PTW riders must be aware of the impact of speed on their crash avoidance capabilities. Furthermore, analyzing the different components of the maneuver indicated that reaction times and distances evolved throughout the day. At 40 kph, the participants reacted slower at 06:00 h than at other times of the day, which resulted in longer distances travelled before initiating the maneuver (+1 m). These observations confirm that the diurnal fluctuation of attention, with increased reaction times in the early morning, evolves closely to the body temperature rhythm (Bougard et al., 2016a; Rabat et al., 2016). The swerving part of the maneuver only changed with 'speed.' According to PTW dynamics, the participants adjusted their trajectory faster, but while traveling more distance at 40 kph than at 20 kph (Shuman et al., 2006). As PTW riders use the right-hand side of the road in France, and usually need to avoid an obstacle by swerving on the left (e.g., door of a car parked could open, right of way violation), we believed that the swerving maneuver would be performed more efficiently on the left than on the right. However, the participants of our study, being quite young, may have lack of exposure to critical situations to develop such preference. Although the effects of SD are well-known, the swerving performances remained globally unchanged in our study. These observations are in agreement with those of Bougard et al. (2012). This maneuver was the more complex in our study, much more than the emergency braking. Indeed, the participants had to maintain speed and control their trajectory while entering the exercise zone, and then to react as quickly as possible to the light signal and perform the correct maneuver according to the side indicated by the LED. Previous studies have shown that compensatory mechanisms may be set up to mitigate SD effects in complex tasks (Drummond, Meloy, Yanagi, Orff, & Brown, 2005). Our results confirm that TOD and SD have different effects on the initial stages of information processing and those responsible for executive functions and motor execution (Frey, Badia, & Wright, 2004; Humphrey et al., 1994; Kraemer et al., 2000). These last mechanisms seem to be more resistant to the effects of both disturbance factors and performances remain mostly unaffected.

Some limitations should be considered in our study. The limited sample size reduces generalization of our results, in particular to other age ranges, since riders were quite young in our study. Moreover, it should be assumed that even if there was some uncertainty related to the side on which the maneuver had to be performed, the participants knew that they had to do something when entering the exercise zone. As a consequence, they were already in an alert state, and waited for a LED signal. Therefore, the practical significance of these results must be interpreted in the context of the contrived experimental conditions. However, this methodological choice was guided by possible comparisons with previous studies conducted in real motorcycling conditions. The total sleep deprivation was rather severe in our study. Nonetheless, previous studies using cumulative partial sleep deprivation, which are more comparable with what happens in everyday life, have shown that the performance decrement increased progressively with the accumulation of the sleep debt (Banks & Dinges, 2007; Van Dongen, Maislin, Mullington, & Dinges, 2003). It appeared that cumulative restriction of sleep to 6 h or less per night across 14 nights produced cognitive performance deficits equivalent to up to two nights of total sleep deprivation. Other studies also reported that 18, 21, and 28 waking hours had the same deleterious effect as a blood alcohol concentration of 0.05%, 0.08%, and 0.1%, respectively (Arnedt, Wilde, Munt, & MacLean, 2001; Williamson, Feyer, Mattick, Friswell, & Finlay-Brown, 2001). In addition, the short duration of the riding sessions may have allowed for compensatory mechanisms enabling limiting time-of-day or sleep deprivation influences (Bougard, Moussay, & Davenne, 2008). Finally, in contrast with results observed in real motorcycling conditions (Bougard et al., 2012), the influence of time-of-day on the simulator remained limited. It is well-known that sleepiness influence on driving capabilities is increased in simulated conditions (Davenne et al., 2012). It can be assumed that the effects of sleep deprivation were strong enough in our paradigm to mask time-of-day influence. In addition, even if motorcycling performance obtained in our dynamic simulator does not precisely replicate real motorcycling conditions, different measurements such as reaction times and stopping distances evolved in the same range between simulated and real driving conditions (Bougard et al., 2012). This tends to confirm the 'relative' validity of riding behaviors in the present study (Godley, Triggs, & Fildes, 2002).

5. Conclusion

This study is the first to demonstrate, on a dynamic PTW simulator, that motorcycling capabilities evolve with sleep deprivation and time-of-day, in reference to emergency braking and swerving maneuvers. The more complex tasks (swerving maneuver) used in this study were only weakly influenced, but the tests performed may have been too short. The PTW simulator, as it mimics the results obtained in real riding conditions at low speed, can be considered a useful tool to focus on more realistic scenarios. Further studies may focus on prolonged motorcycle riding sessions with various road environments (e.g., city, country, and highway) to better understand the impact of fatigue and sleepiness on riding behavior and crashes.

6. Practical Applications

Even during simple tasks at low speed, emergency braking and swerving maneuvers evolved with TOD alone or combined with SD. This suggests that, while riding for a longer duration and in a more complex environment (including numerous distractors), PTW riders should be aware of their limited capabilities at different TOD and/or when SD occurred.

Prevention campaigns for road safety should focus on the dangers of motorcycling while sleepy. As for sensitizing drivers to the dangers of alcohol intoxication using simple equipment (prismatic glasses), specific scenarios could be developed for low-cost PTW simulators such as the Honda Riding Trainer (HRT) which are, for example, already used in motorcycling schools in France.

Acknowledgments

We would like to thank all the participants of this study. Our thanks to Adam Prenzler for proofreading the manuscript.

Funding

This work was supported in part by a PREDIT-GO4 contract. Clément Bougard was granted his Ph.D. thesis by the Conseil Régional de Basse-Normandie (Regional Council of Lower Normandy) and the Institut National de Recherche sur les Transports et leur Sécurité (The French National Institute for Transport and Safety Research).

References

ACEM (2009). MAIDS: Motorcycle Accidents In-Depth Study, Final Report 2.0. http://www.maids-study.eu/pdf/MAIDS2.pdf.

Arnal, P.J., Sauvet, F., Leger, D., van Beers, P., Bayon, V., Bougard, C., Rabat, A., Millet, G.Y., & Chennaoui, M. (2015). Sleep, 1;38(12):1935-1943.

Arnedt, J. T., Wilde, G. J. S., Munt, P. W., & MacLean, A. W. (2001). How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? *Accident Analysis and Prevention, 33*(3), 337–344.

Banks, S., & Dinges, D. F. (2007). Behavioral and physiological consequences of sleep restriction. *Journal of Clinical Sleep Medicine, 03*(05), 519–528.

Basner, M., Rao, H., Goel, N., & Dinges, D. F. (2013). Sleep deprivation and neurobehavioral dynamics. *Current Opinion in Neurobiology, 23*(5), 854–863.

Baxter, C., & Reilly, T. (1983). Influence of time of day on all-out swimming. *British Journal of Sports Medicine, 17*(2), 122–127.

Benedetto, S., Lobjois, R., Faure, V., Dang, N.-T., Pedrotti, M., & Caro, S. (2014). A comparison of immersive and interactive motorcycle simulator configurations. *Transportation Research Part F: Traffic Psychology and Behaviour, 23*, 88–100.

Bougard, C., Davenne, D., Espie, S., Moussay, S., & Léger, D. (2016a). Sleepiness, attention and risk of accidents in powered two-wheelers. *Sleep Medicine Reviews, 25*, 40–51.

Bougard, C., Espié, S., Larnaudie, B., Moussay, S., Davenne, D., & Yamazaki, S. (2012). Effects of time of day and sleep deprivation on motorcycle-driving performance. *PLoS ONE, 7*(6), e39735.

Bougard, C., Moussay, S., & Davenne, D. (2008). An assessment of the relevance of laboratory and motorcycling tests for investigating time of day and sleep deprivation influences on motorcycling performance. *Accident Analysis and Prevention, 40*(2), 635–643.

Bougard, C., Moussay, S., Espié, S., & Davenne, D. (2016b). The effects of sleep deprivation and time of day on cognitive performance. *Biological Rhythm Research, 47*(3), 401–415.

Bougard, C., Moussay, S., Gauthier, A., Espié, S., & Davenne, D. (2009). Effects of waking time and breakfast intake prior to evaluation of psychomotor performance in the early morning. *Chronobiology International, 26*(2), 324–336.

Centola, C., Tagliabue, M., Spoto, A., Palpacelli, M., Giorgetti, A., Giorgetti, R., & Vidotto, G. (2020). Enhancement of unsafe behaviors in simulated moped-riding performance under the influence of low dose of alcohol. *Accident Analysis and Prevention, 136*, 105409.

Chen, Y. C., & Chen, P. L. (2016). Insufficient sleep hours associated with varied types of risky riding behaviors among motorcyclist. European Journal of Public Health, 26 (Suppl 1): ckw175.006.

Cheng, A. S. K., Ng, T. C. K., & Lee, H. C. (2011). A comparison of the hazard perception ability of accident-involved and accident-free motorcycle riders. *Accident Analysis and Prevention, 43*(4), 1464–1471.

Colquhoun, W. P. (1971). *Biological rhythms and human performance* (p. 283). New York: Academic Press.

Corno, M., Savaresi, S. M., Tanelli, M., & Fabbri, L. (2008). n optimal motorcycle braking. *Control Engineering Practice, 16*(6), 644–657.

Corsi-Cabrera, M., Arce, C., Del Rio-Portilla, I. Y., Perez-Garci, E., & Guevara, M. A. (1999). Amplitude reduction in visual event-related potentials as a function of sleep deprivation. *Sleep, 22*(2), 181–189.

Crundall, D., van Loon, E., Stedmon, A. W., & Crundall, E. (2012). Motorcycling experience and hazard perception. *Accident Analysis and Prevention, 50*, 456e64.

Davenne, D., Lericollais, R., Sagaspe, P., Taillard, J., Gauthier, A., Espié, S., & Philip, P. (2012). Reliability of simulator driving tool for evaluation of sleepiness, fatigue and driving performance. *Accident Analysis and Prevention, 45*, 677–682.

Davoodi, S. R., Hamid, H., Pazhouhanfar, M., & Muttart, J. W. (2012). Motorcyclist perception response time in stopping sight distance situations. *Safety Science, 50*(3), 371–377.

Day, L., Lenné, M. G., Symmons, M., Hillard, P., Newstead, S., Allen, T., & McClure, R. (2013). Population based case-control study of serious non-fatal motorcycle crashes. *BMC Public Health, 13*, 72.

de Lapparent, M. (2006). Empirical Bayesian analysis of accident severity for motorcyclists in large French urban areas. *Accident Analysis and Prevention, 38*(2), 260–268.

de Rome, L., Brown, J., Baldock, M., & Fitzharris, M. (2018). Near-miss crashes and other predictors of motorcycle crashes: Findings from a population-based survey. *Traffic Injury Prevention, 19*(sup2), S20–S26.

de Winter, J. C. F., Van Leeuwen, P. M., & Happee, P. (2012). Advantages and disadvantages of driving simulators: a discussion. In: Proceedings of measuring behavior, pp. 47–50 (Utrecht, The Netherlands).

Doran, S. M., vanDongen, H. P., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie, 139*(3), 253–267.

Drummond, S. P. A., Meloy, M. J., Yanagi, M. A., Orff, H. J., & Brown, G. G. (2005). Compensatory recruitment after sleep deprivation and the relationship with performance. *Psychiatry research, 140*(3), 211–223.

Ecker, H., Wassermann, J., Ruspekhofer, R., Hauer, G., & Winkelbauer, M. (2001). Brake reaction times of motorcycle riders. In: Proceedings of the international motorcycle safety conference (Orlando Florida, USA).

European Commission. Road Safety in the European Union – Trends, statistics and main challenges. Luxembourg: Publications Office of the European Union, 2018: available at: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/vademecum_2018.pdf.

Filtness, A. J., Rudin-Brown, C. M., Mulvihill, C. M., & Lenné, M. G. (2013). Impairment of simulated motorcycle riding performance under low dose alcohol. *Accident Analysis and Prevention, 50*, 608–615.

Frey, D. J., Badia, P., & Wright, K. P. (2004). Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation. *Journal of Sleep Research, 13*(4), 305–315.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis and Prevention, 34*(5), 589–600.

Haworth, N., & Mulvihill, C. (2005). *Review of motorcycle licensing and training*. Monash University Accident Research Centre, (Report N°240).

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*(2), 97–110.

Haworth, N., Smith, R., Brumen, I. & Pronk, N. (1997). Case Control Study of Motorcycle Crashes. Monash University Accident Research Centre, (Report N°CR 174).

Hosking, S. G., Liu, C. C., & Bayly, M. (2010). The visual search patterns and hazard responses of experienced and inexperienced motorcycle riders. *Accident Analysis and Prevention, 42*(1), 196–202.

Humphrey, D. G., Kramer, A. F., & Stanny, R. R. (1994). Influence of extended wakefulness on automatic and nonautomatic processing. *Human Factors, 36*(4), 652–669.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep, 14*(6), 540–545.

Kovácsová, N., Grottoli, M., Celiberti, F., Lemmens, Y., Happee, R., Hagenzieker, M. P., & de Winter, J. C. F. (2020). Emergency braking at intersections: A motion-base motorcycle simulator study. *Applied Ergonomics, 82*, 102970.

Kraemer, S., Danker-Hopfe, H., Dorn, H., Schmidt, A., Ehlert, I., & Herrmann, W. M. (2000). Time-of-day variations of indicators of attention: Performance, physiologic parameters, and self-assessment of sleepiness. *Biological Psychiatry, 48*(11), 1069–1080.

Lam, C., Pai, C. W., Chuang, C. C., Yen, Y. C., Wu, C. C., Yu, S. H., ... Guo, Y. (2019). Rider factors associated with severe injury after a light motorcycle crash: A multicentre study in an emerging economy setting. *PLoS ONE, 14*(6), e0219132.

Kuang, Y., Qu, X., Weng, J., & Etemad-Shahidi, A. (2015). How Does the Driver's Perception Reaction Time Affect the Performances of Crash Surrogate Measures?. PloS one, 10(9), e0138617. https://doi.org/10.1371/journal.pone.0138617.

Lobjois, R., & Mars, F. (2020). Effects of motorcycle simulator configurations on steering control and gaze behavior in bends. *Journal of Experimental Psychology, 26*(1), 108–123.

Millar, K. (1992). Some chronic problems in the methodology of performance task applied in clinical settings. In R. Ogilvie & R. Broughton (Eds.), *Sleep, arousal and performance* (pp. 31–53). Boston: Birkhauser.

Möller, H., Senserrick, T., Rogers, K., Sakashita, C., de Rome, L., Boufous, S., ... Ivers, R. (2020). Crash risk factors for novice motorcycle riders. *Journal of Safety Research, 73*, 93–101.

National Highway Traffic Safety Administration (NHTSA). NHTSA's Traffic Safety Facts. Motorcycles. (2019, August). Motorcycles: 2017 data (Updated, Traffic Safety Facts. Report No. DOT HS 812 785). US Department of Transportation. Washington, 2019. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812785.

ONISR (Observatoire national de Sécurité Routière ONISR). La sécurité routière en France, bilan de l'année 2018. (2019) at: https://www.interieur.gouv.fr/Publications/Statistiques/Securite-routiere/Bilan-de-l-accidentologie.

ONISR. (2007). La sécurité routière en France : bilan de l'année 2006. Documentation Française.

Quddus, M. A., Noland, R. B., & Chin, H. C. (2002). An analysis of motorcycle injury and vehicle damage severity using ordered probit models. *Journal of Safety Research, 33*(4), 445–462.

Rabat, A., Gomez-Merino, D., Roca-Paixao, L., Bougard, C., Van Beers, P., Dispersyn, G., ... Chennaoui, M. (2016). Differential kinetics in alteration and recovery of cognitive processes from a chronic sleep restriction in young healthy men. *Frontiers in Behavioral Neuroscience, 10*, 95.

Reilly, T., & Bambaeichi, E. (2003). Methodological issues in studies of rhythms in human performance. *Biological Rhythm Research, 34*(4), 321–336.

Reilly, T., & Waterhouse, J. (2009). Sports performance: Is there evidence that the body clock plays a role? *European Journal of Applied Physiology, 106*(3), 321–332.
Santos, W. J. D., Cêlho, V. M. D. S., Santos, G. B., & Ceballos, A. G. D. C. (2019). Work overload and risk behaviors in motorcyclists. *Rev Bras Enferm, 72*(6), 1479–1484.
Shuman, K., Husher, S., Varant, M., & Armstrong, C. (2006). Do I brake or do I swerve: Motorcycle crash avoidance maneuvering. Available: http://www.msfusa.org/imsc/proceedings/b-Shuman-DoIBrakeorDoISwerve.pdf. Accessed 2020 Sept 6.
Souissi, N., Bessot, N., Chamari, K., Gauthier, A., Sesboüé, B., & Davenne, D. (2007). Effect of time of day on aerobic contribution to the 30-s Wingate test performance. *Chronobiology International, 24*(4), 739–748.
Souissi, N., Sesboüé, B., Gauthier, A., Larue, J., & Davenne, D. (2003). Effects of one night's sleep deprivation on anaerobic performance the following day. *European Journal of Applied Physiology, 89*(3), 359–366.
Van Dongen, H. P., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26*(2), 117–126.
vanElslande P. (2002). Analyse approfondie d'accidents de deux-roues à moteur, défaillances fonctionnelles et scenarios d'erreurs. Rapport de convention DSCR-INRETS.
Vavryn, K. & Winkelbauer, M. (2004). Braking performance of experienced and novice motorcycle riders - Results of a field study. In: Proceedings of the 4th international conference on transport and traffic psychology (Nottingham, England).
Vlahogianni, E. I., Yannis, G., & Golias, J. C. (2012). Overview of critical risk factors in Power-Two-Wheeler safety. *Accident Analysis and Prevention, 49*, 12–22.
Wali, B., Khattak, A. J., & Ahmad, N. (2019). Examining correlations between motorcyclist's conspicuity, apparel related factors and injury severity score: Evidence from new motorcycle crash causation study. *Accident Analysis and Prevention, 131*, 45–62.
Wali, B., Khattak, A. J., & Khattak, A. J. (2018). A heterogeneity based case-control analysis of motorcyclist's injury crashes: Evidence from motorcycle crash causation study. *Accident Analysis and Prevention, 119*, 202–214.
Williamson, A. M., Feyer, A.-M., Mattick, R. P., Friswell, R., & Finlay-Brown, S. (2001). Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. *Accident Analysis and Prevention, 33*(3), 313–326.
Wang, C., Lu, L., Lu, J., & Wang, T. (2016). Correlation between crash avoidance maneuvers and injury severity sustained by motorcyclists in single-vehicle crashes. Traffic injury prevention, 17(2), 188–194. https://doi.org/10.1080/15389588.2015.1045502.
Wu, D., Hours, M., & Martin, J.-L. (2018). Risk factors for motorcycle loss-of-control crashes. *Traffic Injury Prevention, 19*(4), 433–439.
Zheng, Y., Ma, Y., Guo, L., Cheng, J., & Zhang, Y. (2019). Crash involvement and risky riding behaviors among delivery riders in China: The role of working conditions. *Transportation Research Record, 2673*(4), 1011–1022.

Clément Bougard After obtaining his Ph.D. in Sport sciences concerning time-of-day and sleep deprivation influences on motorcycle driving performance, he joined the French Armed Forces Biomedical Research Institute, where he worked for 6 years in the Fatigue and Vigilance team. He then studied the impacts of sleep/sleep-wake rhythm disturbances (acute/chronic) on human behavior (integrated physiology). He is currently a research engineer in the Cognitive Sciences and Human Factors team in STELLANTIS; and the Automotive Motion Lab lead. His works concern the monitoring of driver/passenger to characterize the person's condition. He co-supervises various research works applied to motion sickness, drowsiness, or even neuromuscular fatigue, posture in automotive.

Damien Davenne is Professor at the department of sport and exercise sciences of Caen-Normandy University. He also heads an Inserm research unit (UMR-S 1075). This research unit is specialized in the study of Mobilities, in particular during ageing, both in animal and in healthy and pathological human. He works on the circadian biological rhythmicity, its fundamental mechanisms involved in the body mechanisms with numerous works on the sleep-wake cycles in connection with the motor and cognitive performance. He is the author or co-author of more than 150 articles and book's chapters in fields such as chronobiology, neurophysiology, driving security, and effects of physical activity on health.

Sébastien Moussay Sébastien has a PhD in sports Sciences & Chronobiology from Caen University - France. After graduating in 2002, he started working as an associate professor (INSERM U1075), in charge of the development of technologies for health. Following research work carried out in close collaboration with industrial players, he founded BodyCAP in 2011. Sébastien holds a Habilitation to Direct Research and the author / co-author of 33 publications in international peer-reviewed journals.

Stéphane Espié is Director of Research at Gustave EIFFEL University. His main research topic relates to the design of tools to understand and simulate road users' behaviours (behavioural traffic simulation, driving simulators, vehicle instrumentation). Related to Powered Two Wheelers (PTW) safety, he impulsed and managed various collaborative research projects co-funded by the French National Research Agency (ANR), Insurance Foundations or EEC. He was noticeably the initiator and the scientific director of the 7th FP collaborative project 2BeSafe, and was a partner in the 7th SafeRider ITS focussing project. He currently participates to H2020 SimuSafe and DriveToTheFuture research projects.

Journal of Safety Research 78 (2021) 155–169

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Examining crash injury severity and barrier-hit outcomes from cable barriers and strong-post guardrails on Alabama's interstate highways

Mehdi Hosseinpour *, Kirolos Haleem

School of Engineering & Applied Sciences, Western Kentucky University, 1906 College Heights Blvd, EBS 2122, Bowling Green, KY 42101, United States

ARTICLE INFO

Article history:
Received 16 October 2020
Received in revised form 24 March 2021
Accepted 17 June 2021
Available online 1 July 2021

Keywords:
Safety barrier
Crash injury severity
Barrier-related crashes
Cable barriers
Strong-post guardrails

ABSTRACT

Introduction: This study investigates the impact of several risk factors (i.e., roadway, driver, vehicle, environmental, and barrier-specific characteristics) on the injury severity resulting from barrier-related crashes and also on barrier-hit outcomes (i.e., vehicle containment, vehicle redirection, and barrier penetration). A total of 1,685 barrier-related crashes, which occurred on three major interstate highways (I-65, I-85, and I-20) in the state of Alabama, were collected for a seven-year period (2010–2016), and all relevant information from the police reports was reviewed. Features that were rarely explored before (e.g., median width, barrier length, barrier offset or lateral position, left shoulder width, blockout type, and number of cables) were also collected and examined. Two types of longitudinal barriers were analyzed: high-tension cable barriers installed on medians and strong-post guardrails installed on medians and/or roadsides. *Method:* Two separate mixed logit (MXL) models were used to analyze crash injury severity in median and roadside barrier-related crashes. Two additional MXL models were separately adopted for median and roadside barrier-related crashes to estimate the probability of three barrier-hit outcomes (vehicle containment, vehicle redirection, and barrier penetration). *Results:* The results of crash injury severity MXL models showed that, for both median and roadside barrier crashes, barrier penetration, female drivers, and driver fatigue were associated with a higher probability of injury or fatal crashes. The results of barrier-hit MXL models showed that longer barrier length, Brifen cable barrier system, and barrier lateral position were significant predictors of median barrier-hit outcomes, whereas dark lighting condition, driving under the influence (DUI), presence of curved freeway sections, and right shoulder width significantly contributed to roadside barrier-hit outcomes. *Conclusions:* The MXL model succeeded in identifying several contributing factors of crash severity and barrier-hit outcomes along Alabama's interstate highways. *Practical applications:* One study application is to design longer barrier run length (greater than 1230 feet or 0.2 miles) to reduce the barrier penetration likelihood.

1. Introduction

Roadside appurtenances, such as guardrails and cable barriers, are designed to prevent crossover crashes or roadway departure (RwD) crashes. RwD crashes are defined as a crash in which an errant vehicle leaves its travel way to the left (median crossover) or to the right (run-off-road "ROR"), potentially colliding with a vehicle or fixed object. RwD crashes are among the most severe traffic crashes because a vehicle's roadway departure mainly results in striking a rigid fixed object, overturning on roadside embankments, or colliding head-on with another vehicle in the opposite direction after crossing the median (Chitturi et al., 2011; Roque & Jalayer, 2018; Zou, Tarko, Chen, & Romero, 2014). As indicated in Russo and Savolainen (2018), per Park et al. (2015), RwD crashes are responsible for over half of all traffic fatalities in the United States. Longitudinal safety barriers (such as W-beam/thrie-beam guardrails and cable barriers) are widely installed on medians and roadsides to prevent errant vehicles from severe consequences following roadway departure (Hu & Donnell, 2010; Molan et al., 2019; Park et al., 2016; Russo & Savolainen, 2018).

There exist three possible outcomes for any vehicle after leaving the roadway and striking a median or roadside barrier. These are: (i) *vehicle containment* (i.e., vehicle is contained by the barrier; which is an indication of barrier success or barrier non-crossover), (ii) *vehicle redirection* (i.e., vehicle is redirected back onto the roadway after colliding with the barrier; which is an indication of barrier success or barrier non-crossover), and (iii) *barrier penetration* (i.e., vehicle penetrates through, under, or over the barrier; which is an indication of barrier crossover or non-success).

* Corresponding author.
E-mail address: mehdi.hosseinpour@wku.edu (M. Hosseinpour).

https://doi.org/10.1016/j.jsr.2021.06.009
0022-4375/

The effectiveness of a safety barrier system in reducing or eliminating harmful events following a vehicle's roadway departure depends on a number of factors, such as barrier-specific conditions (e.g., barrier type, barrier placement, post type and spacing, barrier length, etc.), road geometric conditions (e.g., number of traffic lanes, shoulder width, road alignment), traffic volume, vehicle type, and weather conditions (Burns & Bell, 2016). It is thus recommended to estimate the performance of different barrier alternatives on barrier-hit outcomes and resulting injury severity (Park et al., 2016; Zou & Tarko, 2018). However, the effect of barrier-specific characteristics (i.e., barrier length, barrier lateral position, type of guardrail blockout, cable barrier type, and number of cable or strands) has not been extensively documented when analyzing the severity and hit outcomes of barrier-involved crashes (notable exceptions are Zou, Tarko, Chen, and Romero (2014), Russo and Savolainen (2018), Molan et al. (2019), and Rezapour et al. (2019)).

Recognizing the aforementioned gap, this study has two main objectives: (1) to analyze the impact of roadway, driver, vehicle, environmental, and barrier-specific factors on the severity of barrier hits using the mixed logit (MXL) model, and (2) to estimate the probability of three different barrier-hit outcomes (vehicle containment, vehicle redirection, and barrier penetration) given a barrier-related crash on median or roadside using the MXL model. This model has the advantage of varying the parameter estimates across the crash observations for reliable coefficient estimation. In this study, two barrier types were investigated: high-tension cable barriers installed on medians and strong-post W-beam/thrie-beam guardrails installed on both medians and roadsides.

To achieve the study objectives, extensive data collection effort has been made through reviewing, in detail, police crash reports and capturing relevant geometric roadside features. This study was based on a sample of 1,685 barrier-related crashes that occurred on three major interstate highways (I-65, I-85, and I-20) in the state of Alabama over the seven-year period from 2010 to 2016.

2. Literature review

2.1. Studies on the safety performance of road barriers

A number of studies have been conducted to investigate the in-service performance of road barriers in reducing the consequences of RwD crashes. Hunter et al. (2001), for example, applied the negative binomial (NB) regression models to investigate the effects of three-strand median cable barriers on crash rates for several collision types. Crash data were collected on North Carolina Interstate System for the years 1990 through 1997. The authors found that the number of severe crashes declined on the sections where cable median barriers were installed, though the frequency of some crash types, such as run-off-road-left and fixed object-related, increased. Ray and Weir (2001) evaluated the safety performance of four different guardrail systems, including the G1 cable guardrail, the G2 weak-post W-beam guardrail, and the G 4(1S) and G4(1 W) strong-post W-beam guardrails. Crash data were collected on different highways in the states of Connecticut, Iowa, and North Carolina between 1997 and 1999. The study found that there were no significant differences in the safety performance of the G1 and the G2 or the G1 and the G4(1 W). Nevertheless, occupant injuries were less likely in collisions with a G1 guardrail than in collisions with the G4 guardrails.

Zou and Tarko (2016) analyzed the probabilities of various hazardous events under different road and barrier conditions when a vehicle departed the roadway. The study focused on the performance of four types of road barriers (median concrete barriers, median W-beam guardrails, median high-tension cable barriers, and roadside W-beam guardrails) in preventing high-risk crashes from occurring. A multinomial logit (MNL) model was applied using 10-year crash records collected on freeway and non-freeway segments from 2003 to 2012 in the state of Indiana. The results showed that no median crossovers were observed where a median barrier was present. Cable barriers installed on the far-side edge of a median reduced the probabilities of barrier collisions and of vehicles' redirection into traffic. Roadside guardrails were found to diminish the risk of hazardous off-road crashes. Chimba et al. (2017) conducted a before-and-after analysis using paired t-test and empirical Bayes (EB) methods to examine the safety effectiveness of median cable barriers installed on Tennessee highways between 2006 and 2010. The overall findings indicated that the installation of cable barriers significantly reduced crash frequency and severity.

Using 10 years (2003–2012) of crashes on arterial roads in Indiana, Zou and Tarko (2018) estimated crash modification factors (CMFs) and average crash costs for different barrier-related crash types (concrete barriers, W-beam guardrails, and high-tension cable barriers). The NB and MNL models were, respectively, applied to estimate the number of barrier-related crashes and the probability of barrier-related harmful events. The study concluded that the total number of barrier-related crashes increased where median barriers were present. In terms of crash cost, the unit cost of a crash was reduced for cable barriers installed in medians wider than 50 feet and for concrete barriers and guardrails installed in medians narrower than 50 feet. Compared to other barrier types, median cable barriers were found to be the most cost-effective.

2.2. Studies on barrier-related crash severity analysis

Hu and Donnell (2010) adopted a nested logit model to assess the effects of roadway, environmental, and driver-related factors on median barrier crash severity on North Carolina interstate highways. The information on roadway and crash data was collected from the North Carolina Highway Safety Information System (HSIS) dataset between 2000 and 2004. Several risk factors (such as median barrier offset, travel speed, overturning, and curved road segment indicator) were found to be associated with median barrier crash severity. Furthermore, collisions with cable barriers were found to be less serious compared to collisions with other barriers. Zou, Tarko, Chen, and Romero (2014) assessed the effectiveness of barriers in reducing the injury severity of vehicle occupants in single-vehicle crashes that occurred on Indiana highways. A mixed-effects binary logistic regression model was developed using crash data collected between 2008 and 2012. The results indicated that striking all barrier types, including cable barriers, guardrails, and concrete barriers, reduced the probability of injury when compared to colliding with high-risk roadside object.

Park et al. (2016) conducted a before-after analysis using EB and full Bayesian (FB) techniques to examine the safety effectiveness of the installation of roadside barriers (including W-beam guardrails and concrete barriers) installed on freeway segments in Florida. The crash data were collected from the crash analysis reporting system (CARS) for the four-year before (2003–2006) and four-year after (2008–2011) periods. The results showed that roadside barriers were more effective in reducing severe ROR crashes that occurred during nighttime, under rainy weather conditions, and for middle and old age drivers. Russo and Savolainen (2018) compared the performance of different median barrier types in terms of crash frequency, severity, and barrier-hit outcomes on Michigan freeway segments. A random-parameters negative binomial (RPNB) model and a random-parameters ordered logit (RPOL) model were developed to estimate crash frequency and severity, respectively. An MNL model was also fitted to estimate barrier strike outcomes. The studied median barrier types included high-tension cable barriers, thrie-beam guardrails, and concrete barriers. The authors found that the effects of factors related to roadway, traffic, and environmental attributes on the frequency, severity, and barrier-hit outcomes varied across the three different barrier types.

Rezapour et al. (2019) investigated the factors affecting the severity of barrier-hitting crashes on two-lane highways in Wyoming. A mixed logit model was applied to traffic barrier crashes that occurred between 2007 and 2017. The results showed that non-normal (e.g., anger and anxious) driving conditions, driver's citation record, older drivers, rollover crashes, barrier offset distance, barrier height, shoulder width, and dry surface condition were associated with higher severity of crashes involving roadside safety barriers.

Molan et al. (2019) developed a random-parameters ordered logit model to examine the effects of driver attributes, environmental conditions, roadway geometric design, and barrier-specific characteristics (e.g., barrier type: cable barriers, box-beam barriers, W-beam barriers, and concrete barriers), height, post-spacing, side slope ratio, and lateral offset, on the severity of median barrier crashes that occurred on interstate highways in Wyoming between 2016 and 2018. The results showed that female drivers, unbelted drivers, drivers with a record of alcohol citation, sharp horizontal curves, motorcycles, dry surface conditions, road segments with higher trucks, and rollover crashes were associated with more severe injury crashes involving median barriers. In terms of barrier-specific conditions, barrier height, barrier type, barrier lateral offset, and barrier post-spacing were found to significantly affect the risk of severe injury upon crash occurrence. In a similar study, Molan et al. (2020) investigated the severity of single-vehicle crashes involving median and roadside barriers. Using crashes occurred in the state of Wyoming from 2007 through 2016, three separate ordered logistic regression models were developed for three barriers including cable, guardrail, and rigid barriers. The authors demonstrated that the effects of different explanatory variables varied across safety barriers. For example, higher speed limits had a positive effect on crash severity of guardrails, while it had a reverse impact on the severity of crashes involving cable barriers.

2.3. Summary of the literature review

Table 1 summarizes previous studies discussed in the aforementioned two sub-sections. The review of the existing literature indicates that few studies have been carried out to investigate the probability of different barrier-hit outcomes (e.g., vehicle containment and barrier penetration) in the case of colliding with road safety barriers. Furthermore, relatively limited studies have investigated the effects of barrier-specific characteristics (such as barrier length, barrier lateral position, type of guardrail blockout, cable barrier type, embankment distance behind the guardrail posts, and number of cable or strands) on consequences following hitting a barrier, including injury severities or barrier-hit outcomes. For this, the current study aims to examine the effects of various barrier-specific factors, along with roadway, vehicle, driver, and environmental characteristics, on the consequences of barrier-related crashes, including the crash injury severity and barrier-hit outcomes (i.e., vehicle containment, vehicle redirection, and barrier penetration).

3. Study sites & data collection and preparation

The data used in this study were collected on three major interstate highways in the state of Alabama. These are I-65, I-85, and I-20. Interstate-65 is the most used interstate corridor in Alabama. It runs north–south and spreads for 366 miles in the state of Alabama. Its southmost point ends in Mobile, Alabama and it extends north through the Alabama/Tennessee border. I-65 runs through major cities in Alabama, including Mobile, Montgomery (the capital city of Alabama), Birmingham, and Huntsville. Interstate-85 is a major corridor that has wide existence of strong-post guardrails and median high-tension cable barriers. I-85 connects Montgomery with other cities in the south and mid-Atlantic, including Atlanta in Georgia, Charlotte in North Carolina, and Petersburg in Virginia. It runs nominally north–south and has an 80-mile extension in Alabama. Finally, Interstate-20 travels approximately 214 miles through the center of the state. I-20 enters Alabama from the Mississippi/Alabama border on the west and travels northeastward through Tuscaloosa and Birmingham until it crosses the Alabama/Georgia border on the east. A map highlighting the three interstate study sites in Alabama is shown in Fig. 1.

3.1. Crash data collection & review of police reports

Two types of longitudinal barriers were investigated in the study: high-tension cable barriers (installed on medians) and strong-post W-beam/thrie-beam guardrails (installed on medians and/or roadsides). A total of 1,685 crashes (2010 through 2016), involving median and roadside guardrail hits along three major interstates in the state of Alabama (I-65, I-85, and I-20), were downloaded from Alabama's Critical Analysis Reporting Environment (CARE) and their corresponding police reports were retrieved from ALDOT's internal web portal.

The police report's narrative and sketch were used to categorize crashes as crossover (i.e., barrier penetration or over-ride) and non-crossover crashes (i.e., vehicle either contained or redirected by the barrier), identify involved vehicle type, and record the crash injury severity. Note that penetrations and over-rides were aggregated together due to the limited sample size for each.

In this study, the crash injury severity for the most-severely injured occupant was used as the response variable. The injury severities are categorized into the five-level KABCO scale, where "K" represents fatal, "A" represents incapacitating injury, "B" represents non-incapacitating injury, "C" represents minor injury, and "O" represents no injury or property damage only (PDO). Due to the low number of observations for fatal and incapacitating injury outcomes, the five-level KABCO scale was further grouped into two main categories: (1) injury/fatal crash (while combining the KABC injury categories) and (2) non-injury (O) crash. Such dichotomous injury outcomes have also been used in previous studies to provide sufficient sample size for modeling (Chimba et al., 2017; Rezapour et al., 2019; Sacchi et al., 2012; Saleem & Persaud, 2017). Of the 1,685 barrier-related crashes, 387 (23%) resulted in injury/fatal crashes and 1,298 (77%) had no injuries. Crash data were then split into median- and roadside-barrier crashes, where the former and the latter accounted for 211 (55%) and 176 (45%) of injury/fatal crashes, respectively. In terms of barrier-hit outcomes, median cable barriers resulted in the least injury outcomes, involving only 24% of the total injury/fatal crashes, while median guardrails and roadside guardrails comprised 31% and 45% of the total injury/fatal crashes, respectively.

3.2. Roadway geometric and Barrier-Related variables data collection

In addition to police reports review, the following variables were collected by the research team:

1. Barrier-related variables: These include the type of traffic barrier hit by the vehicle during the crash (i.e., guardrail or cable barrier) and the length of the traffic barrier. If the crash involved hitting a cable barrier, then the type of cable barrier (Gibraltar, Brifen, and CASS) and the number of strands were recorded. If the

Table 1
Summary of previous studies on safety of road barriers.

Authors	Safety Barrier Type	Case Study (Years)	Method of Analysis	Main Findings
Hunter et al. (2001)	Median: three-strand cable barriers	North Carolina interstate highways (1990–1997)	Negative binomial (NB) model	• The number of severe crashes declined on road sections with median cable barriers. • The frequency of other crash types, such as run-off-road-left and fixed object-related, increased.
Ray and Weir (2001)	Median and roadside: cable barriers, and strong-post and weak-post W-beam guardrails	Roadways of different functional classifications in Connecticut, Iowa, and North Carolina (1997–1999)	Observational analyses of crash frequency, occupant injury, and barrier damage	• No significant differences were found in the safety performance of the G1 and G2 or G1 and G4(1 W) guardrails. • Occupant injuries were less likely in collisions with G1 guardrails than in collisions with G4 guardrails.
Hu and Donnell (2010)	Median: cable barriers, guardrails, and concrete barriers	North Carolina interstate highways (2000–2004)	Nested logit model	• Collisions with cable barriers resulted in less severe outcomes compared to collisions with other barriers. • Median barrier offset affected the probability of severe injuries.
Zou, Tarko, Chen, and Romero (2014)	Median and roadside: cable barriers, guardrails, and concrete barriers	Indiana highways (2008 – 2012)	Mixed-effects binary logistic model	• Striking all studied barrier types reduced the probability of injuries compared to those strikes with a high-risk roadside object.
Park et al. (2016)	Roadside: W-beam guardrails and concrete barriers	Florida freeways (2003–2011)	Empirical Bayes (EB) and full Bayesian (FB)	• Roadside barriers were more effective in reducing severe run-off-road (ROR) crashes that occurred during nighttime, under rainy weather conditions, and for middle and old age drivers.
Zou and Tarko (2016)	Median: concrete barriers, high-tension cable barriers, and roadside W-beam guardrails	Indiana freeway and non-freeways (2003–2012)	Multinomial logit (MNL) model	• No median crossovers occurred where a median barrier was present. • Roadside guardrails reduced the risk of hazardous off-road crashes.
Chimba et al. (2017)	Median: cable barriers	Tennessee highways (2006–2010)	Paired *t*-test and EB method	• Median cable barriers significantly reduced the crash frequency and severity along the studied road segments.
Russo and Savolainen (2018)	Median: high-tension cable barriers, thrie-beam guardrails, and concrete barriers	Michigan freeways	Random-parameters negative binomial (RPNB), random-parameters ordered logit (RPOL), and MNL models	• The effects of roadway, traffic, and environmental attributes on the frequency, severity, and barrier-hit outcomes varied across the three different barrier types (i.e., high-tension cable barriers, thrie-beam guardrails, and concrete barriers).
Zou and Tarko (2018)	Median: concrete barriers, W-beam guardrails, and high-tension cable barriers	Indiana roadway arterials (2003–2012)	NB and MNL models	• Total number of barrier-related crashes was higher where median barriers were present. • Relative to other barrier types, median cable barriers were found to be the most cost-effective.
Molan et al. (2019)	Median: cable barriers, box-beam barriers, W-beam barriers, and concrete barriers	Wyoming interstate highways (2016–2018)	RPOL model	• Female drivers, unbelted drivers, drivers with a record of alcohol citation, sharp horizontal curves, motorcycles, dry surface conditions, road segments with higher trucks, and rollover crashes were related to more severe injury crashes involving median barriers. • Barrier type, barrier height, barrier lateral offset, and barrier post-spacing were found to significantly affect the risk of severe injury crashes.
Rezapour et al. (2019)	Roadside: box-beam barriers, W-beam barriers, and concrete barriers	Wyoming two-lane highways (2007–2017)	Mixed logit model	• Non-normal (e.g., anger and anxious) driving conditions, driver's citation record, older drivers, rollover crashes, barrier offset distance, barrier height, shoulder width, and dry surface condition were associated with a higher severity of crashes involving roadside safety barriers.
Molan et al. (2020)	Median and roadside: cable barriers, guardrails, and rigid barriers	Wyoming highways (2007–2016)	Ordered logistic model	• The impact of the explanatory variables varied by safety barrier type. • Higher speed limits had a positive effect on the crash severity involving guardrails, while it had a reverse impact on the severity of crashes involving cable barriers.

Fig. 1. Map of three studied interstates (I-65, I-85, and I-20) in the State of Alabama.

crash involved hitting a guardrail, then the type of blockout was recorded. The location of the barrier, either on the roadside or on the median, was also recorded. The information about the aforementioned variables was collected via Google Maps (2018) back to the year of crash occurred.

2. *Geometric design variables*: These include median and roadside elements, such as: median width, median type, posted speed limits, right and left shoulder widths, presence of left/right rumble strips, barrier offset distance or lateral position (i.e., distance from the barrier to the median edge), number of lanes, and embankment distance behind the guardrail posts (for roadside guardrail crashes). All these variables were collected using Google Maps (2018) back to the specific year where the crash occurred using the Street View time slider on Google Maps (2018) (refer to Fig. 2).

3. *Traffic volume:* This includes annual average daily traffic (AADT) collected on the three interstates' segments. Traffic volumes were counted by traffic counter stations installed along the three studied interstate highways (I-20, I-65, and I-85). The relevant information on AADT was obtained from the Alabama Department of Transportation (ALDOT) and overlaid on each interstate's segments using ArcMap 10.2 (ESRI, 2019).

Fig. 2 shows four sample variables obtained using Google Maps (2018). The figure also shows sample strong-post guardrail system (having a composite blockout). Blockouts are mainly installed on guardrail barrier systems to allow for proper separation of the rail from the post and to absorb the energy resulted from the collision. A screenshot of the Street View time slider on Google Maps (2018), which allows traveling back to the specific crash year, is also shown in Fig. 2. In all, close to 90 variables were collected by the research team and stored in an Excel Spreadsheet. Table 2 presents descriptive statistics of all explanatory variables.

The total lengths of safety barriers scanned in this study were 553 miles. Of which, I-65, I-20, and I-85 constituted with 66% (366 miles), 19% (107 miles), and 15% (80 miles) of the total lengths, respectively. In terms of barrier-related crashes, a total of 696 roadside guardrail crashes (RGCs) occurred on these three

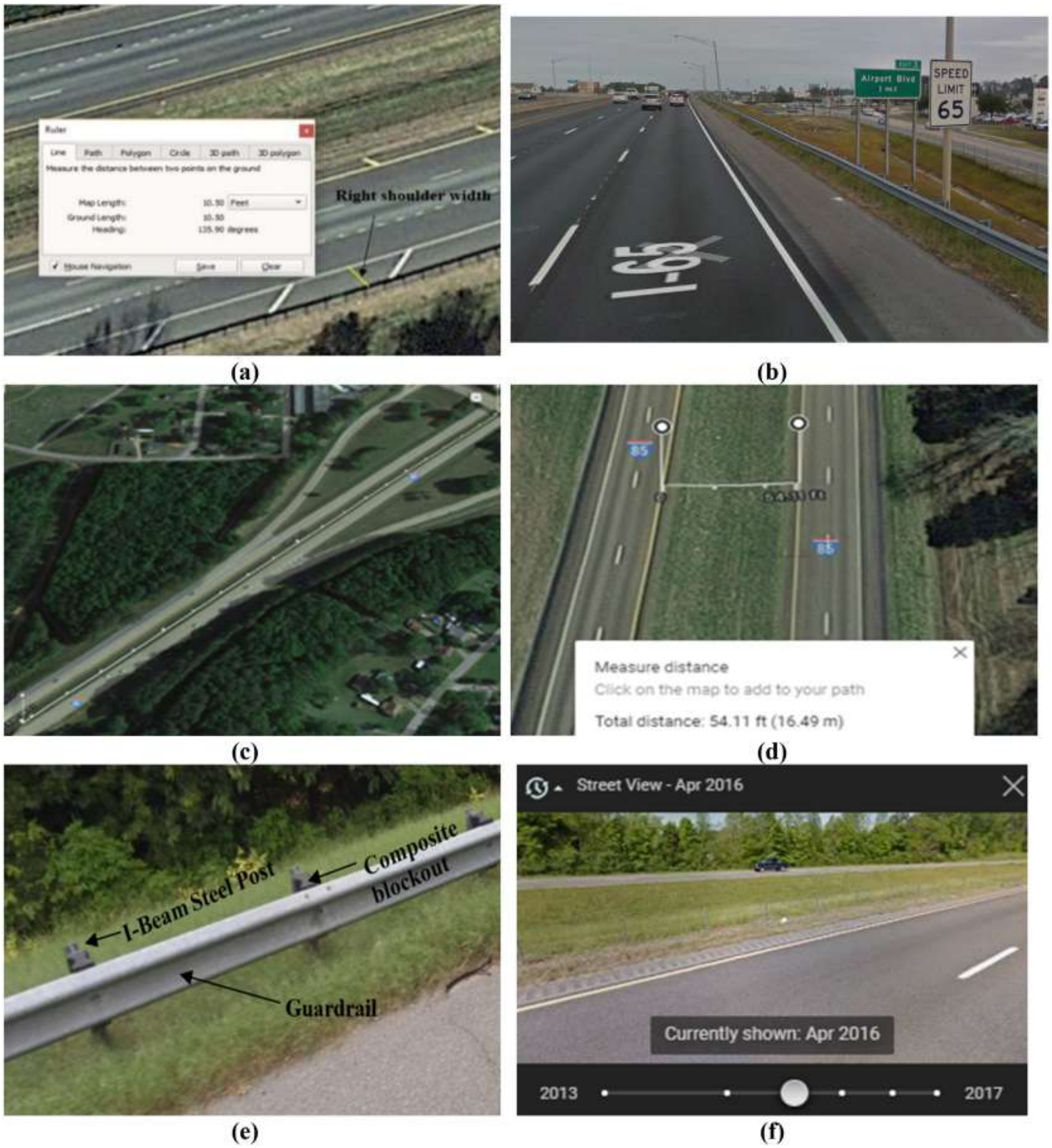

Fig. 2. Sample data collection for some variables on the interstates using Google Maps (2018): (*a*) measuring right shoulder width, (*b*) measuring posted speed limit, (*c*) measuring barrier length, (*d*) measuring median width, (*e*) strong-post guardrail system with composite blockout, and (*f*) Google street view time slider.

interstates during the study period, while 989 were attributed to median barrier crashes. Of the latter, 589 resulted in median cable barrier crashes (MCBCs), while the rest were median guardrail crashes (MGCs).

For the studied I-20 running from milepost 107 to 214.4, aggregately 10 miles of the interstate experienced MCBCs during the study period. Also, about 41 and 93 miles of I-20 involved MGCs and RGCs, respectively. For I-65, entirely 85 miles experienced MCBCs, over 153 miles experienced MGCs, and 282 miles experienced RGCs. For I-85, aggregately 53.8, 50.7, and 74.9 miles of the interstate experienced MCBCs, MGCs, and RGCs, respectively.

4. Methodology

This study adopted the MXL modeling approach to investigate the effects of various factors on crash injury severity, as well as barrier-hit outcomes in median and roadside barrier-related crashes. Regarding the injury severity modeling, two MXL models were separately developed for median and roadside barrier crashes to estimate the probability of killed or injured outcome versus no-injury outcome. Two separate MXL models were also fitted, one for median barrier and another for roadside barrier crashes, to estimate the barrier-hit outcomes, which were classified into three different outcomes (vehicle containment, vehicle redirection onto roadway, and barrier penetration). Compared to the traditional multinomial logit (MNL) model, MXL model can appropriately capture unobserved heterogeneity, which may result from unmeasured factors, by allowing a subset of explanatory variables to vary across observations. This enables the MXL model to yield more reliable parameter estimates than the traditional MNL model, where the latter restricts the effects of all explanatory variables to be fixed and identical across the observations.

The MXL model has widely been used in previous road safety research (Haleem et al., 2015; Hosseinpour et al., 2018; Islam & Brown, 2017; Li et al., 2019; Maistros et al., 2014; Manner & Wünsch-Ziegler, 2013; Russo & Savolainen, 2018; Yu & Abdel-Aty, 2014). The propensity function (S_{ni}) of a barrier-related crash n falling into category i is written as

$$S_{ni} = \beta_i X_n + \varepsilon_{ni} \quad (1)$$

Table 2
Descriptive statistics of explanatory variables.

Variable	Variable Description	Frequency	Percentage
Crash Characteristics			
Crash Injury Severity			
PDO	1 if true, otherwise 0	1298	77.0%
Injury/fatal	1 if true, otherwise 0	387	23.0%
Seatbelt Use	If seatbelt was used = 1, otherwise = 0	1543	91.7%
Airbag Deployment	If airbag deployed = 1, otherwise = 0	366	21.8%
Barrier-Hit Outcome			
Contained by the barrier	If vehicle was contained by the barrier = 1, otherwise = 0	701	41.7%
Redirected onto roadway	If vehicle was redirected onto the roadway = 1, otherwise = 0	799	47.5%
Penetrated the barrier	If vehicle penetrated through the barrier = 1, otherwise = 0	156	9.3%
Unknown	If barrier-hit outcome is unknown = 1, otherwise = 0	29	1.7%
Barrier Characteristics			
Barrier Type			
Median cable barrier	If type of median barrier was high-tension cable barrier = 1, otherwise = 0	589	35.0%
Median guardrail	If type of median barrier was guardrail = 1, otherwise = 0	400	23.8%
Roadside guardrail	If type of roadside barrier was guardrail = 1, otherwise = 0	696	41.4%
Cable Barrier System			
Gibraltar	If cable system is Gibraltar = 1, otherwise = 0	450	26.8%
Brifen	If cable system is Brifen = 1, otherwise = 0	100	5.9%
CASS	If cable system is CASS = 1, otherwise = 0	39	2.3%
Number of Cable Strands	If four-strand cable = 1, if three-strand cable = 0	294	17.5%
Barrier Length			
Short	If barrier length of coverage is shorter than 616 feet = 1, otherwise = 0	573	34.1%
Medium	If barrier length of coverage is between 616 and 1230 feet = 1, otherwise = 0	195	11.6%
Long	If barrier length of coverage is longer than 1230 feet = 1, otherwise = 0	917	54.5%
Guardrail Blockout Type			
Composite	If composite blockout = 1, otherwise = 0	812	48.19%
Steel	If steel blockout = 1, otherwise = 0	70	4.15%
Wooden	If wooden blockout = 1, otherwise = 0	29	1.72%
Mixed (Mix of more than one type)	If mixed blockout = 1, otherwise = 0	185	10.98%
Guardrail Configuration	If double-faced = 1, if single-faced = 0	160	9.5%
Barrier Lateral Position	If the distance between median barrier and median edge "or yellow stripe" is less than or equal to 8 feet = 1, otherwise = 0	446	26.5%
Roadway/Roadside Characteristics			
AADT/10,000*	Min. = 1.83, Max. = 13.32, Mean = 5.05, Std. Dev. = 2.59	–	–
Posted Speed Limit (mph)*	Min. = 45, Max. = 70, Mean = 68.69, Std. Dev. = 4.21	–	–
Area Type	If rural = 1, otherwise = 0	1,242	73.8%
Right Shoulder Width (ft)*	Min. = 1.46, Max. = 19.7, Mean = 10.41, Std. Dev. = 1.97	–	–
Right Shoulder Rumble Strips	If not present = 1, otherwise = 0	276	16.4%
Left Shoulder Width (ft)*	Min. = 0, Max. = 29.4, Mean = 5.86, Std. Dev. = 3.16	–	–
Left Shoulder Rumble Strips	If not present = 1, otherwise = 0	200	11.9%
Work Zone	If crash occurred at work zone = 1, otherwise = 0	106	6.3%
Road Alignment	If curved = 1, otherwise = 0	175	10.4%
Vertical Alignment	If upgrade or downgrade exists = 1, otherwise = 0	480	28.5%
Total Number of Traffic Lanes			
Four lanes	If four lanes = 1, otherwise = 0	1,349	80.0%
More than four lanes	If more than four lanes = 1, otherwise = 0	336	20.0%
Median Width	If greater than 55 feet = 1, otherwise = 0	700	41.6%
Roadside Embankment	If present = 1, otherwise = 0	545	32.4%
Embankment Width			
EMB2	If embankment width is less than 2 feet	176	10.5%
EMB2-8	If embankment width is between 2 and 8 feet	78	4.6%
EMB8	If embankment width is greater than 8 feet	1,431	85.1%
Environmental Characteristics			
Season of Year			
Winter	If crash occurred in winter = 1, otherwise = 0	414	24.6%
Spring	If crash occurred in spring = 1, otherwise = 0	419	24.9%
Summer	If crash occurred in summer = 1, otherwise = 0	495	29.4%
Autumn	If crash occurred in autumn = 1, otherwise = 0	357	21.2%
Day of Week	If crash occurred on weekend = 1, otherwise = 0	576	34.2%
Time of Day			
Daytime	If crash occurred in daytime (7:00 AM – 6:59 PM) = 1, otherwise = 0	1,005	59.8%
Evening	If crash occurred in the evening/at night (7:00 PM–11:59 PM) = 1, otherwise = 0	249	14.8%
Midnight	If crash occurred at late night/early morning (12:00 AM – 6:59 AM) = 1, otherwise = 0	431	25.6%
Weather Condition			
Clear/Cloudy	If clear/cloudy = 1, otherwise = 0	1,041	61.9%
Foggy	If foggy = 1, otherwise = 0	246	14.6%
Rainy/Snowy	If rainy/snowy = 1, otherwise = 0	366	21.8%
Unknown	If weather was unknown = 1, otherwise = 0	32	1.9%
Lighting Condition	If dark = 1, otherwise = 0	685	40.7%
Road Surface Condition	If road surface was wet or slippery at the time of crash = 1, otherwise = 0	540	32.1%

(*continued on next page*)

Table 2 (*continued*)

Variable	Variable Description	Frequency	Percentage
Driver Characteristics (Driver of Vehicle Hitting the Barrier)			
Driver Gender	If female = 1, otherwise = 0	634	37.7%
Driver Age**			
Young	If driver was 30 years or under	755	44.9%
Middle	If driver was between 31 and 49 years old	518	30.8%
Old	If driver was 50 years or above	376	22.4%
Unknown	If driver age was unknown	36	2.1%
Driver License Status	If license issued in Alabama (in-state) = 1, otherwise (out-of-state) = 0	1,105	65.7%
Crash Primary Cause			
Fatigue/sleepiness	If "driver was fatigued or asleep" was the primary cause = 1, otherwise = 0	302	18.0%
DUI "Driving under the Influence"	If "driving under the influence of alcohol or drug" was the primary cause = 1, otherwise = 0	92	5.5%
Aggressive driving	If "aggressive driving (e.g., speeding, improper lane change, following too close)" was the primary cause = 1, otherwise = 0	131	7.8%
Distraction	If "distraction or inattention" was the primary cause = 1, otherwise = 0	118	7.0%
Loss of control	If "vehicle loss of control" was the primary cause = 1, otherwise = 0	788	46.8%
Defective equipment	If "defective equipment (tire blow-out, improperly loaded trailer, load shift)" was the primary cause = 1, otherwise = 0	160	9.5%
Others	If other factor was the primary cause = 1, otherwise = 0	94	5.6%
Vehicle Characteristics			
Vehicle Type			
Passenger car	If vehicle was passenger car = 1, otherwise = 0	891	53.0%
SUV/pick-up	If vehicle was sport utility vehicle (SUV) or pick-up = 1, otherwise = 0	568	33.8%
Heavy vehicle	If vehicle was heavy vehicle (e.g., large trucks or buses) = 1, otherwise = 0	183	10.9%
Others	If other vehicle types = 1, otherwise = 0	43	2.6%

* *Continuous variable (non-discrete).*
** *Similar studies (e.g., Baker et al. (2003) and Li et al. (2018)) used "30 years" and "50 years" as thresholds to separate different age groups.*

where X_{ni} is a vector of K explanatory variables; β is a vector of K estimable coefficients and may vary across crash observations; and ε_{ni} denotes the error term.

The term ε_{ni} is assumed to follow the generalized extreme value distribution. The probability that a barrier-related crash n falls into category i (P_{ni}) is specified as follows (Islam & Brown, 2017):

$$P_{ni} = \frac{\exp(\beta_i X_{ni})}{\sum_i \exp(\beta_i X_{ni})} \tag{2}$$

The MXL model is a generalized form of the MNL model, where the former is able to accommodate unobserved heterogeneity by allowing some of explanatory variables to vary across observations. The MXL model is thus now specified as follows:

$$P_{ni} = \int \frac{\exp(\beta_i X_{ni})}{\sum_i \exp(\beta_i X_{ni})} f(\beta_i|) d\beta \tag{3}$$

where $f(\beta_i|)$ is the density function; and φ is a vector of parameters that is associated with the density function (mean and variance).

In this study, the model parameters were specified to follow the normal distribution, that is, $\beta\ N(b, \sigma)$, where b is the mean and σ is the standard deviation. A parameter was considered as random if the corresponding standard deviation was statistically significant. This study adopted the maximum simulation likelihood method to estimate the MXL models. The parameter estimates were based on 200 Halton draws, which have been reported to produce reliable parameter estimates (Anastasopoulos & Mannering, 2009; Hosseinpour et al., 2018; Russo & Savolainen, 2018; Wang et al., 2011). The MXL models in this study were developed using a user-written command that is referred to as *mixlogit* in the statistical software STATA (Hole, 2007).

Due to the inclusion of random parameters in the MXL model, the model coefficients alone are not sufficient to explain the real impacts of significant explanatory variables on the probability of a specific barrier-crash category (e.g., injury severity or barrier-hit outcome; Kim et al., 2013; Hosseinpour et al., 2014). To better interpret the effects of significant variable on the probabilities of barrier-hit consequences, the corresponding marginal effects were calculated and presented in this paper. The marginal effects reflect the change in the probability of a specific barrier crash outcome due to one unit change in the value of a continuous variable or the effect on the injury outcome of an indicator variable changing from zero to one (Al-Bdairi & Hernandez, 2017; Behnood & Mannering, 2017). The marginal effects of continuous and binary indictor variables were computed using Eqs. (4) and (5), respectively (Wu et al., 2014):

$$E^{P_{ni}}_{X_{nik}} = \frac{\partial P_{ni}}{\partial X_{nik}} \frac{X_{nik}}{P_{ni}} \tag{4}$$

$$E^{P_{ni}}_{X_{nik}} = \frac{P_{ni}[X_{nik} = 1] - P_{ni}[X_{nik} = 0]}{P_{ni}[X_{nik} = 0]} \tag{5}$$

where X_{nik} is the kth variable value for barrier-related crash n in the propensity function of barrier crash category i.

5. Results and discussion

5.1. Analysis of injury severity outcome

Table 3 presents the results of two separate MXL models estimated for the injury severity of each of median and roadside barrier-related crashes. The no-injury outcome was selected as the reference case. Hence, the estimated coefficients represent the probability of injury or fatal outcomes relevant to no-injury level. A deviance statistic was used in this study to show the overall goodness-of-fit for the developed models. The test follows the Chi-squared (x^2) distribution with degrees of freedom equal to the number of predictors in the final model. The values of deviance statistic were found to be highly significant for the median and roadside barrier-related severity models, rejecting the null hypothesis that "the resulting models had the same explanatory power in relation to their intercept-only (or null) models."

From Table 3, for the median barrier-related crash model, two indicator variables "*Redirected onto Roadway*" and "*Young Driver*" turned to be random across crash observations. For the roadside

Table 3
MXL model results for crash injury severity of median and roadside barrier-related crashes.

Variable	Median Barrier Crashes*			Roadside Barrier Crashes*		
	Coef.	Std. Err.	*P*-Value	Coef.	Std. Err.	*P*-Value
Constant	−2.113	0.382	0.000	−2.325	0.275	0.000
Crash Characteristics						
Seatbelt Use	−1.019	0.315	0.001	–	–	–
Airbag Deployment	0.726	0.206	0.000	1.000	0.217	0.000
Barrier Characteristics						
Barrier-Hit Outcome:						
Redirected onto roadway	1.035	0.244	0.000	–	–	–
*Redirected onto roadway (SD)***	*0.558*	*0.274*	*0.042*	–	–	–
Penetrated barrier	2.073	0.292	0.000	1.655	0.289	0.000
Median Guardrail Barrier	1.075	0.213	0.000	–	–	–
Median Barrier Position (≤8 feet)	0.470	0.191	0.014	–	–	–
Roadway Characteristics						
Work Zone-Related	−0.959	0.526	0.068	–	–	–
Presence of Roadside Embankment	–	–	–	0.422	0.222	0.058
Driver Characteristics						
Female Driver	0.855	0.193	0.000	0.539	0.205	0.008
Young Driver (≤30 Years)	−0.338	0.268	0.206	−0.537	0.202	0.008
*Young Driver (SD)***	*0.874*	*0.259*	*0.001*	–	–	–
Driver License Issued in Alabama	–	–	–	0.432	0.236	0.068
*Driver License Issued in Alabama (SD)***	–	–	–	*0.477*	*0.197*	*0.015*
Crash Primary Cause:						
Fatigue/Sleepiness	0.759	0.223	0.001	0.658	0.251	0.009
DUI	0.668	0.359	0.063	–	–	–
Aggressiveness	–	–	–	0.843	0.351	0.016
Vehicle Characteristics						
Passenger Car	−0.585	0.192	0.002	–	–	–
Goodness-of-Fit Statistics						
Number of Observations	989			696		
Number of Significant Parameters	15			10		
Log-Likelihood at Zero (LL_0)	−512.082			−393.567		
Log-Likelihood at Convergence (LL)	−422.068			−348.764		
Deviance x^2 Statistic $= -2[LL_0 - LL]$(*P-Value*)	180.028 (*<0.001*)			89.61 (*<0.001*)		
AIC (Akaike Information Criterion)	874.14			717.53		
BIC (Bayesian Information Criterion)	947.59			762.98		

* *"No injury" used as the reference category & parameters at 10% significance level were retained.*
** *SD means standard deviation of estimated parameter.*

barrier-related crash model, "*Driver License Issued in Alabama*" was the only variable found to have a random effect on the severity outcome.

The effects of each variable on the severity of barrier crashes are discussed in the following subsections. To explore the actual effects of the variables on the probability of injury/fatal crash severity outcomes, the marginal effects were estimated and are provided in Table 4. It should be noted that, in addition to investigating the main effects of variables of interest, several interaction effects among the independent variables were examined (e.g., the interaction effect of speed limit groups and airbag deployment in the crash injury severity model). However, no interaction effects were found to be significant in the final models. As a result, only main effects of the independent variables were included and discussed.

5.1.1. Crash-specific characteristics

From Table 3, seatbelt use was found to be correlated with a reduced probability of injury severity in median barrier crashes, which is expected. Wearing a seatbelt protects vehicle occupants from serious injuries when a crash occurs. According to the estimated marginal effect, wearing a seatbelt reduced the risk of sustaining fatal or injury severity outcomes by 15.6% in median barrier crashes. Airbag deployment was associated with an increased probability of injury or fatal outcomes in both median and roadside barrier crashes. Airbags are usually activated in the event of high-

Table 4
Marginal effects results for crash injury severity of median and roadside barrier-related crashes.

Variable	Median Barrier Crashes % Change in Injury/ Fatal	Roadside Barrier Crashes % Change in Injury/ Fatal
Seatbelt Use	−15.6	
Airbag Deployment	10.4	18.1
Barrier-Hit Outcome:		
Redirected onto roadway	14.6	–
Penetrated barrier	35.0	33.0
Median Guardrail Barrier	15.0	–
Median Barrier Position (≤8 feet)	6.3	–
Work Zone-Related	−10.7	–
Presence of Roadside Embankment	–	6.5
Female Driver	11.6	8.9
Young driver (≤30 Years)	−2.7	−8.4
Driver License Issued in Alabama	–	7.5
Crash Primary Cause:		
Fatigue/Sleepiness	10.9	11.5
DUI	9.8	–
Aggressiveness	–	15.2
Passenger Car	−7.9	–

speed related crashes, which could result in serious injuries to the vehicle occupant. Other studies found similar results (Amarasingha & Dissanayake, 2014; Behnood & Mannering, 2017; Khorashadi et al., 2005; Savolainen & Ghosh, 2008; Schneider et al., 2009). The marginal effects indicate that airbag deployment increased the probability of injury consequences by 10.4% and 18.1% for median and roadside barrier-related crashes, respectively.

5.1.2. Barrier-specific characteristics

Regarding the effect of barrier-hit outcomes, the results indicated that both the category "penetrated the barrier" was positively associated with an increased probability of injury or fatal barrier-related crashes, while "redirected onto the roadway" contributed to the severity of median-barrier crashes. These findings are expected because when a vehicle penetrates the barrier or is redirected onto its travel lane after striking the barrier, the risk of severe injury outcomes increases, as compared to the situation where the vehicle is contained by the barrier. According to the marginal effects (in Table 4), "penetrated the barrier," on average, increased the probability of killed or injury outcomes in median and roadside barrier crashes by 34%. For the median barrier type, the effect of vehicle redirection onto the roadway was found to be random and normally distributed with a mean of 1.035 and standard deviation of 0.558, indicating that over 97% of the distribution is above zero. This means that vehicle redirection onto the roadway increased the probability of injury or fatal outcomes for over 97% of the median barrier crashes and reduced the likelihood of fatalities or injuries for 3% of the sample.

For median barrier crashes, the probability of injury or fatal outcomes is higher for guardrails compared to high-tension cable barriers, which is an expected result. This may be due to the fact that cable barriers are more flexible than guardrails, and hence absorb more impact energy than guardrails when being struck by an errant vehicle and could result in less injury outcome. This finding is consistent with that of previous studies (Alberson et al., 2003; Hu & Donnell, 2010; Hunter et al., 2001; Zou & Tarko, 2018). Based on the marginal effect, a collision with median guardrails had a 15% increase in the probability of sustaining fatal or injury outcomes.

The result for median barrier lateral position shows that a nearside median barrier (where the distance between the barrier and edge of the nearest travel lane is less than or equal to 8 feet) was associated with an increased probability of injury or fatal outcomes in median barrier crashes. This might be due to the fact that farside median barriers (distance >8 ft) could provide drivers with more recovery room, so that they can reduce their speed or regain control of the vehicle in the case of striking a median barrier; hence, lowering the risk of fatal or injury outcomes. This finding is consistent with that of previous studies (Hu & Donnell, 2010; Zou, Tarko, Chen, & Romero, 2014). From the marginal effect estimated, there was a 6.3% increase in the probability of fatal or injury outcomes for a nearside median barrier (distance to the nearest travel lane $\leq$8 ft).

5.1.3. Roadway characteristics

From Table 3, regarding road work zones, the result showed that median barrier crashes that occurred at work zones were associated with a reduced probability of injury or fatal severity outcomes. This finding could be explained by the fact that drivers tend to travel more cautiously nearby road work and construction zones, resulting in lower vehicle impact speeds in median barrier-related crashes. The corresponding marginal effect shows that work zones reduced the probability of sustaining injury outcomes by 10.7%.

The presence of roadside embankment was positively associated with the severity of roadside barrier crashes. Roadside embankments increase the risk of rollover crashes when an errant vehicle runs off the roadway, which mainly results in severe injuries. This finding is supported by previous studies (Anarkooli et al., 2017; Martin et al., 2013; Zou, Tarko, Chen, & Romero, 2014). From the marginal effect, the presence of roadside embankments increased the probability of fatal or injury outcomes by 6.5%.

5.1.4. Driver characteristics

Female drivers were found to be more involved in injury or fatal crashes in both median and roadside barrier collisions. In general, females have lesser physiological strength compared to males. As a result, they are more likely to sustain severe injuries in traffic crashes. Yasmin et al. (2004), for example, found that female drivers increased the probability of severe injury crashes involving guardrails. A similar finding was also reported in other studies (Li et al., 2019; Molan et al., 2019; Schneider et al., 2009; Wu et al., 2016; Yasmin et al., 2014; Zou, Tarko, Chen, & Romero, 2014). The marginal effect indicates that female drivers increased the probabilities of fatal or injury outcomes by 11.6% and 8.9% for median and roadside barrier crashes, respectively.

The indicator for young drivers (30 years or under) was negatively associated with injury severity in median and roadside barrier crashes. In median barrier crashes, the young driver indicator was found to be random (mean of -0.338 and standard deviation of 0.874). This implies that for about 65% of the median barrier crashes involving young drivers, hitting a median barrier is less likely to result in injury or fatal outcomes. One possible reason is that young drivers possess fitter physical conditions than middle and elderly age groups, which enable them to avoid severe injuries. This result was also supported by Liu and Fan (2019) and Rezapour et al. (2019). From the marginal effect, there were, respectively, 2.7% and 8.4% reductions in fatal or injury probability for median and roadside barrier crashes involving young drivers.

The result for driving license status showed that in-state drivers (with an Alabama-issued driving license) were more likely to be involved in severe roadside barrier crashes. The effect of this factor was random (mean of 0.432 and standard deviation of 0.477). These values suggested that for 82% of roadside barrier-related crashes involving drivers with an Alabama-issued driving license, the probability of being killed or injured is higher. A potential explanation is that native or in-state drivers are more familiar with the road environment; hence, they tend to drive more aggressively and at relatively higher speeds than out-of-state drivers (who tend to drive more cautiously due to their unfamiliarity of the roadway conditions). The marginal effect indicated that Alabama native drivers increased the risk of injury-resulting roadside barrier crashes by 7.5%.

Driver fatigue/sleepiness was found to be as a primary factor affecting the injury severity of median and roadside barrier crashes. The marginal effects showed that being fatigued or falling asleep increased the probability of injury severities, on average, by 11.2% in median and roadside barrier-related crashes. Driving under the influence (DUI) of alcohol or drugs was another driver factor that affected injury severity in median barrier crashes. This finding was expected as DUI drivers tend to drive aggressively and they are more often unable to react in time to avoid barrier-related crash. Based on the marginal effects, DUI drivers experienced a 9.8% increase in the risk of severe injuries in median-related barrier crashes.

Aggressive driving (or aggressiveness) was found to increase the probability of injury outcomes in roadside barrier crashes. This finding is not surprising as driving aggressive behaviors (such as speeding or following too closely) significantly contribute to severe injury consequences. A similar finding was found by Zou and Tarko (2018). Based on the marginal effect, a roadside barrier collision involving aggressive driving was associated with a 15.2% increase in the probability of fatal or injury outcomes.

5.1.5. Vehicle characteristics

Compared to other vehicle types, passenger cars were less likely to be involved in injury or fatal outcomes in median-related barrier crashes. One possible explanation is that passenger cars have lower mass compared to other heavy vehicles. As such, they are more likely to be contained by the barrier in the event of a median barrier crash, lowering the risk of an injury or fatal outcome (Russo & Savolainen, 2018). From the marginal effect, passenger cars reduced the probability of injury outcomes by 7.9% in median barrier crashes.

5.2. Analysis of Barrier-Hit outcome

This section describes the results of MXL models estimating the probabilities of different barrier-hit outcomes for median and roadside barrier crashes. As previously illustrated, there were three different barrier-hit outcomes (vehicle containment, vehicle redirection onto the roadway, and barrier penetration). The first two types denote barrier success (or non-crossover consequence), whereas the latter denotes barrier failure (i.e., a crossover). Note that "barrier penetration" served as the reference case. It should be noted that information regarding barrier-hit outcomes was not available for 29 crash records. As a result, the total number of observations for barrier-hit outcomes was reduced to 1,656 cases (as opposed to 1,685). Also, as previously illustrated in the injury severity model, the interaction effects of several independent variables were examined in the barrier-hit outcome model (e.g., the interaction effect of passenger cars and speed limit groups). However, no interaction effects were found to be significant in the final models.

The modeling results are shown in Table 5 and the corresponding marginal effects are presented in Table 6. Note that two coefficients (for vehicle containment and vehicle redirection after hitting the barrier) are shown for each parameter since the response variable had three levels and one level was used as the reference case (barrier penetration). The deviance statistic values (that follow the Chi-square distribution) for the two models were found to be statistically significant, which reject the null hypothesis that "the resulting models had the same performance as their intercept-only (or null) models, implying an overall good statistical fit for the fitted MXL models."

Table 5
MXL model results for barrier-hit outcomes for median and roadside barriers.

Variable	Median Barrier Crashes*						Roadside Barrier Crashes*					
	Vehicle Contained			Vehicle Redirected			Vehicle Contained			Vehicle Redirected		
	Coef.	Std. Err.	P-Value	Coef.	Std. Err.	P-Value	Coef.	Std. Err.	P-Value	Coef.	Std. Err.	P-Value
Constant	0.888	0.698	0.203	−0.943	0.67	0.159	4.964	2.292	0.030	1.280	2.262	0.571
Barrier Characteristics												
Brifen Cable Barrier	0.685	0.923	0.458	−1.112	0.586	0.058	–	–	–	–	–	–
Long Barrier Length (>1230 feet or 0.2 miles)	2.123	0.467	0.000	1.118	0.375	0.003	–	–	–	–	–	–
*Long Barrier Length (>1230 feet) (SD)***	*3.475*	*1.571*	*0.027*	–	–	–	–	–	–	–	–	–
Median Barrier Position (≤8 feet)	−1.010	0.361	0.005	−0.919	0.318	0.004	–	–	–	–	–	–
Roadway Characteristics												
Posted Speed Limit	–	–	–	–	–	–	−0.061	0.033	0.065	−0.012	0.032	0.712
Rural Area Type	−1.941	0.511	0.000	−0.077	0.415	0.852	−0.720	0.383	0.060	−0.095	0.338	0.778
Right Shoulder Width	–	–	–	–	–	–	0.114	0.061	0.063	0.125	0.057	0.029
*Right Shoulder Width (SD)***	–	–	–	–	–	–	*0.054*	*0.015*	*0.000*	–	–	–
Curved Road Alignment	–	–	–	–	–	–	−0.774	0.42	0.066	−0.353	0.387	0.361
Environmental Characteristics												
Winter Season	−0.781	0.42	0.063	0.126	0.358	0.725	–	–	–	–	–	–
Weekend	–	–	–	–	–	–	−0.868	0.303	0.004	−0.586	0.303	0.053
*Weekend (SD)***	–	–	–	–	–	–	–	–	–	*0.551*	*0.257*	*0.032*
Time of Day:												
Daytime (7:00 AM to 6:59 PM)	0.739	0.514	0.151	1.033	0.472	0.029	–	–	–	–	–	–
*Daytime (SD)***				*2.225*	*0.983*	*0.024*	–	–	–	–	–	–
Midnight (12:00 AM to 6:59 AM)	0.262	0.559	0.639	1.001	0.439	0.023	–	–	–	–	–	–
Dark Lighting Condition	–	–	–	–	–	–	−0.596	0.290	0.039	−0.341	0.274	0.213
Driver Characteristics												
Young Driver (≤30 Years)	0.898	0.373	0.016	0.787	0.328	0.017	–	–	–	–	–	–
DUI Crash Cause	–	–	–	–	–	–	−1.138	0.602	0.059	−0.634	0.505	0.209
Vehicle Characteristics												
Vehicle Type:												
Passenger car	1.273	0.512	0.013	2.142	0.469	0.000	1.192	0.325	0.000	1.312	0.311	0.000
SUV/Pick-up	0.902	0.531	0.089	1.426	0.445	0.001	–	–	–	–	–	–
Goodness-of-Fit Statistics												
Number of Observations	978						678					
Number of Parameters	24						20					
Log-Likelihood at Zero (LL_0)	−902.162						−636.748					
Log-Likelihood at Convergence (LL_β)	−824.866						−593.789					
Deviance x^2 *Statistic* = $-2[LL_0 - LL_\beta]$*(P-Value)*	154.592 (*<0.001*)						85.92 (*<0.001*)					
AIC (Akaike Information Criterion)	1697.73						1227.58					
BIC (Bayesian Information Criterion)	1814.98						1317.96					

* *"Barrier penetration" used as the reference category & parameters at 10% significance level were retained.*
** *SD means standard deviation of estimated parameter.*

Table 6
Marginal effects results for barrier-hit outcomes for median and roadside barriers.

Variable	Median Barrier Crashes			Roadside Barrier Crashes		
	Contained*	Redirected*	Penetrated*	Contained**	Redirected*	Penetrated*
Brifen Cable Barrier System	16.5	−17.2	0.7	–	–	–
Long Barrier Length (>1230 ft)	16.2	−13.9	−2.3	–	–	–
Median Barrier Position (≤8 feet)	−2.1	−0.8	2.9	–	–	–
Posted Speed Limit	–	–	–	−11.0	15.4	4.5
Rural Area Type	−19.6	15.3	4.3	−13.4	10.5	2.9
Right Shoulder Width	–	–	–	0.3	0.8	−1.1
Curved Road Alignment	–	–	–	−7.5	2.9	4.5
Winter Season	−9.2	7.9	1.2	–	–	–
Weekend	–	–	–	−7.4	0.7	6.7
Time of Day:						
Daytime	−2.1	6.3	−4.1	–	–	–
Midnight	−5.4	8.3	−2.9	–	–	–
Dark Lighting Condition	–	–	–	−5.1	1.4	3.7
Young driver (<30 Years)	2.0	0.3	−2.4			
DUI Crash Cause	–	–	–	−10.8	5.5	5.2
Vehicle Type:						
Passenger car	−3.7	12.6	−8.8	1.5	9.2	−10.7
SUV/Pick-up	−3.4	6.6	−3.1	–	–	–

* *Denotes % change in each category.*

Among others, four variables were found to be random and had varying effects across the median and roadside crash observations, including barrier length longer than 1230 feet (median-contained), daytime (median-redirected), right shoulder width (roadside-contained), and weekend (roadside-redirected). A variety of variables was found to significantly influence barrier-hit outcomes. The effects of those significant variables are discussed in the following subsections.

5.2.1. Barrier-specific characteristics

From Table 5, for median cable barriers, the results showed that Brifen cable barriers were associated with a lower probability of vehicle redirection after being hit (i.e., higher probability of barrier penetration) when compared to other barrier systems. This result warrants further investigation, as very few studies have been conducted to investigate the effect of different cable barrier types on barrier-hit outcomes. As one of those few studies, the Utah DOT investigated 2-mile-long Brifen and 8-mile-long CASS high-tension systems on I-15 in Ohio in 2003 (Clayton, 2005). The results showed that the Brifen cable barrier system was more likely than the CASS system to be involved in cable barrier penetrations. The marginal effects in Table 6 showed that Brifen cable barrier type increased the barrier penetration probability by 0.7%, but, on the contrary, it reduced the probability of vehicle redirection onto the roadway by 17.2%.

The barrier length indicator (longer than 1230 feet or 0.2 miles) was found to reduce the probability of barrier penetration in median cable barrier collisions. This is because more strike forces are absorbed as the barrier length increases; thus, the probability of barrier penetration is reduced. According to the American Association of State Highway and Transportation Officials (AASHTO) Roadside Design Guide, shorter barrier lengths may not effectively prevent penetration or provide adequate redirection capability (AASHTO, 2011). The long barrier length indicator was random and normally distributed for median contained outcome (mean of 2.123 and standard deviation of 3.475), which indicates that 73% of the median-hit observations had a mean greater than zero. That is, longer barrier lengths were more likely to result in vehicle containment for 73% of the median-hit observations, but less likely for 27%. From the marginal effects in Table 6, longer barrier lengths increased the probability of vehicle containment by 16.2% and reduced the probabilities of the vehicle being redirected onto the roadway and barrier penetration by 13.9% and 2.3%, respectively.

For median barrier crashes, the effect of barrier lateral position less than or equal to 8 feet was significant and had negative coefficients for the outcomes of vehicle containment and redirection onto the roadway. This can be attributed to the fact that as the lateral distance between the safety barrier and edge of the nearest travel lane increases, drivers have more room to avoid a more serious strike, like barrier penetration, with median barrier. The marginal effect showed a 2.9% increase in the probability of barrier penetration by an errant vehicle in the case of median barrier crashes.

5.2.2. Roadway characteristics

The probability of a vehicle being contained or redirected onto the roadway for roadside barrier crashes declined with increasing posted speed limit. An argument for this finding is that higher speed limits are posted in rural areas, where drivers tend to drive at high speeds, and thus they are more likely to lose control of the vehicle and hit a roadside barrier at high speeds resulting in barrier penetration. Interestingly, increasing the speed limit by one mile per hour increased the probability of barrier penetration by 4.5%.

Confirming the "posted speed limit" finding, rural areas had a negative impact on all barrier-hit outcomes for median and roadside categories, indicating that the risk of vehicle penetration through the barrier was higher in rural areas. As previously indicated, drivers tend to travel at high speeds in rural areas, so that the risk of a serious barrier-hit outcome, like barrier penetration, increases when the vehicle leaves the roadway and hits a barrier on the median or roadside. The marginal effects in Table 6 indicate that median and roadside barrier-related crashes that occurred in rural areas increased the probabilities of vehicle redirection onto the roadway and barrier penetration, but reduced the probability of vehicle containment by the barrier.

Wider right shoulders were associated with a higher probability of vehicle being contained or redirected back onto the roadway compared to barrier penetration. This finding is consistent with expectations as wider shoulder widths provide more room for drivers to avoid a barrier hit resulting in serious outcomes, such as barrier penetration. For the vehicle containment outcome of roadside barrier hits, the effect of right shoulder width was found to be random (mean of 0.114 and standard deviation of 0.054). These values indicate that wider right shoulder widths increase the probability of vehicle containment for over 98% of the observations.

Curved freeway sections were more likely, than straight or tangent sections, to experience a barrier penetration outcome in the event of roadside barrier crashes. This might be since drivers are more likely to lose control of their vehicles and run off the roadway when negotiating a horizontal curve; hence, increasing their risk of barrier penetration (especially at high speeds). This finding is consistent with that of Jama et al. (2011). From the marginal effect, the presence of curved sections increased the probability of roadside barrier crashes leading into barrier penetration by 4.5%.

5.2.3. Environmental characteristics

Median barrier hits that occurred in winter were less likely to result in a vehicle being contained. A possible explanation is that winter driving conditions are typically associated with slippery road surface, which increases the risk of RwD crashes involving median barriers. A similar finding was reached in previous studies (Donnell et al., 2002; Sicking et al., 2009). According to the marginal effects, barrier crashes in wintertime were associated with a 9.2% reduction in the probability of vehicle containment and a 1.2% increase in the probability of barrier penetration. Roadside barrier crashes that occurred during the weekend increased the probability of barrier penetration. This result is consistent with previous studies (Jama et al., 2011; Yu et al., 2019). The weekend indicator was found to be random (mean of −0.586 and a standard deviation of 0.551) for the vehicle redirection outcome of roadside barrier hits. These values suggested that 85% of the roadside barrier crashes that occurred on weekends were less likely to result in a vehicle redirection onto the roadway. A possible explanation is the relatively lower traffic volume during weekends, where drivers were more likely to travel at high speeds. In addition, there might be a higher likelihood of DUI driving during the weekend (Roque & Jalayer, 2018). Based on the marginal effects, roadside barrier crashes that occurred on weekends increased the probability of barrier penetration by 6.7%, but reduced the probability of vehicle containment by 7.4%.

Regarding the time of day of crash, daytime (7 a.m. to 6:59 p.m.) and midnight (12.00 a.m. to 6:59 a.m.) periods of the day lead to higher probability of a vehicle being redirected onto the roadway for median-related barrier hits. The estimated parameter for daytime period was found to be random (mean of 1.033 and standard deviation of 2.225). This indicates that the probability that a vehicle being redirected onto the roadway in daytime median barrier crashes was higher for 68% of the observations. The marginal effects suggested that barrier crashes occurring during the daytime or midnight reduced the probabilities of vehicle containment and barrier penetration, but increased the probability of vehicle redirection onto the roadway. Dark lighting condition was associated with a higher probability of a barrier being penetrated in roadside barrier crashes, though it had no significant effect on vehicle redirection. This might be due to the reduced sight distance associated with dark lighting condition. Similar results were reported by Islam and Hernandez (2013), Anarkooli et al. (2017), and Uddin and Huynh (2017). From the marginal effects, roadside barrier crashes occurring in dark lighting condition reduced the probability of vehicle containment by 5.1%, but increased the probability of barrier penetration by 3.7%.

5.2.4. Driver characteristics

Young drivers (≤30 years old) were found to be positively associated with the probability that a vehicle being contained or redirected by the barrier in the event of median barrier crashes. A possible reason for this finding is that young drivers have faster reaction times and more ability than old drivers to regain control of an errant vehicle; hence, avoiding a barrier hit with severe outcomes. The marginal effects suggested that median barrier crashes involving young drivers increased the probability of vehicle containment by 2%, but reduced the probability of barrier penetration by 2.4%. DUI of alcohol or drugs was found to increase the risk of barrier penetration in roadside barrier crashes. This finding is not surprising because DUI drivers tend to drive aggressively and take risky maneuvers; hence, in the case of leaving the roadway, DUI drivers are more likely to be involved in a severe-impact barrier crash.

5.2.5. Vehicle characteristics

Regarding the vehicle type, the results showed that passenger cars were less likely to be involved in a crash resulting in barrier penetration in both median and roadside barrier-related crashes. A similar finding was reached for SUVs/pick-ups involving in median barrier crashes. The reason for these findings is that passenger cars and SUVs are categorized as light vehicles (compared to large truck and truck-trailers); hence, imposing lower impact forces on the barrier in an event of barrier crash. Therefore, such light vehicles were less likely than heavy vehicles (e.g., large trucks and buses), to penetrate the barrier. From the marginal effects, passenger cars reduced, on average, the probability of barrier penetration by about 10% in median and roadside barrier-hit crashes. For SUVs and pick-ups, there was a lesser percentage (3.1%) reduction in the risk of barrier penetration in median barrier-related crashes.

6. Conclusions and recommendations

This study provided a comprehensive investigation of several factors related to barrier-specific, roadway, driver, vehicle, and environmental characteristics that affected the consequences of barrier-related crashes including crash injury severity and barrier-hit outcomes (e.g., vehicle containment, vehicle redirection, and barrier penetration). Two types of safety barriers were investigated in this study, including high-tension cable barriers installed on medians and strong-post guardrails installed on medians and/or roadsides. Separate MXL models were developed for crash injury severity and barrier-hit outcomes for median and roadside barrier-related collisions. The MXL model has the advantage to capture unobserved heterogeneity by allowing a subset of explanatory variables to vary across observations. A total of 1,685 crashes that occurred on three major interstates in the state of Alabama (I-65, I-85, and I-20) between 2010 and 2016 were used.

When modeling the crash injury severity from median and roadside barrier-related crashes, it was found that barrier penetration, female drivers, and being fatigued or falling asleep were associated with a higher probability of injury or fatal crashes. The factors vehicle redirection onto the roadway, median barrier lateral position (>8 feet), median guardrails, and DUI resulted in higher injury or fatal probability outcomes in median barrier-related crashes, while presence of work zones, young drivers, seatbelt use, and passenger cars increased the probability of no-injury crashes. Presence of roadside embankment, drivers with Alabama-issued driving license, and aggressive driving contributed only to the increase in the crash injury severity probability in roadside barrier crashes.

The results of barrier-hit models showed that different factors were associated with the probability of barrier-hit outcomes in median and roadside crashes. Longer barrier length (greater than 1230 feet or 0.2 miles), daytime, right shoulder width, and weekend were heterogeneous across crash observations. For median barrier crashes, the results showed that longer barrier length (greater than 0.2 miles), barrier's far-side lateral position (>8 feet), passenger cars, SUVs and pick-ups, urban areas, daytime, midnight, and young drivers were associated with a lower probability of barrier penetration. For roadside barrier-hit outcomes, barrier-related crashes that occurred on high speed-limit segments, in rural areas, on curved sections, on segments with narrower shoulder widths, on weekends, in dark lighting conditions, due to the influence of alcohol or drugs, and involving vehicles other than passenger cars were correlated with a higher probability of barrier penetration.

The findings obtained in this study could help road safety practitioners improve the design of road barriers with the aim of minimizing the consequences of barrier-related collisions. Two design-related study applications are to design longer barrier run length (greater than 1230 feet or 0.2 miles) to reduce barrier penetration outcomes and to also ensure a minimum 8-feet lateral offset distance between the median barrier and median edge (or edge of the nearest travel lane) to reduce the resulted severity and barrier penetration outcomes. Moreover, the significant impacts of DUI, aggressive driving, fatigued driving, and seatbelt use on the crash injury severity and/or barrier-hit outcomes highlight the importance of launching public awareness campaigns that aim at minimizing the association of these driver-specific factors on the consequences of barrier-related crashes. Another effective solution is to intensify law enforcement targeting DUI driving and aggressive driving behaviors. Because of more forgiving performance of cable barriers, they were found to be less associated with crash injury severity. As such, cable barriers could be installed where the historical number of severe median guardrail hits is high, though requiring wider medians to allow for increased deflection.

Future studies can investigate the impacts of the aforementioned barrier-specific factors on the effectiveness of cable and guardrail barriers using crash observations from other U.S. regions, and then compare the impact of the significant variables with this study. Another research avenue is to compare the results of the flexible and semi-rigid barriers in this study with rigid-type barriers (e.g., concrete barriers) to see how the impact of the significant variables might concur or differ.

Acknowledgments

The authors would like to acknowledge the Alabama Department of Transportation (ALDOT) for the grant provided to conduct this research. The opinions, findings, and conclusions in this paper are those of the authors and not necessarily those of the State of Alabama Department of Transportation. The authors would also like to acknowledge the dedication and effort made by the following research team members in the data collection and police reports review process: Ms. Abigail Swartz, Mr. Israel Anastasio, Mr. Austin Montgomery, and Mr. Matthew Murphy.

References

AASHTO (2011). *Roadside Design Guide*. Washington, D.C.: American Association of State Highways and Transportation Officials.

Al-Bdairi, N. S. S., & Hernandez, S. (2017). An empirical analysis of run-off-road injury severity crashes involving large trucks. *Accident Analysis & Prevention, 102*, 93–100.

Alberson, D. C., Bligh, R. P., Buth, C., & Bullard, D. L. Jr, (2003). Cable and wire rope barrier design considerations. *Transportation Research Record, 1851*(1), 95–104.

Amarasingha, N., & Dissanayake, S. (2014). Gender differences of young drivers on injury severity outcome of highway crashes. *Journal of Safety Research, 49*. 113. e1-120.

Anarkooli, A. J., Hosseinpour, M., & Kardar, A. (2017). Investigation of factors affecting the injury severity of single-vehicle rollover crashes: A random-effects generalized ordered probit model. *Accident Analysis & Prevention, 106*, 399–410.

Anastasopoulos, P. C., & Mannering, F. L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention, 41*(1), 153–159.

Behnood, A., & Mannering, F. (2017). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 16*, 35–47.

Blincoe, L., Miller, T. R., Zaloshnja, E., & Lawrence, B. A. (2015). *The economic and societal impact of motor vehicle crashes, 2010*. Washington, DC: National Highway Traffic Safety Administration.

Burns, K., & Bell, K. (2016). Performance evaluation of a cable median barrier system on an Oregon highway with a narrow median. *Transportation Research Record, 2588*(1), 137–144.

Chimba, D., Ruhazwe, E., Allen, S., & Waters, J. (2017). Digesting the safety effectiveness of cable barrier systems by numbers. *Transportation Research Part A: Policy and Practice, 95*, 227–237.

Chitturi, M. V., Ooms, A. W., Bill, A. R., & Noyce, D. A. (2011). Injury outcomes and costs for cross-median and median barrier crashes. *Journal of Safety Research, 42*(2), 87–92.

Clayton, R. (2005). *Cable barrier: A high tension transformation- the Utah experience*. Salt Lake City: Utah Department of Transportation.

Donnell, E. T., Harwood, D. W., Bauer, K. M., Mason, J., Jr, & Pietrucha, M. T. (2002). Cross-median collisions on Pennsylvania interstates and expressways. *Transportation Research Record, 1784*(1), 91–99.

ESRI, 2019. ArcMap 10.2. http://support.esri.com/en/Products/Desktop/arcgis-desktop/arcmap/10-2-2.

Google Maps, 2018. Satellite Imagery from the Google Maps. https://www.google.com/maps.

Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention, 81*, 14–23.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal, 7*(3), 388–401.

Hosseinpour, M., Sahebi, S., Zamzuri, Z. H., Yahaya, A. S., & Ismail, N. (2018). Predicting crash frequency for multi-vehicle collision types using multivariate Poisson-lognormal spatial model: A comparative analysis. *Accident Analysis & Prevention, 118*, 277–288.

Hosseinpour, M., Yahaya, A. S., & Sadullah, A. F. (2014). Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: Case studies from Malaysian Federal Roads. *Accident Analysis & Prevention, 62*, 209–222.

Hu, W., & Donnell, E. T. (2010). Median barrier crash severity: Some new insights. *Accident Analysis & Prevention, 42*(6), 1697–1704.

Hunter, W. W., Richard Stewart, J., Eccles, K. A., Huang, H. F., Council, F. M., & Harkey, D. L. (2001). Three-strand cable median barrier in North Carolina: In-service evaluation. *Transportation Research Record, 1743*(1), 97–103.

Islam, M. B., & Hernandez, S. (2013). Modeling injury outcomes of crashes involving heavy vehicles on Texas highways. *Transportation Research Record, 2388*(1), 28–36.

Islam, S., & Brown, J. (2017). A comparative injury severity analysis of motorcycle at-fault crashes on rural and urban roadways in Alabama. *Accident Analysis & Prevention, 108*, 163–171.

Jama, H. H., Grzebieta, R. H., Friswell, R., & Mcintosh, A. S. (2011). Characteristics of fatal motorcycle crashes into roadside safety barriers in Australia and New Zealand. *Accident Analysis & Prevention, 43*(3), 652–660.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, F. (2005). Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. *Accident Analysis & Prevention, 37*(5), 910–921.

Kim, J.-K., Ulfarsson, G. F., Kim, S., & Shankar, V. N. (2013). Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accident Analysis & Prevention, 50*, 1073–1081.

Li, Z., Ci, Y., Chen, C., Zhang, G., Wu, Q., Qian, Z. S., ... Ma, D. T. (2019). Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accident Analysis & Prevention, 124*, 219–229.

Liu, P., & Fan, W. (2019). Modeling head-on crash severity on NCDOT freeways: A mixed logit model approach. *Canadian Journal of Civil Engineering, 46*(6), 322–328.

Maistros, A., Schneider Iv, W. H., & Savolainen, P. T. (2014). A comparison of contributing factors between alcohol related single vehicle motorcycle and car crashes. *Journal of Safety Research, 49*(129), e1–135.

Manner, H., & Wünsch-Ziegler, L. (2013). Analyzing the severity of accidents on the German Autobahn. *Accident Analysis & Prevention, 57*, 40–48.

Martin, J.-L., Mintsa-Eya, C., & Goubel, C. (2013). Long-term analysis of the impact of longitudinal barriers on motorway safety. *Accident Analysis & Prevention, 59*, 443–451.

Molan, A. M., Moomen, M., & Ksaibati, K. (2019). Investigating the effect of geometric dimensions of median traffic barriers on crashes: Crash analysis of interstate roads in Wyoming using actual crash datasets. *Journal of Safety Research, 71*, 163–171.

Molan, A. M., Rezapour, M., & Ksaibati, K. (2020). Investigating the relationship between crash severity, traffic barrier type, and vehicle type in crashes involving traffic barrier. *Journal of Traffic and Transportation Engineering (English Edition), 7*(1), 125–136.

Park, J., Abdel-Aty, M., & Lee, J. (2016). Use of empirical and full Bayes before–after approaches to estimate the safety effects of roadside barriers with different crash conditions. *Journal of Safety Research, 58*, 31–40.

Ray, M. H., & Weir, J. A. (2001). Unreported collisions with post-and-beam guardrails in Connecticut, Iowa, and North Carolina. *Transportation Research Record, 1743*(1), 111–119.

Rezapour, M., Wulff, S. S., & Ksaibati, K. (2019). Examination of the severity of two-lane highway traffic barrier crashes using the mixed logit model. *Journal of Safety Research, 70*, 223–232.

Roque, C., & Jalayer, M. (2018). Improving roadside design policies for safety enhancement using hazard-based duration modeling. *Accident Analysis & Prevention, 120*, 165–173.

Russo, B. J., & Savolainen, P. T. (2018). A comparison of freeway median crash frequency, severity, and barrier strike outcomes by median barrier type. *Accident Analysis & Prevention, 117*, 216–224.

Sacchi, E., Persaud, B., & Bassani, M. (2012). Assessing international transferability of highway safety manual crash prediction algorithm and its components. *Transportation Research Record, 2279*(1), 90–98.
Saleem, T., & Persaud, B. (2017). Another look at the safety effects of horizontal curvature on rural two-lane highways. *Accident Analysis & Prevention, 106*, 149–159.
Savolainen, P., & Ghosh, I. (2008). Examination of factors affecting driver injury severity in Michigan's single-vehicle–deer crashes. *Transportation Research Record, 2078*(1), 17–25.
Schneider, W. H., Savolainen, P. T., & Zimmerman, K. (2009). Driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways. *Transportation Research Record, 2102*(1), 85–92.
Sicking, D. L., De Albuquerque, F. D. B., Lechtenberg, K. A., & Stolle, C. S. (2009). Guidelines for implementation of cable median barrier. *Transportation Research Record, 2120*(1), 82–90.
Uddin, M., & Huynh, N. (2017). Truck-involved crashes injury severity analysis for different lighting conditions on rural and urban roadways. *Accident Analysis & Prevention, 108*, 44–55.
Ulfarsson, G. F., & Mannering, F. L. (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention, 36*(2), 135–147.
Wang, C., Quddus, M. A., & Ison, S. G. (2011). Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. *Accident Analysis & Prevention, 43*(6), 1979–1990.
Wu, Q., Chen, F., Zhang, G., Liu, X. C., Wang, H., & Bogus, S. M. (2014). Mixed logit model-based driver injury severity investigations in single-and multi-vehicle crashes on rural two-lane highways. *Accident Analysis & Prevention, 72*, 105–115.
Wu, Q., Zhang, G., Chen, C., Tarefder, R., Wang, H., & Wei, H. (2016). Heterogeneous impacts of gender-interpreted contributing factors on driver injury severities in single-vehicle rollover crashes. *Accident Analysis & Prevention, 94*, 28–34.
Yasmin, S., Eluru, N., Pinjari, A. R., & Tay, R. (2014). Examining driver injury severity in two vehicle crashes–A copula based approach. *Accident Analysis & Prevention, 66*, 120–135.
Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science, 63*, 50–56.
Yu, R., Wang, Y., Quddus, M., & Li, J. (2019). A marginalized random effects hurdle negative binomial model for analyzing refined-scale crash frequency data. *Analytic Methods in Accident Research, 22* 100092.
Zou, Y., & Tarko, A. P. (2016). An insight into the performance of road barriers–redistribution of barrier-relevant crashes. *Accident Analysis & Prevention, 96*, 152–161.
Zou, Y., & Tarko, A. P. (2018). Barrier-relevant crash modification factors and average costs of crashes on arterial roads in Indiana. *Accident Analysis & Prevention, 111*, 71–85.
Zou, Y., Tarko, A. P., Chen, E., & Romero, M. A. (2014). Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury. *Accident Analysis & Prevention, 72*, 55–65.

Dr. Mehdi Hosseinpour is a Postdoctoral Research Fellow at Western Kentucky University. His main interests are traffic crash analysis, road user safety behavior, and the application of state-of-the-art statistical models to transportation studies.

Dr. Kirolos Haleem is an Assistant Professor of Transportation Engineering at Western Kentucky University. He has more than 10-year experience in transportation engineering. His areas of expertise include traffic safety analysis, application of statistical models and data mining techniques in transportation data analysis, traffic operations, and ITS.

Journal of Safety Research 78 (2021) 80–95

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Exploring relationships between microscopic kinetic parameters of tires under normal driving conditions, road characteristics and accident types

Changjian Zhang, Jie He *, Xintong Yan, Ziyang Liu, Yikai Chen, Hao Zhang

[a] School of Transportation, Southeast University, Nanjing 210018, China
[b] School of Automobile and Traffic Engineering, Hefei University of Technology, Hefei 230009, China

ARTICLE INFO

Article history:
Received 10 August 2020
Received in revised form 8 December 2020
Accepted 27 May 2021
Available online 9 June 2021

Keywords:
Microscopic kinetic parameters of tires
Normal driving conditions
SOL model
FPL model
RPL model
Mechanism of traffic accidents

ABSTRACT

Introduction: Freeway accidents are a leading cause of death in China, which also triggers substantial economic loss and an emotional burden to society. However, the internal mechanism of how microscopic kinetic parameters of vehicles influenced by road characteristics determine the occurrence of different types of accidents has not been explicitly studied. This research aimed to explore the "link role" of tire microscopic kinetic parameters in road characteristic variables and traffic accidents to aid in facilitating the traffic design and management, and thus to prevent traffic accident. *Method:* A mountain freeway in Zhejiang Province, China was used as the research object and the data used in this paper were obtained through a real-time vehicle experiment. Multiple estimation models, including the standard ordered logit (SOL) model, fixed parameters logit (FPL) model, and random parameters logit (RPL) model were established. *Results:* The findings show that road characteristics will affect the longitudinal kinetic characteristics of the vehicle and, consequently, map the level of risk of rear-end accidents. Driving compensation effects were also identified in this paper (i.e., the drivers tend to be more cautious in complicated driving circumstances). Another finding relating to the mountain freeway is that different tunnel characteristics (e.g., tunnel entrance and tunnel exit) have different effects on different types of traffic accidents. *Practical Applications:* The framework proposed in this article can provide new insight for researchers to enlarge the research subjects of both explanatory and outcome variables in accident analysis. Future research could be implemented to consider more driving conditions.

1. Introduction

Freeway crashes are recognized as a leading causes of death in China, which also triggers substantial economic loss and an emotional burden to society. According to data from the National Bureau of Statistics, more than 240,000 traffic accidents occur each year in China, with about 60,000 people killed, ranking first in the world for many consecutive years. With the advantages of a large capacity and low delay, the freeway has developed into a major artery of road transportation in China. However, due to the high traffic volume and high speed, the freeway, which only accounts for 3% of the total mileage, bears 10% of the death toll (Liu, He, Zhang, Xing, & Zhou, 2020). This phenomenon is particularly evident on mountain highways, with complex and changeable road features being a latent precursor to increase the occurrence and frequency of severe crashes. Under the aforementioned background, much analyses has been conducted to investigate the causes of freeway traffic accidents (Dadashova, Ramírez, McWilliams, & Izquierdo, 2016; Ellison, Greaves, & Bliemer, 2015). Past studies have identified the diversity of the causes and conditions of different types of traffic accidents, which deserves more exploration for a better understanding of the inherent mechanism of various accidents and, thus, to provide the targeted strategies and countermeasures of traffic safety management and control.

Ye, Pendyala, Washington, Konduri, and Oh (2009) proposed that traffic accidents occurring at specific locations tend to be due to the joint impact of the complex roadway geometric characteristics and traffic flow conditions. The study also demonstrated that location-related accidents may not occur frequently, whereas some specific location-related accidents have a propensity to share a preponderate proportion, among which mountain freeway accidents have been considered as significant but not adequately studied subjects. Likewise, empirical evidence has proven that there is a critical necessity to further investigate the potential factors that

* Corresponding author.
E-mail addresses: 230189680@seu.edu.cn (C. Zhang), hejie@seu.edu.cn (J. He), 230198699@seu.edu.cn (X. Yan), lzyseu6257@outlook.com (Z. Liu), yikaichen@hfut.edu.cn (Y. Chen), andyhao@seu.edu.cn (H. Zhang).

https://doi.org/10.1016/j.jsr.2021.05.010
0022-4375/

induce the frequency of mountain freeway accidents, given the complicated mountain freeway circumstances of varied topographies and intricate road facilities such as tunnels, ramps, and interchange.

1.1. Accident analysis and in-vehicle data research

Traffic accidents are classified according to injury-severity in the existing accident analysis studies, which have focused on the investigation of how the variation in the numerous indicator variables such as human factors, vehicle factors, and environmental factors can be determinant in the injury-severity of accidents (Mannering & Bhat, 2014; Mothafer, Yamamoto, & Shankar, 2016). It should be illustrated that accident type such as rear-end accidents, fixed-object collision accidents, and overturning accidents are often identified as indicator variables instead of outcome variables in the previous studies, whereas further research could attempt to use the accident type as an outcome variable in accident analysis to reveal the internal mechanism of the occurrence of different accident types for better prevention of these accidents. Additionally, past work has been restricted to the analyses of the macroscopic and after-event accident-related factors such as weather, lighting condition, and road surface condition, neglecting the microscopic kinetic parameters of real-time tire output when the vehicle is driving normally.

Ye et al. (2009) found that the traffic flow on the main road has a significant impact on rear-end accidents, and the larger shoulder width was more likely to cause side collision and skidding. Park and Lord (2007) concluded that the brightness is negatively correlated with fatal accidents, and as the severity of the accident decreased, the correlation gradually turned into a positive correlation and increased. De Oña, Mujalli, and Calvo (2011) took the Spanish freeway as the research object, studied the causes of accidents of different severity from the aspects of accident characteristics, freeway characteristics, vehicle characteristics, driver characteristics, and environmental factors, and found the characteristics of the accident, driver age, light conditions, and number of injuries are highly correlated with severe accidents. Li, Wang, Liu, Bigham, and Ragland (2013) studied the determinant influence of country-level in fatal accidents and found that an increasing proportion of freeways in the region is correlated with fewer fatal accidents, and a larger proportion of drivers under the age of 18 tend to cause more fatal accidents. Beshah and Hill (2010) focused on research on the impact of road conditions on the severity of accidents with data mining technology.

In recent years, the method of using in-vehicle data for accident analysis is developing rapidly, and the most representative one is Naturalistic Driving Studies (NDS), which has been verified by the Virginia Tech Transportation Institute (VTTI) and US Strategic Highway Research Program Phase 2 (SHRP 2) to be effectively applied to road traffic safety analysis (Hao, Eric, Medina, Gibbons, & Wang, 2020; Seacrist et al., 2020; Simons-Morton et al., 2020). Naturalistic driving data (such as the vehicle's driving trajectory, speed, heading angle, and road infrastructure characteristics) over a period of time were captured with data collection equipment installed on volunteers' vehicles, and some scholars have utilized the relevant data relating to certain road accidents to quantitatively analyze the location of high accident risk in a road or a road network (Dingus & Klauer, 2006; Ghasemzadeh & Ahmed, 2018; Molnar, Eby, Bogard, LeBlanc, & Zakrajsek, 2018). Benefiting from the characteristics of easy access, large amount, and high quality in-vehicle data analysis approach, researchers have been more and more inclined to use these data in various fields of traffic safety, such as: driver factors (Molnar et al., 2018), features of road facilities (Gitelman et al., 2018), surrogate events (Wu & Jovanis, 2013), the relationship between regional economic level, and vehicle speeding (Ghasemzadeh & Ahmed, 2019). Although the safety research on in-vehicle data has been continuously expanded, the related studies involving in-vehicle data to explore the causes of different types of traffic accidents are still insufficient.

1.2. Research on the association between in-vehicle data and other factors

The generation of real-time in-vehicle data when driving under the naturalistic environment is the result of the joint effects of various factors such as road conditions, driver behavior, and traffic flow environment. Therefore, uncovering the inherent correlation between in-vehicle data and other factors will facilitate better comprehension of the internal mechanism of accidents.

Garber and Gadiraju (1989) proposed that changes in speed have a direct impact on traffic safety, and on a well-shaped road, speeding tends to occur more easily, but it does not cause a higher accident rate. Speed variance was found to have a significant positive correlation with accident rate in this study. Mahmud, Ferreira, Hoque, and Tavassoli (2019) took two-lane and two-way freeways as the research object to explore the relationship of vehicle speed, road geometry characteristics, and roadside environment. Their findings provided a more detailed insight to better understand driver's speed tendency in different road segments. Medina and Tarko (2005) developed free-flow speed models based on the analysis of speed distribution influenced by road characteristics of tangent segments and horizontal curves of two-lane freeways in rural areas. Gitelman et al. (2018) employed the in-vehicle data record (IVDR) system to obtain the critical events occurring in the process of naturalistic driving, such as braking, steering braking, and speed warning and constructed a negative binomial regression model to explore the relationship between the occurrence of these events and the operation of road infrastructure characteristics. It was found that the increase of speed warning and decrease of emergency braking could occur under a better road condition, while the opposite effect could be observed when the constraints of the road increased or the vehicle was approaching the intersection.

It can be concluded that the correlation between some macroscopic vehicle driving data, such as speed, braking, and other road characteristics has been fully identified by previous studies, whereas the microscopic kinetic parameter response of the tire of the vehicle still lacks explicit unraveling, which can also serve as a powerful instrument in the field of traffic safety. Compared with their macroscopic factor counterparts, the microscopic kinetic parameters may reflect the real-time driving state of the vehicle from a more intuitive perspective and subsequently help to provide a statistical basis for the establishment of a more comprehensive real-time accident risk prevention system.

1.3. Modeling approach

The accident estimation models can be categorized into two aspects based on different modelling purposes: accident frequency prediction and accident cause analysis. Regarding quantitative analysis of the effects of different levels of factors on the frequency of accident types, univariate Poisson regression models are widely used (Geedipally & Lord, 2010; Lyon, Oh, Persaud, Washington, & Bared, 2003; Milton, Shankar, & Mannering, 2008). However, one limitation should be emphasized that the univariate models tend to ignore the correlation between different types of accidents, resulting in deviations in the independently estimated results (Park & Lord, 2007). The multivariate regression models, compared to the univariate models, exhibit better performance in simultaneously dealing with multiple indicator variables and, thus, significantly ameliorate the prediction accuracy, whereas the complexity of the models may increase the difficulty of model estimation.

Logit models are generally utilized to conduct the analysis of the contribution of different factors to accidents. Abdel-Aty and Pemmanaboina (2006) constructed a log-linear regression model to fit the impact of real-time traffic flow and weather on accidents. Tay, Choi, Kattan, and Khan (2011) established a multinomial logit (MNL) model to investigate the main factors that affected the severity of pedestrian traffic accidents. Abrari et al. (2020) constructed a MNL model to study the motorcycle crash severity at Australian intersections. Chen et al. (2015) explored the remarkable causes of rear-end accidents based on driver behavior, vehicle factors, natural environment, and road geometry parameters through the MNL model. Fan, Kane, and Haile (2015) divided the severity of accidents and employed an MNL model to explore the effect of explanatory variables on traffic accidents of different levels of injury-severity. Fixed parameters logit (FPL) models were generally estimated in numerous previous studies, while they were found to be incapable of observing the heterogeneity in the data, leading to biased estimation results of the coefficients and inferences of the models. The random parameters logit (RPL) models/ mixed logit models were subsequently proposed as a useful and effective instrument to better identify the unobserved heterogeneity and gradually became a research hotspot in recent studies (Dong, Ma, Chen, & Chen, 2018; Haleem & Gan, 2013). Yu and Abdel-Aty (2014) compared the fit effect of FPL models and RPL models using ROC and found that the performance of RPL models was better. Xing, He, Abdel-aty, Wu, and Yuan (2020) constructed a RPL model considering the effects of time and space and analyzed the factors that caused traffic conflicts in the upstream area of the toll station. Li et al. (2019), Li et al. (2019) and Behnood and Mannering (2015) used a mixed logit model in the accident study to explain the heterogeneity of unobservable variables.

According to the research results of the existing literature, the macroscopic factors, such as traffic flow, weather, road conditions, driver characteristics, and vehicle operating parameters (speed, acceleration/deceleration, heading, etc.) have been extensively discussed, while there is a lack of research on the microscopic force characteristics of tires when the vehicle is driving under a normal condition. This paper aims to explore the "link role" of tire microscopic kinetic parameters in road characteristic variables and traffic accidents. The authors are not aware of any previous research similar to this study. Using a mountain freeway in Zhejiang Province, China as the research object, by installing the six-component force meter on the vehicle, the real-time microscopic kinetic parameter response generated by the tire was obtained, and then the microscopic kinetic parameters were used as the indicator variables to build the standard ordered logit (SOL) model, FPL model, and RPL model for the exploration of the quantitative relationship among microscopic kinetic parameters, road characteristics, and accident types.

To aid in the real-time traffic accident analysis and prevention, the study will reveal different road characteristics and different types of traffic accidents corresponding to the changes in the microscopic kinetic parameters. The model estimation can be employed to intuitively analyze the probability of different road characteristic variables to cause abnormal fluctuations in the microscopic kinetic parameters as well as the type of traffic accidents that tend to occur.

2. Data

The accident statistics and road characteristic data were provided by the Zhejiang Provincial Department of Transportation (relying on the science and technology project "Application Research on Rapid Identification of Highway Safety Defects." Related studies have been successfully published based on these data (Li & He, 2016; Liu et al., 2020).

2.1. Accident data

Wenli Freeway is located in Zhejiang Province, China, also known as "bridge and tunnel club" because up to 90% of its length is tunnels and bridges; there are also numerous mountains and complicated geographical environment along the route. The accident database used in this paper contains the accident data of K117-K189 segments (two directions: Wenzhou direction and Lishui direction) of Wenli Freeway from 2007 to 2013, including the following two types of statistics: (1) the location of each accident; (2) the type of each accident; The types of accidents include rear-end, overturning, fixed-object hitting, and so forth.

The total number of accidents was 2,026, including 670 rear-end accidents, 1,087 fixed-object hitting accidents, 200 overturning accidents, and 69 other types of accidents (e.g., scratch, fire). From the perspective of accident statistics, rear-end, fixed-object hitting, and overturning accidents are the three most common types of accidents. Hence, this study will focus on an in-depth analysis of these three types of accidents. The distribution of accident types are shown in Fig. 1.

2.2. Highway characteristic data

Road characteristic data for the K117-K189 segments of the Wenli Freeway were obtained from the Zhejiang Transportation Administration. The road characteristics database mainly includes: curve radius, length of transition curve, elevation difference, elevation standard deviation (STD), tunnel, riverside bridge, conflict zone (diversion/merge zone), etc. (see Table 1). Referring to Liu et al. (2020) for the division method and results, the left and right lines of K117 $\sim$ K189 were divided into 144 road segments with 1 km as the unit ($i = 1, 2, \cdots 144$).

Table 2 provides the mean, STD, minimum, and maximum values of the road characteristic variables. Among them, "*geometry*" is a variable describing road geometry alignment. When the values 0, 1, and 2, respectively, represent that the road segment is a straight one, a circular curve one, and a composite alignment one (circular curve + straight), it is recorded as "*straight*," "*circular curve*," and "*composite alignment*."

To analyze the impact of the tunnel on the driving vehicle, the "*tunnel*" is innovatively decomposed into five categorical variables: when there is no tunnel in a road segment, the value is 0 ("*no tunnel*"); when there is only tunnel exit, the value is 1 ("*tunnel exit*"); when there is only tunnel entrance, the value is 2 ("*tunnel entrance*"); the value is 3 when there is both the exit and entrance of the tunnel ("*both exit and entrance*"); and the value is 4 when a certain segment is completely inside the tunnel (*inside the tunnel*").

When a bridge along the river is built on a certain road segment, the value of "*bridge along the river*" is 1, otherwise it is 0.

Some road segments have some areas called "*conflict zone,*" including interchange hubs, service areas, and other areas where there is a separation and convergence of traffic flows. The discrete transform approach was also used to describe whether there are diverging and merging zones. When there is no diversion/merging zone in a road segment, the value is 0 when there is only existing diversion zone, the value is 1 ("*diversion zone*"); when there is only existing merging zone, the value is 2 ("*merge zone*"); when there is existing both diversion and merge zone at the same time, the value is taken as 3 ("*diversion and merge*").

Fig 1. Different type of accidents distribution on Wenli Freeway.

Table 1
Road characteristic statistics.

Segment	Elevation change (m)	Elevation STD	Horizontal radius (m)	Length of transition curve (m)	Transition curve ratio	Tunnel	Bridge along the river	Conflict zone
1	15.69	15.96	12,000	0	0	Y	N	N
2	−25.63	12.86	2000	240	0.120	Y	N	N
3	11.66	8.56	2000	0	0	Y	N	N
...	...	...	...	...	...	...	...	...
142	2.52	1.91	2100	250	0.119	N	Y	N
143	1.58	2.07	12,000	0	0	N	Y	N
144	7.25	3.88	1250	170	0.136	N	Y	N

Table 2
Descriptive statistics.

Variable	Obs	Mean	Std.Dev.	Min	Max
Elevation_change	144	0	22.525	−78.6	68.81
Elevation_STD	144	7.398	7.455	0	28.07
Geometry	144	0.917	0.642	0	2
Horizontal radius (m)	144	4422.361	4649.364	500	12,000
Length of transition curve (m)	144	85.42	87.721	0	250
Tunnel	144	1.076	1.49	0	4
Bridge along the river	144	0.431	0.497	0	1
Conflict zone	144	0.208	0.699	0	2

2.3. Tire kinetic data

Fig. 2 shows the equipment used to obtain kinetic data: a six-component force meter. A vehicle loaded with a six-component force meter (installed on the tire) was employed to conduct a real vehicle experiment. After 3-round driving on the target freeway segments, the microscopic kinetic parameters of the tire traversing 144 road segments at a frequency of 100HZ were collected. The experimental conditions are shown in Table 3. The obtained microscopic kinetic parameters include: longitudinal force "*Fx*," lateral force "*Fy*," vertical force "*Fz*," turning moment "*Tx*," rolling moment "*Ty*," homing moment "*Tz*," and vehicle speed "*V*." In order to avoid the influence of traffic flow status on experimental results, the experiment selected non-holiday time with free traffic flow. Table 4 displays some of the original data obtained from the experiment.

An experimental HONDA CR-V car and automotive instrumentation (six-component force meter) were used for the experimental study. Automotive instrumentation consisted of a 6-component force axle load sensor SLW-ND, a wheel alignment measuring system, a wireless signal receiver DT-24R, a 6-component force measuring analyzer MFT-306T, control software WAM-701A, a multifunction car recorder TMR-200, and a laptop PC. The wheel alignment measuring system consisted of wheel alignment sensor WAD-1A and special measuring analyzer WAM-1A installed together with the SLW-ND in the right front wheel of the HONDA CR-V car, as shown in Fig. 2.

Lateral tire forces were measured by the SLW-ND and then were transferred to the MFT-306T by the DT-24R. The WAD-1A ensured the signal acquisition precision of the SLW-ND, while the SLW-ND ensured the signal processing precision for the WAM-1A.

In this paper, the microscopic kinetic parameters of vehicles were used as intermediate variables to explore the joint mechanism of road characteristics and traffic accidents. Therefore, in the experimental design, strict control of the driver characteristics was carried out to weaken the interference of the driver factors on the research results. The driver, with 15-year driving experience, was explicitly informed of the purpose of this experiment and had sufficient rest before driving. There was no behavior such as

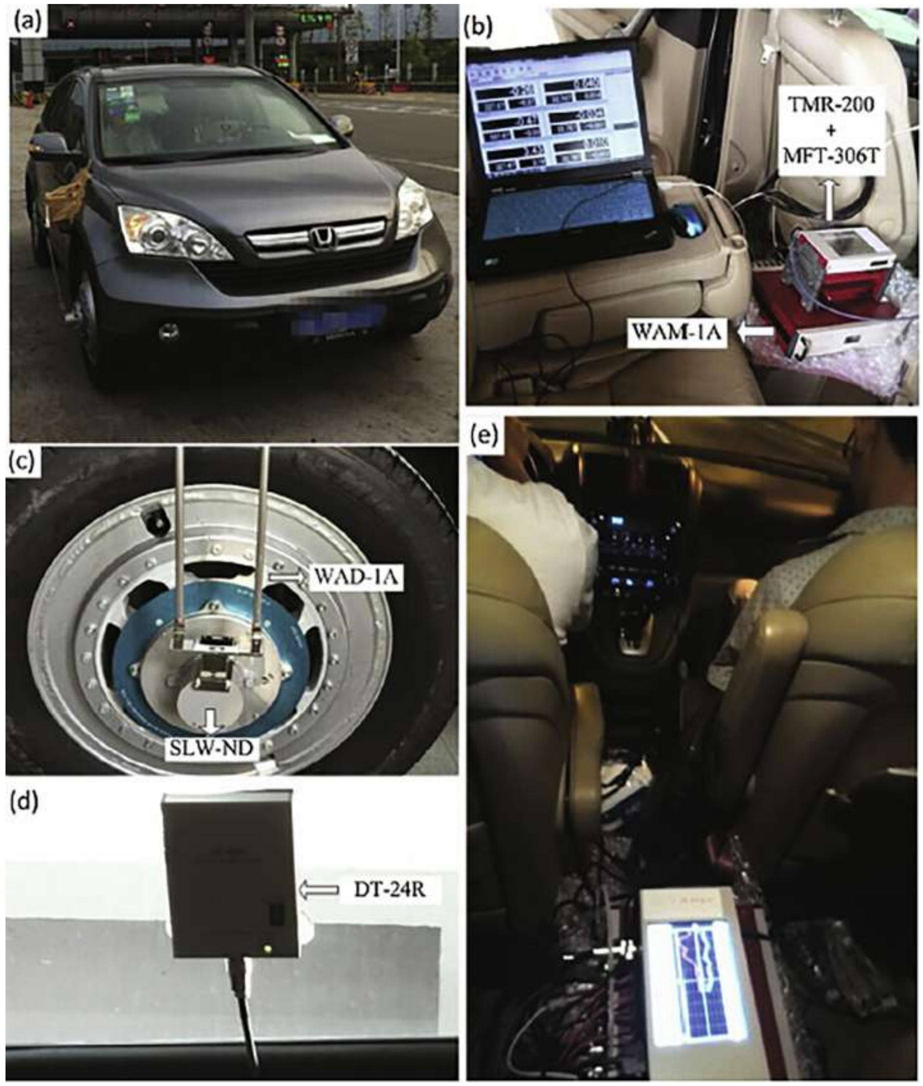

Fig 2. Experimental setup and testing: (a)–(d) Experimental HONDA CR-V car and automotive instrumentation; (e) field test.

Table 3
Experimental conditions.

Conditions	Content
Weather	Sunny, breeze
Traffic flow state	Free flow
Road section	Wenli Freeway South and North Lines, K117-K189
Road surface	Class A highway pavement
Experimental vehicle	Dongfeng Honda CR-V
Collection data type	Longitudinal force / moment, lateral force / moment, vertical force / moment, speed
Data collection interval	0.01 s
Test wheel	Right front wheel

drinking and drug use that had a detrimental influence on driving control. The driver remained sober and focused throughout the driving. Besides, to further ensure that the relationship among microscopic kinetic parameters, road characteristics, and accidents is accurately extracted, cruise mode was applied throughout the experiment. Such an experimental design can not only contribute to more accurate research results by minimizing the interference of driver factors on the relationship between kinetic parameters and accidents, but also ensure that the results will be of convenience, feasibility, and versatility.

However, such experimental design also has some defects. Due to the strict control of the driver and other realistic factors, the microscopic kinetic parameters may be insensitive to certain types of accident risks (such as traffic accidents dominated by driver characteristics) or road characteristic (such as road characteristics significantly related to driver characteristics). But as a preliminary investigation, this research has reached the expected goal. In the follow-up, larger-scale experiments covering different drivers, different vehicle types, and different weather could be organized to build a more complete traffic safety analysis theory with microscopic kinetic parameters as the core.

To ensure the validity and reliability of the data analysis, the raw data of various microscopic kinetic parameters obtained from the experiment were filtered and pretested before the model estimation. To eliminate the significant influence of the vehicle's weight on tire force, this paper intended to use the STD of various microscopic kinetic parameters in the time series to carry out the research. Selecting 1 s as the time window unit to calculate the STD sequence, the total STD of various microscopic kinetic parameters of each road segment was calculated to represent the fluctuation level of microscopic kinetic parameters of each segment (see Fig. 3).

Table 4
Part of the original data.

Sample clock				**10 msec**			
Type	Fx	Fy	Fz	Tx	Ty	Tz	V
Unit	kN	kN	kN	N·m	N·m	N·m	km/h
0	−0.612	−0.132	4.58	0.1232	0.0848	0.0144	105.915
0.01	−0.36	−0.048	4.588	0.1216	0.0768	0.0168	106.119
0.02	−0.436	0.068	4.848	0.072	0.0832	0.0288	106.46
0.03	−0.544	0.016	4.652	0.1096	0.0888	−0.0056	106.324
0.04	−0.348	0.068	4.68	0.1088	0.0808	0.0224	105.983
0.05	−0.372	−0.132	4.528	0.1344	0.0864	0.0096	105.983
0.06	−0.624	−0.06	4.62	0.1048	0.0848	0	105.983
0.07	−0.62	−0.168	4.344	0.1512	0.0832	−0.004	106.392
0.08	−0.412	−0.076	4.484	0.1088	0.0856	0.0352	106.528
0.09	−0.34	0.016	4.504	0.1216	0.0792	0.0296	106.392

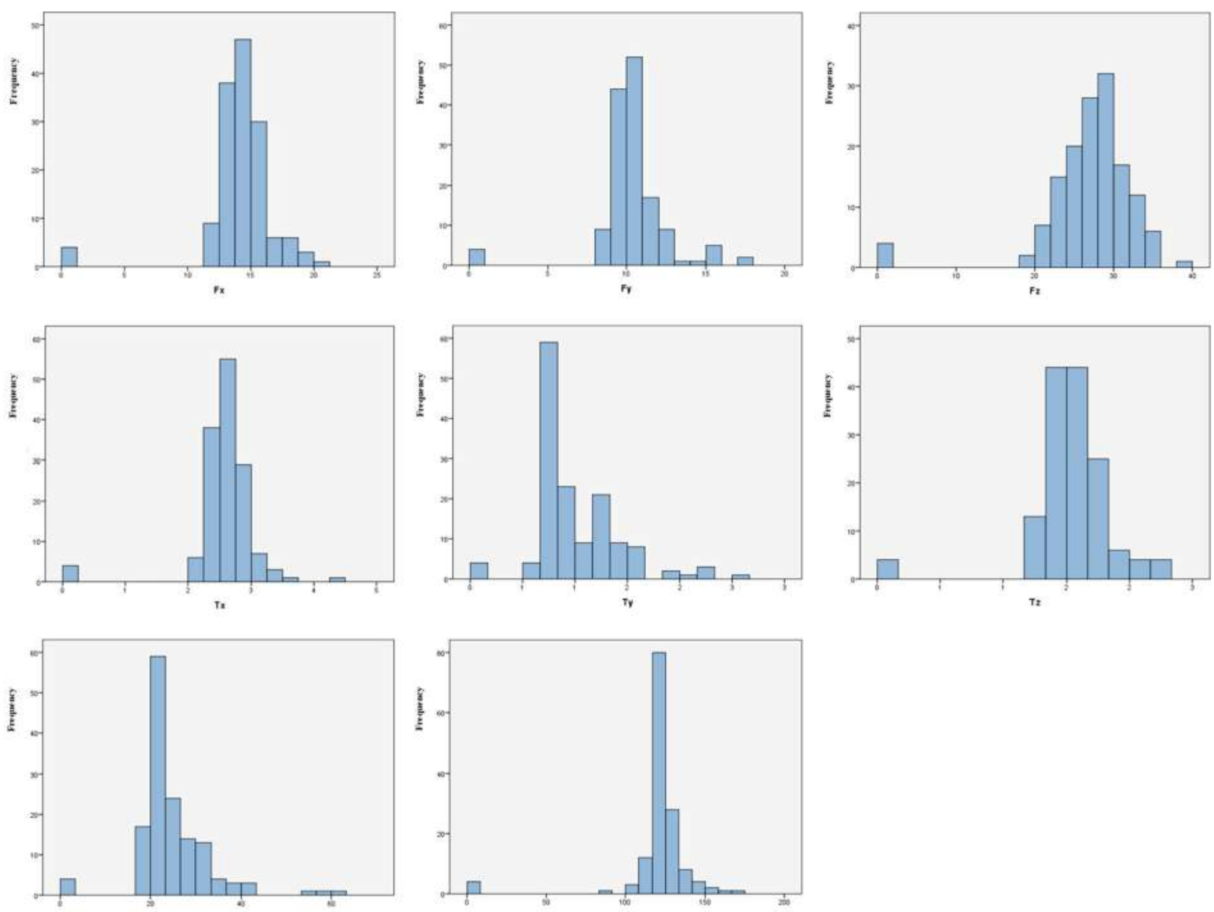

Fig 3. Frequency distribution of STD of microscopic kinetic parameters.

It can be seen from the frequency analysis that there are four road segments with missing kinetic data with the frequency of 0. In addition, the distribution of the STD level of the kinetic parameters of most segments is relatively concentrated. Only a few segments have large fluctuations in tendency, and the exposure of the kinetic data (refers to the number of STD for each type of kinetic parameters on each road segment) is mostly concentrated around 120, which implies that the vehicle is in a state of stable operation over most periods.

3. Methodology

3.1. Data processing

To eliminate the effects of different dimensions of indicator variables and ensure all continuous indicator variables will be on the same scale, the Max-Min standardized approach was used to in this paper transform the STD level of various kinetic parameters into the interval [0,1]:

$$std\_d_{ij} = \frac{d_{ij} - \min[d_j]}{\max[d_j] - \min[d_j]}\ j \in \{Fx, Fy, Fz, Tx, Ty, Tz, V\} \tag{1}$$

where std_d_{ij} is the standardized sizes of microscopic kinetic parameter j on segment i; d_{ij} is the total STD of microscopic kinetic parameter j on segment i; $\max[d_j]$ and $\min[d_j]$ are the maximum and minimum STD of the microscopic kinetic parameter j on all segments.

Then, the kinetic data are discretized by dividing the kinetic sequence $std\_d_{ij} = \{d_{1j,} d_{2j,} d_{3j,} \cdots, d_{ij}\}(j = 1, 2, \cdots 7)$ into three equal parts reflected by $K \sim \{0, 1, 2\}$(note as $D\_d_{ij}^{k}$, 0 means the divergence of the STD values in $std\_d_{ij} = \{d_{1j,} d_{2j,} d_{3j,} \cdots, d_{ij}\}$ is the smallest, which represents the most stable road segment condition; while 2 represents the largest divergence of this variable set, which denotes the least stable road segment condition). Likewise, to weaken the collinearity and heteroscedasticity between indicator variables, reduce the data scale, and make the data more stable, the logarithmic method (which will not change the correlation between variables and the monotonicity of variables) was used to balance variables "*horizontal radius*" and "*length of transition curve:*"

$$\begin{aligned} r_i &\sim Ln\ r_i \\ l_i &\sim Ln\ l_i \end{aligned} \tag{2}$$

where $r_i$and $l_i$represent the radius of the "*horizontal radius*" and "*length of transition curve*" of the road segment *i*.

3.2. Modeling approach

3.2.1. Binary logit model (FPL model)

Discrete choice model represented by Logit model has been widely used in accident cause analyses in the field of traffic safety (Abdel-aty, Hassan, Ahmed, & Al-ghamdi, 2012), of which the binary logit is the most basic one. For the accident type *m*, the risk probability of road segment *i* for this accident type being relatively high and relatively low can be, respectively, defined by the binary logit models as follows:

$$\begin{aligned} P(Y_{mi} = 1|X_i) &= \tfrac{1}{1+e^{-g(X_i)}} \\ P(Y_{mi} = 0|X_i) &= 1 - \tfrac{1}{1+e^{-g(X_i)}} \end{aligned} \tag{3}$$

where $X_i$is the sequence of explanatory variables; $Y_{mi} = 1$and $Y_{mi} = 0$indicate the segment being with relatively high accident risk and relatively low accident risk. The variable*odds*denotes the ratio of the observed "relatively high risk" to "relatively low risk" probability, after taking a logarithm of *odds*, the linear function is as follows:

$$\begin{aligned} In(odds) = In(\frac{P(Y_{mi} = 1|X_i)}{P(Y_{mi} = 0|X_i)}) = In(e^{g(x_i)}) = g(X_i) \\ = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots \end{aligned} \tag{4}$$

Additionally, to calculate the intercept β_0 and coefficients in Eq. (4), the maximum likelihood estimation method was used.

3.2.2. Standard ordered logit model

To deal with the multi-classification variables with specific orders such as the test scores (measured by ordinal variables such as A, B, C, D) and the economic income (measured by ordinal variables such as low, medium, and high), the standard ordered logit model (SOL model) is more powerful with the ability to analyze the probability of indicator variables determining the ordered outcome variables. In this paper, SOL model estimation will be performed to investigate the road segment stability $K(0,\ 1,\ 2)$ determined by the microscopic kinetic parameters $(j = \{Fx, Fy, Fz, Tx, Ty, Tz, V\})$. The utility function of the SOL model is as follow (Haghighi, Liu, Zhang, & Porter, 2018):

$$U_{ik} = \beta_k X_{ik} + \varepsilon_{ik} \tag{5}$$

where, β_κ is the coefficient of the explanatory variables $X_{i\kappa}$ and $\varepsilon_{i\kappa}$ is a disturbance term. For a certain type of microscopic kinetic parameter, the road segment stability *K* can be expressed as follows:

$$K = \left\{ \begin{array}{c} 0,\ if\ U_i \leqslant \mu_0 \\ 1,\ if\ \mu_0 < U_i \leqslant \mu_1 \\ 2,\ if\ U_i > \mu_1 \end{array} \right\} \tag{6}$$

where μ_0, μ_1 are the threshold values of different road segment stability categories 0, 1, 2 (stable, slightly fluctuating, seriously fluctuating) and the probability of different road segment stability categories can be calculated by:

$$\begin{aligned} P(K = 0) &= 1 - \tfrac{\exp(x_i\beta - \mu_0)}{1+\exp(x_i\beta - \mu_0)} \\ P(K = 1) &= \tfrac{\exp(x_i\beta - \mu_0)}{1+\exp(x_i\beta - \mu_0)} - \tfrac{\exp(x_i\beta - \mu_1)}{1+\exp(x_i\beta - \mu_1)} \\ p(K = 2) &= \tfrac{\exp(x_i\beta - \mu_1)}{1+\exp(x_i\beta - \mu_1)} \end{aligned} \tag{7}$$

3.2.3. Random parameters logit model

The RPL model can explain the heterogeneity among unobservable variables (Islam & Mannering, 2020; Mannering & Bhat, 2014). According to previous research (Li et al., 2019a, 2019b), the utility functions of different accident types are generally expressed linearly:

$$U_{mi} = \beta_m X_{mi} + \varepsilon_{mi} \tag{8}$$

where,β_{mi} is the parameter vector of the explanatory variable X_{mi} that affects the judgment of the accident type; ε_{mi} is the disturbance term indicating the unobserved heterogeneity. Assuming that ε_{mi} follows the generalized extreme value distribution, the probability of the road segment *i* being a segment with relatively high accident risk for type *m* can be calculated as follows:

$$P_{mi}|\beta_m = \frac{\exp(\beta_m X_{mi})}{1 + \exp(\beta_m X_{mi})} \tag{9}$$

Regarding the random distribution of unobserved factors among different observations, the probability of the RPL model can be expressed as follows (Liu & Fan, 2020):

$$P_{mi} = \int (P_{mi}|\beta_m) f(\beta|\varphi) d\beta \tag{10}$$

where $f(\beta|\varphi)$ represents the probability density function of the random parameter vectorβ, and φdenotes a vector of parameters describing the probability density function (mean and variance). In general, the distribution of random parameters varies, including lognormal distribution, normal distribution, uniform distribution, triangular distribution, and so forth. In addition, since the RPL model involves substantial integration of the logit formula over observation-specific parameters, which can be extremely time-consuming, its estimation is undertaken based on stimulated maximum likelihood method (Behnood & Mannering, 2015).

3.3. Margins effects analysis

For multi-classification models, the estimated results of model parameters cannot truly reflect the actual impact that the slight change in any specific explanatory variable has on the probability of the outcome variable (Osman, Paleti, Mishra, & Golias, 2016), so further marginal effect analysis of the model is required. Since there are both continuous and discrete variables in this study, different formulas are needed for calculation (Chang, Xu, Zhou, Chan, & Huang, 2019; Liu & Fan, 2020):

$$E_X^P = P(X = 1) - P(X = 0) \tag{11}$$

$$E_X^P = \frac{\partial P}{\partial X} \tag{12}$$

where, the marginal effects of discrete variables and continuous variables can be calculated using Eq. (9) and Eq. (10), respectively.

4. Results

4.1. Relationship between microscopic kinetic parameters and highway characteristics

4.1.1. SOL model results

The SOL model was used to analyze the fluctuation of the microscopic kinetic parameter variables ($Fx, Fy, Fz, Tx, Ty, Tz, V$). The estimated results of the model are shown in Table 5. As illustrated in Table 5, the road characteristics "*elevation_change,*" "*tunnel exit,*" and "*inside the tunnel*" have no significant effect on all the microscopic kinetic parameter variables.

With regard to "*Fx,*" it was found that the variables "*elevation_STD,*" "*bridge along the river,*" and "*diversion and merge*" have a positive effect on "*Fx*" (coefficients = 0.072, 1.056, and 2.690, respectively), that is, when the vehicle is driving on these segments, the longitudinal force on its tires may produce large fluctuations. In terms of "*Fy,*" "*horizontal radius,*" "*length of transition curve,*" and "*circular curve*" were found to have a negative effect on "*Fy*" (coefficients = -1.590, −1.142, and −0.298, respectively). On the contrary, "*bridge along the river,*" "*merge zone,*" and "*diversion and merge*" could be regarded to be positively related to "*Fy*" (coefficients = 0.993, 2.145 and 2.071), indicating that when the vehicle is driving on these segments, the lateral force on the tire may suffer the corresponding fluctuations. Additionally, "*Fz*" was found to be positively correlated with the variables "*both exit and entrance,*" "*bridge along the river,*" and "*diversion and merge*" (coefficients = 1.44, 1.379, and 2.948, respectively), and the increase of these variables will bring about the increase of the fluctuations in the vertical force on the tire.

Unlike the effects of force, the influence of road characteristics on the torque of the tire is majorly negative. For instance, it was found that as the variables "*elevation_STD,*""*circular curve,*""*horizontal radius,*""*length of transition curve,*" increase, the torque "*Tx*" tends to decrease (coefficients = -0.069, −1.512, and −0.900, respectively), and similar conclusions can also be observed in outcome variables "*Fy,*" indicating "*Tx*" and "*Fy*" may be of certain consistency. The variables "*elevation_STD,*" "*circular curve,*" "*composite alignment,*" and "*diversion zone,*" were found to have a restrictive influence on "*Ty*" (coefficients = -1.082, −1.735, −1.939, −2.076, respectively). Besides, "*horizontal radius*" appeared negative association with "*Tz.*" Among all the significant explanatory variables, the variable "*tunnel exit*" is the only road characteristic that has a positive effect on the torque variables, whose increase may cause a large fluctuation in the "*Tx*" (coefficient = 1.145).

With reference to the vehicle speed "*V,*" it was found that only the variable " *length of transition curve*" had a significant effect on the fluctuation of this outcome variable (coefficient = -0.290). This is consistent with the empirical evidence that a longer transition curve will reduce the difficulty of driving and make it easier for the driver to maintain speed stability.

Although the coefficients of the SOL model can reflect the quantitative relationship between explanatory variables and outcome variables to some extent, it is not enough to simply make a cross comparison of different coefficients among three outcome variables. Further exploration of the inherent mechanism of how one-unit change in any specific explanatory variable influences the outcome variables needs to be undertaken and this paper used marginal effect analysis to carry out more precise discussions.

4.1.2. Margins effects analysis

The marginal effect analysis shown in Table 6 can reflect the quantitative relationship between each explanatory variable and different level of outcome variables. In this study, seven microscopic kinetic parameter variables were divided into three categories (0-stable, 1-slightly fluctuating, 2-seriously fluctuating) according to the degree of fluctuation. The two categories of "0″ and "2″ were found to be sensitive to most of the road characteristic variables, while the category of "1″ was hardly affected. One possible explanation is that the category "1″ corresponds to the well-designed road segments, where the operation of vehicles is less likely to be disturbed by adverse road conditions, and the generation of abnormal kinetic behavior of vehicles is mainly attributed to human errors.

Every additional unit of the explanatory variable "*elevation_STD*" was found to generate the increase of the possibility of the category "2″ (seriously fluctuating) in the outcome variable "*Fx*" by 1.5%. When there is a river bridge within the road segment ("*bridge along the river*"=1), the probability of the category "2" (seriously fluctuating) in the outcome variable "*Fx* " is 21.1% higher than the road segment without it ("*bridge along the river*"=0). Likewise, the positive coefficient value of "*diversion and merge*" indicating that segments with such road characteristics generally sustained larger fluctuation degree in "*Fx.*"

Compared with the straight line, the "*circular curve*" showed a decline of 29.2% in the possibility of the category "2″ in the outcome variable "*Fy,*" indicating the increase of the stability of "*Fy*" in the road segment with "*circular curve.*" This situation can be explained by the driving compensation principle (i.e., the driver has a more cautious operation and a lower driving speed in a more

Table 5
Parameter estimation of the SOL model.

Variables	Fx	Fy	Fz	Tx	Ty	Tz	V
Elevation_change	-	-	-	-	-	-	-
Elevation_STD	0.072***	-	-	−0.069**	−1.082***	-	-
Geometry (base: straight segments)							
1.Circular curve (Geometry = 1)	-	−1.590**	-	−1.512**	−1.735**	-	-
2.Composite geometry(Geometry = 2)	-	-	-	-	−1.939**	-	-
Ln_Horizontal radius	-	−1.142***	-	−0.900***	-	−0.833**	-
Ln_Length of transition curve	-	−0.298***	-	-	-	-	−0.290***
Tunnel (base: without tunnel)							
1.Tunnel exit (Tunnel = 1)	-	-	-	1.145**	-	-	-
2.Tunnel entrance(Tunnel = 2)	-	-	-	-	-	-	-
3.Both exit and entrance(Tunnel = 3)	-	-	1.440***	-	-	-	-
4.Inside the tunnel(Tunnel = 4)	-	-		-	-	-	-
Along the river bridge	1.056**	0.993**	1.379***	-	-	-	-
Conflict zone (base: without conflict)							
1.Diversion zone(Conflict zone = 1)	-	-	-	-	−2.076**	-	-
2.Merge zone(Conflict zone = 2)	-	2.145**	-	-	-	-	-
3.Diversion and merge(Conflict zone = 3)	2.690***	2.071***	2.948***	-	-	-	-

Note:* means $p \leq 0.1$, ** means $p \leq 0.05$, *** means $p \leq 0.01$.

Table 6
The result of margins effects.

Category	Variables	Fx	Fy	Fz	Tx	Ty	Tz	V
0	Elevation_STD	−0.015	-	-	0.013	-	-	-
	Geometry (base: straight segments)							
	Circular curve	-	0.265	-	0.250	0.264	0.239	-
	Composite geometry	-	-	-	-	0.302	-	-
	Ln_Horizontal radius	-	0.215	-	0.171	0.208	0.156	-
	Ln_Length of transition curve	-	0.056	-	-	-	-	0.056
	Tunnel (base: without tunnel)							
	Tunnel exit	-	-	-	−0.184	-	-	-
	Both exit and entrence	-	-	−0.246	−0.159	-	-	-
	Inside the tunnel	-	-	−0.178	-	-	-	-
	Along the river bridge	−0.212	−0.185	−0.268	-	-	-	-
	Conflict zone (base: without conflict)							
	Diversion zone	-	-	-	-	0.438	-	0.414
	Merge zone	-	−0.278	−0.219	-	-	-	-
	Both the diversion and merge	−0.290	−0.274	−0.310	-	-	-	−0.184
1	Bothe diversion and merge	−0.256	-	−0.255	-	-	-	-
2	Elevation_STD	0.015	-	-	−0.014	-	-	-
	Geometry (base: straight segments)							
	Circular curve	-	−0.292	-	−0.291	−0.322	−0.269	-
	composite geometry	-	-	-	-	−0.355	-	-
	Ln_Horizontal radius	-	−0.208	-	−0.177	−0.215	−0.156	-
	Ln_Length of transition curve	-	−0.054	-	-	-	-	−0.058
	Tunnel (base: without tunnel)							
	Tunnel exit	-	-		0.234	-	-	-
	Both exit and entrence	-	-	0.296	-	-	-	-
	Along the river bridge	0.211	0.179	0.267	-	-	-	-
	Conflict zone (base: without conflict)							
	Diversion zone	-	-	-	-	−0.264	-	−0.252
	Merge zone	-	0.437	-	-	-	-	-
	Both the diversion and merge	0.546	0.423	0.565	-	-	-	-
Note:		Indicates a positive relationship Indicates a negative relationship						

complicated environment such as a segment with circular curve; Haleem & Gan, 2013), which makes it easier for the vehicle to keep running smoothly and generates more moderate longitudinal and vertical force fluctuations of the tires. With regard to the influence of log variables "*horizontal radius*" and "*length of transition curve*" in the outcome variables, one-unit increase of them will bring about the decrease of the possibility of the category "2" of "*Fy*" by 20.8% and 5.4%, respectively, implying that the gentle curve helps vehicles keep stable operation. Besides, the road characteristics of "*bridge along the river,*" "*merge zone,*" and "*diversion and merge*" were found to make the lateral force "*Fy*" suffer greater fluctuation (the increase in one-unit will lead to the probability of "2" of "*Fy*" increasing by 17.9%, 43.7%, and 42.3%, respectively). This may be the result that drivers tend to adopt more lane-changing behaviors in the conflict zone, triggering the increase in the lateral force fluctuation of the tire (Jiang & Dong, 2012).

Considering "*Fz,*" when there are both tunnel exits and entrances in the road segment, the probability of "2″ (seriously fluctuating) in "*Fz*" was associated with an increase of 29.6% compared to the ordinary segments without tunnels. In addition, the other two variables ("*bridge along the river*" and "*diversion and merge*") have similar effects with the aforementioned, causing the probability of "seriously fluctuating" increased by 26.7% and 56.5%. There may be three possible reasons for the above situation. The first is the abnormal reaction of the driver caused by the sudden change of the driving environment (such as brightness and lane width; Hoeven, 2011; Rudin-brown, Young, Patten, Lenné, & Ceci, 2013). Secondly, the road surface materials of the tunnel and bridge are generally different, resulting in different road response to the tires. The third explanation is that different external conditions will bring about different pavement performances., that is, the tunnel environment is characterized with minor temperature difference, high humidity, abundant groundwater and surface water and being such long-term humid state causes the adhesion performance and damage degree of tunnel pavement to be significantly different from that of ordinary road segments (Cong, Chen, Zheng, & Zhou, 2020), resulting in differences of tire force. With regard to the conflict zone, whose surrounding differs from the tunnel, it was also found to have a tendency of serious fluctuation of "*Fz,*" with the reasons mainly attributed to the frequent braking and acceleration operations caused by the traffic flow intersection.

Through the fluctuation analysis of "*Fx,*" "*Fy,*" and "*Fz,*" it was found that the existence of "*bridge along the river*" will lead to an increase in the probability of "seriously fluctuating" of all three outcome variables. Currently, regarding this phenomenon, some scholars have put forward relevant views on this: although bridges have promoted the development of transportation in areas with mountains and rivers, they also bring about the problems of bridge vibration and strong winds on passing vehicles, while the sudden crosswise wind and the resonance of wind-vehicle-bridge will cause uneven forces in X, Y, and Z directions of tires, resulting in a significant increase in the fluctuation degree (Han, Hu, Cai, & Li, 2013).

Similarly, some road characteristic variables also have significant impacts on the probability of being seriously fluctuating in "*Tx,*" "*Ty,*" "*Tz,*" and "*V.*" It is worth mentioning that as the proportion of "*diversion and merge*" being "1″ value increases, the probability of being "stable" in "*V*" will be reduced by 18.4%, while it will not lead to more occurrences of being "seriously fluctuating." The same situation also occurs in "*Fz*" and "*Tx*" under the influence of "*merge zone*" and "*both exit and entrance,*" respectively.

From the aforementioned results, it can be seen that the microscopic kinetic characteristics of tires under normal driving circumstances are generally related to the complex road characteristics such as tunnels, bridges, diverging/merging zones, circular curves, but it is still unclear what the connection between the abnormal fluctuation of these micro-kinetic variables and driving safety is.

To address this issue, the relationship between the microscopic kinetic parameters and different kinds of traffic accidents will be further discussed.

4.2. Relationship between microscopic kinetic parameters and accidents

This section first used factor analysis to extract two mutually independent kinetic factors, which are named tire cornering kinetic characteristic (TCKC) and tire longitudinal kinetic characteristic (TLKC), from several highly correlated microscopic kinetic parameters. Then we built the FPL model and the RPL model to explore the relationship between microscopic kinetic parameters and different types of traffic accidents, and found some interesting phenomena. The flowchart in this section is provided in Fig. 4.

4.2.1. Factors analysis

The relationship of microscopic kinetic parameters and accidents was examined through a correlation analysis of seven microscopic kinetic parameter variables to avoid model estimation deviation caused by high collinearity and correlation in modeling. Since the variables of the microscopic kinetic parameters are all discrete variables, the Spearman correlation coefficient method was adopted. As presented in Fig. 5, most of the microscopic kinetic parameter variables are positively correlated, indicating high collinearity exiting among variables; only "*Fz*" is weakly correlated with others.

To deal with the collinearity of microscopic kinetic parameters, factor analysis was conducted. Factor analysis (FA) is based on the principle of least information loss (generally with the eigenvalue > 1 or cumulative variance > 80%) to extract unobservable factors affecting explanatory variables with high correlation and convert them into linear combinations of independent factors

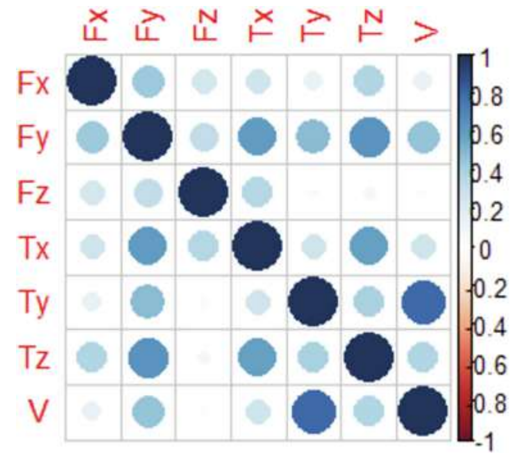

Fig 5. Correlation analysis of micro kinetic parameters.

(Johnson & Wichern, 1992). The model construction and matrix expression is shown in Eq. (11) and Eq. (12):

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \ldots + \ell_{1h}F_h + \varepsilon_1 \\ X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \ldots + \ell_{2h}F_h + \varepsilon_2 \\ &\vdots \; \vdots \; \ddots \; \vdots \; \vdots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \ldots + \ell_{ph}F_h + \varepsilon_p \end{aligned} \quad (11)$$

$$(\mathrm{X} - \mu)_{p\times 1} = \mathrm{L}_{p\times h}\mathrm{F}_{h\times 1} + \varepsilon_{p\times 1} \quad (12)$$

where, F_h is the extracted new factor; ℓ_{ph} is the factor load with the value between 0 and 1, while 1 indicates the greatest the influence

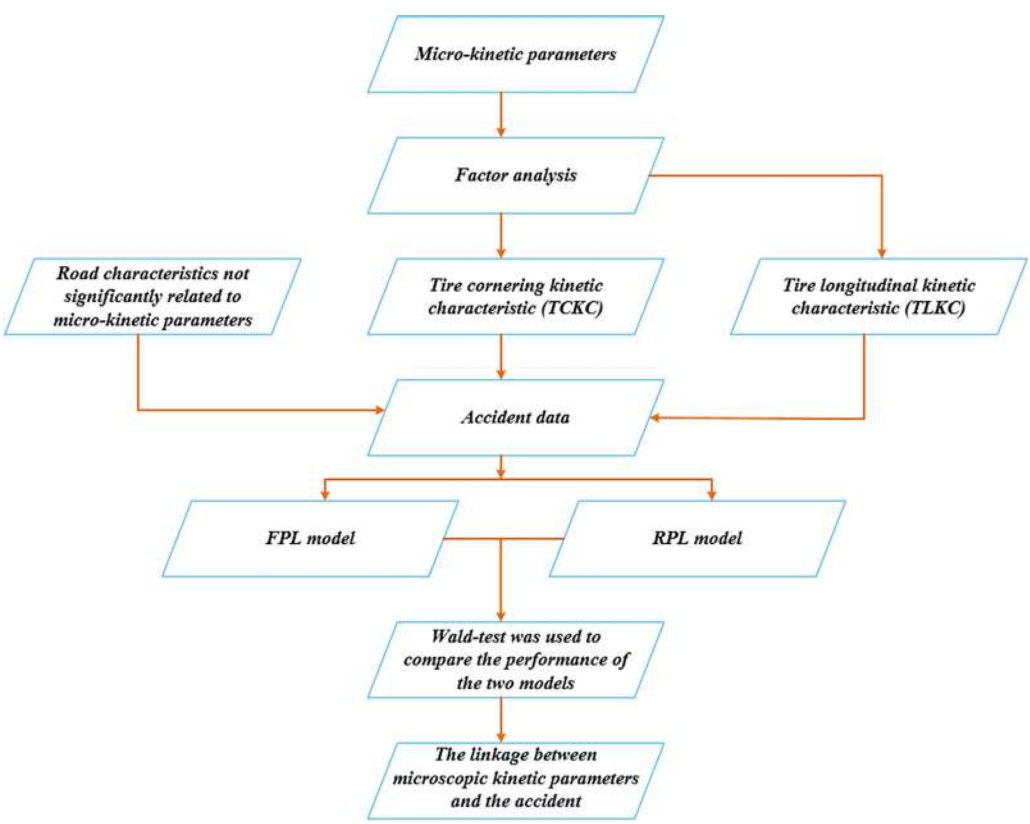

Fig 4. The framework of Section 4.2.

of ℓ_{ph} has on the original variable (Washington, Karlaftis, & Mannering, 2010); ε_p is the random error term only related to X_p.

KMO (Kaiser-Meyer-Olkin) test was carried out before FA. According to statistics experience, when the KMO value is more than 0.6, FA could be conducted (De Castro et al., 2017; Yang, Feng, Zhao, Jiang, & Huang, 2020). In this case, the original KMO value of seven variables is less than 0.6. Considering the correlation of "*Fz*" and other variables is not significant, "*Fz*" was excluded in FA process and the KMO value of 0.7456 was finally obtained, indicating an acceptable internal consistency reliability (Bian, 2017; Wei et al., 2012).

As illustrated in Table 7, factor 1 highly reflects the cornering kinetic characteristic including "*Fy,*" "*Tx,*" and "*Tz,*" and factor 2 highly reflects the longitudinal kinetic characteristic including "*Ty*" and "*V*." The two factors were thus interpreted as tire cornering kinetic characteristic (TCKC) and tire longitudinal kinetic characteristic (TLKC), respectively, according to the distribution law of high loading variables and the tire mechanical characteristics (as shown in Fig. 6). To have a primary understanding of the road segment performance of these two factors, the corresponding factor scores were calculated and shown in Fig. 7.

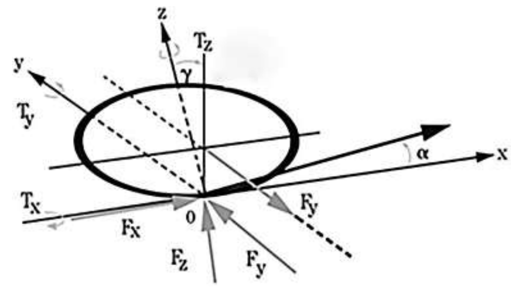

Fig 6. Schematic diagram of tire kinetics.

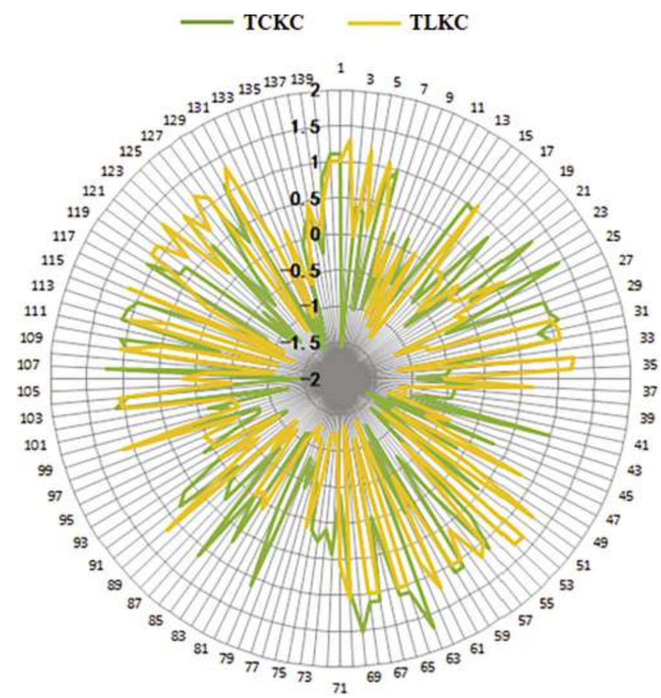

Fig 7. Radar chart of factor scores.

4.2.2. FPL model results

A primary FPL model was first performed to analyze the quantitative relationship of microscopic kinetic parameter variables and traffic accidents. In order to improve the accuracy and feasibility of the model, three road characteristic variables with no correlation with microscopic kinetic parameters were incorporated into the model, including the "*elevation_change,*" "*inside the tunnel,*" and "*tunnel entrance.*"

Previous to model estimation, the term relating to "*road segment accident rate*" was defined based on the accident numbers of certain types happening in the chosen road segment; for example, if the number of rear-end accidents on a road segment exceeds the average number of the samples, the road segment can be defined as a "*road segment with a high rear-end accident rate*" (variable value = 1 if "yes" and = 0 if "no"), otherwise it is a "*road segment with a low rear-end accident rate*" (variable value = 1 if "yes" and = 0 if "no"). Likewise, this method was used to complete the definition of the other two accident types (fixed-object hitting and overturning accidents).

Table 8 presents the estimated results of the FPL model for rear-end, fixed-object hitting, and overturning accidents. The model results show that under normal driving conditions, the real-time output of TCKC has no remarkable relationship with all three accident types, which does not indicate that the tire cornering kinetics characteristics are not correlated with safety. A possible explanation may be that the driving test was conducted under a windless, rainless, and snowless environment, which decreased the fluctuation of the operation of the vehicle and led to the ambiguous results of the model estimation, while the severe weather conditions such as rain, snow, and crosswinds are more likely to cause bad kinetic behaviors of high-speed vehicles such as side deflection, skidding, and deflection (He, Liu, Chen, & Zhao, 2011). Therefore, more comprehensive real-time driving experiments including various driving condition need to be further developed. Additionally, the variables "*elevation_change*" and "*Fz*" were also found to be unrelated to all three accident types .

Table 7
Results of FA of the remaining six variables.

Variable	Factor 1:tire cornering kinetic characteristic (TCKC)	Factor 2: tire longitudinal kinetic characteristic(TLKC)
Fx	0.4730	0.0655
Fy	**0.7432**	0.3632
Tx	**0.6769**	0.1624
Ty	0.2175	**0.8350**
Tz	**0.7116**	0.2867
V	0.1956	**0.8259**

The coefficient value of TLKC was negative related with the rear-end accidents ($p \leq 0.01$). It indicates that when TLKC increases by one unit, the occurrence of the value "1″ of the variable *"road segment with a high rear-end accident rate*" will be reduced by 0.64 times (the value "1" means that the road segment is with a high rear-end accident rate). This result is original but significant, since it reveals that the more unstable the tire is, the less likely it is that a rear-end accident will occur, contrary to the traditional concept that the stable operation of vehicles symbolizes higher traffic safety. A more in-depth analysis will be shown in the "Discussion" section. TLKC will be used as an intermediate variable between road characteristics and traffic accidents, so as to more accurately interpret the results of the models.

In terms of road characteristics, compared with those segments without the tunnel, when there is only a tunnel entrance (the tunnel begins at certain part of the targeted road segment while ending at another road segment), it's more likely for this segment to suffer a higher probability of rear-end accidents, fixed-object hitting and overturning accidents (coefficients = 1.92, 1.8 and 2.11); when the road segment is completely contained in the tunnel (the tunnel both begins and ends at other road segments), the

Table 8
Estimation results of FPL model.

Variables	Rear-end		Fixed-object hitting		Overturning	
	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Constant	−0.647985***	−3.138	−0.3248408*	−1.715	−0.537994***	−2.705
TCKC	−0.125863	−0.818	0.0420165	0.301	0.004820	0.034
TLKC	−0.443961***	−3.198	0.0884184	0.703	0.145536	1.103
Fz	−0.079634	−0.497	0.0055200	0.038	−0.037168	−0.247
Elevation_change	0.002377	0.511	−0.0003913	−0.083	0.003730	0.749
Tunnel entrance	0.652842**	2.286	0.5902991**	2.164	0.748793**	2.719
Inside the tunnel	0.654032*	1.856	0.4297278	1.207	0.934856***	2.635
Log Likelihood	−77.32		−93.24		−86.55	

Note:* means $p \le 0.1$, ** means $p \le 0.05$, *** means $p \le 0.01$.

probability of "*road segment with a high rear-end accident rate*" as well as "*road segment with a high overturning accident rate*" will increase (coefficients = 1.92, 2.55)

It is demonstrated that the tunnel has laid a serious hidden danger for traffic safety. This is because the environment of the tunnel is closed and dark, and drivers will have physiological pressure such as tension and fear to different degrees when entering and leaving the tunnel (Mehri et al., 2019; Zhou et al., 2020). In other words, the sudden changes in environmental conditions and abnormal driver responses may cause the tunnel to become a place where various types of accidents occur frequently and accumulate, especially in tunnel entrances (Dai, Guo, Ma, & Ni, 2010).

Nevertheless, the kinetic characteristics of the tires are undesirable in identifying fixed-object hitting and overturning accidents. It can be speculated that these two types of accidents are more correlated with driver characteristics, such as driving distraction, driver personality differences, and driving habits. However, in the experimental design of this paper, the driver was familiar with the purpose of the experiment and focused their attention on driving throughout, which is consistent with previous studies that believed that participants will be influenced by expectancy effect and tend to show an inherent need to perform well under test conditions (Harvey & Burnett, 2019). Moreover, the use of cruise mode might further restrict the impact of driver characteristics. In addition, the vehicle in this experiment is a small car, and its kinetic characteristics are different from that of a larger one, which is more likely to cause traffic accidents such as overturning, skidding, and deflection because of the higher center of gravity and larger windward area on the side (Miller, Davis, Reed, Doraiswamy, & Fu, 2003). On the other hand, overturning, skidding, and deflection may further cause vehicles to hit the roadside guardrails, isolation piles, and other fixed facilities, thereby increase the occurrence of fixed-object hitting accidents.

In other words, driver characteristics and vehicle types may offset the influence that vehicle kinetic factors have on the occurrence and frequency of fixed-object hitting and overturning accidents, indicating that more relevant research should be implemented in future research; that is, it is necessary to carry out extensive ranges of the naturalistic driving experiments in China to explore more detailed driving information relating to Chinese road conditions, especially the relationship between the kinetic characteristics of vehicles and traffic accidents.

4.2.3. RPL model results

This paper focused on the "link role" of microscopic kinetic parameters of vehicle tires in the road characteristics and accidents, and based on the previous results of FPL model, it was found that fixed-object hitting and overturning accidents were unrelated to the microscopic kinetic parameters. Therefore, only rear-end accident model will be estimated in this section with three significant variables TLKC, "*inside the tunnel,*" and "*tunnel entrance*" being identified as random parameters. Table 9 shows the comparison results of the RPL model and the FPL model.

Wald-test was used to compare the performance of the two models (Sarrias, 2016). The resulting χ^2 statistic can be used to determine if the null hypothesis that the performance of the RPL model and the FPL model are equal can be rejected. As illustrated in Table 10, *p*-value ($\le$0.01) shows that the null hypothesis is rejected and the performance of the RPL model is superior to the FPL model.

The RPL model estimation results indicate that the explanatory variable TLKC with random parameters follows the normal distribution (mean = -0.52, standard error = 0.36) which is recorded as $X \sim N(-0.52, 0.36^2)$. This reveals that under a probability of 92.64%, TLKC may have a negative influence on the occurrence of rear-end accidents, while under a probability of 7.36% it will exhibit the opposite tendency.

The road characteristic variables "*inside the tunnel*" and "*tunnel entrance*" were identified to obey uniform distribution. This paper also found that 99.99% of the segments with only tunnel entrance were more likely to trigger rear-end accidents compared to those segments without tunnels. Likewise, random parameter estimation of "*inside the tunnel*" illustrates that under a probability of 53.77%, the segments completely contained inside the tunnel have a tendency to be positively related to rear-end accidents (while the remaining 46.23% probability leads to the opposite conclusion). This also explains why the variable "*inside the tunnel*" is not significant in the mean value. It may be due to the counteracting effect of the random parameter distribution, which is balanced between the negative and positive coefficients. From this result, it can also be demonstrated that RPL model is more explanatory in understanding the heterogeneity problems.

5. Discussion

Since only rear-end accidents were found to be significantly correlated with the microscopic kinetic parameters generated by tires during normal driving, this section only discussed rear-end accidents.

Only considering the relationship between the microscopic kinetic parameters and traffic accidents may draw fragmentary conclusions that the increase of divergences of longitudinal kinetic characteristic indicators such as "*Ty*" and "*V*" will lead to decrease in the rear-end accidents, which is obviously unrealistic since limited by the driving road environment, the microscopic kinetic characteristics of the tire have thresholds for increasing and decreasing. On the other hand, the relationship between road characteristics and rear-end accidents cannot fully reflect the changes of vehicle driving status (i.e., although the results can be obtained, the inherent mechanism is still ambiguous).

Therefore, it is necessary to consider the simultaneous effects of the road characteristics, microscopic kinetic parameters, and traffic

Table 9
Estimation results of RPL model.

Variables	FPL model		RPL model	
	Coefficient	t-stat.	Coefficient	t-stat.
Constant	−0.647985***	−3.138	−0.789294 ***	2.582
TCKC	−0.125863	−0.818	−0.163777	−0.860
TLKC	−0.443961***	−3.198	−0.519908**(0.358710)	−2.195
Fz	−0.079634	−0.497	−0.005616	−0.029
Elevation_change	0.00237	0.511	0.001939	0.370
Tunnel entrance	0.652842**	2.286	0.697090**(0.001443)	2.031
Inside the tunnel	0.654032*	1.856	0.703528(7.435321)	0.373
Log Likelihood	−77.32		−76.27	

Note:* means $p \le 0.1$, ** means $p \le 0.05$, *** means $p \le 0.01$; values in parentheses indicate the standard error of the random parameters.

Table 10
Result of Wald-test.

Model	χ^2	P-value
FPL model	-	-
RPL model	18.854	0.0002931 ***

accidents, by employing the microscopic kinetic parameters as a critical link to explain the internal mechanism of how the road characteristics affect traffic safety.

It can been seen in Table 11 that microscopic kinetic parameters (*Ty* and *V*) has different impacts on determining the occurrence of rear-end accidents under different road circumstances. Take straight segments as the base group, the segments with the geometry characteristics of "*circular curve*" and "*composite alignment*" will reduce the possibility of "*Ty*," producing large fluctuations and thus lead to the increase of the probability for this segment to suffer rear-end accidents. Likewise, when there is a diversion zone in a road segment, the "*Ty*" and "*V*" tend to decline compared to the segment without any conflict area, which reflects the increased probability of the segment becoming a high-risk segment of a rear-end accident.

Another interesting finding is that when there are both diversion and merging zones existing in the road segment, the occurrence of rear-end will be reduced, presumably because the drivers will be more cautious and focused when the road environment is complicated to a certain degree. The similar findings were also obtained considering the other two road characteristics, *"horizontal radius*" and *"length of transition curve.*" Generally speaking, the larger the radius and the longer the transition curve, the lower the driver's operation requirements, and the easier the vehicle drives smoothly. But in this environment, the driver is more likely to be distracted, and the concentration of driving will be significantly reduced, resulting in an increased risk of rear-end accidents.

Also, road characteristics such as tunnels can directly affect the road traffic safety. As exhibited in Fig. 8, tunnel-related variables can increase the accident risks of all three types (by setting the segments without tunnels as a base [note the value is 1] according to the frequency of the segments with relatively high accident rate). Additionally, it can be seen intuitively that the tunnel entrance is more likely to cause rear-end accidents, and the tunnel exit is more likely to cause fixed-object hitting accidents. When a road segment has both a tunnel exit and a tunnel entrance, the risk of accidents is greater than in other cases. Therefore, how to ameliorate the design and management of the tunnel traffic is of extraordinary research significance for accident prevention.

Table 11
Relationship between microscopic kinetic parameters, road characteristics and rear-end accident.

Variables	TLKC		Risk of rear-end
	Ty	V	
Circular curve (base: straight segments)	Decrease	-	Increase
Composite geometry (base: straight segments)	Decrease	-	Increase
Ln_Horizontal radius	Decrease	-	Increase
Ln_Length of transition curve	–	Decrease	Increase
Diversion zone(base: without conflict zone)	Decrease	Decrease	Increase
Diversion and merging zone(base: without conflict zone)	-	Increase	Decrease

6. Conclusion

This paper took a mountainous freeway in Zhejiang Province, China as an object. By installing a six-component force meter on the tire, the microscopic kinetic parameters ($Fx, Fy, Fz, Tx, Ty, Tz, V$) of the tire's real-time output under normal driving conditions were obtained. For statistical analysis, a SOL model was primarily built to analyze the quantitative relationship between microscopic kinetic parameters and road characteristics (such as tunnels, geometry alignments, bridges along rivers, conflict zones). Then, a FPL model and a RPL model were estimated and compared to further explore the internal correlation of microscopic kinetic parameters and traffic accidents with different categories (rear-end, fixed-object hitting, overturning). Lastly, the link function of microscopic kinetic parameters was also investigated to unravel the inherent mechanism of how road characteristics lead to roads having certain types of accident-prone segments by combining the results of the above models

The findings of this paper demonstrated that the road characteristics including "*circular curve,*" "*composite alignment,*" "*tunnel exit,*" and "*diversion zone*" will cause longitudinal kinetic characteristics to stabilize (shown as a smaller value). At the same time, the kinetic variables TLKC have a significant negative relationship on the occurrence of rear-end accidents on the road segments. It can be concluded that the above-mentioned road characteristics can be defined as potential risk-trigger factors that result in higher occurrence and frequency of rear-end accidents by affecting the longitudinal kinetic of the tires. In addition to road characteristics, other external conditions (e.g., stagnant water, crosswind, turbulent traffic flow, and driver distraction) may also be triggers, which could be further studied in the future.

Driving compensation principles were also found in this study. For example, the road elements "*circular curve*" and "*diversion and merging,*" generally considered to be related with higher driving risk, were found to decrease the fluctuations of *"Fy"* and the probability of the occurrence of rear-end accidents. This can be explained by driving compensation principles, that is, when driving environments such as road geometry design and traffic flow become complicated, the drivers will be more cautious and con-

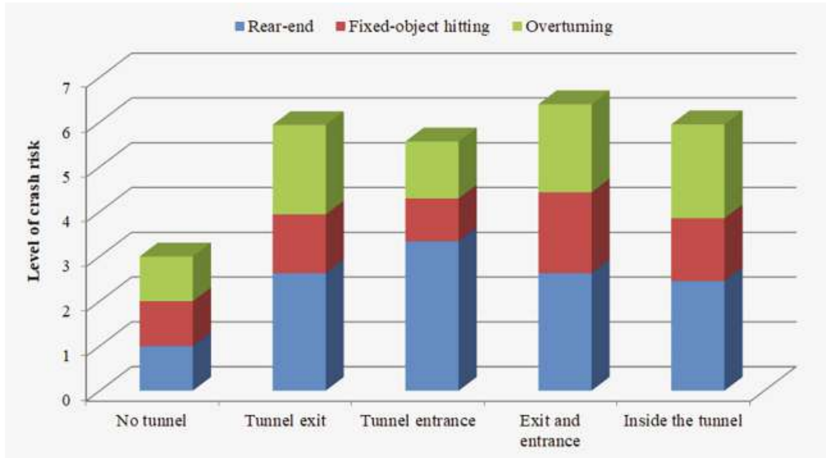

Fig 8. Accident risk under different tunnel-related settings.

centrated. It is an interesting finding and further research could be conducted from this perspective.

The main contribution of this study can be summarized from the following three aspects:

- The method proposed in this paper provides an innovative insight of how to employ the data obtained from a real-time vehicle experiment to analyze the accident risk on different road segments based on microscopic kinetic characteristics of the vehicle. The traditional accident analyses tended to utilize the historical accident report, which is less useful to provide real-time perceptions and warnings of accidents compared to the vehicle microscopic kinetic parameter. The results show that this was successful in the is study, as longitudinal kinetic characteristics are highly correlated with the risk of rear-end accidents, which can be used to predict whether any specific location is of high risk of some specific accidents, such as rear-end, fixed-object hitting, overturning.
- The "link role" in traffic accidents was investigated. Unlike previous studies that directly used road characteristics and traffic accidents to build models for analysis, this paper used microscopic kinetic parameters as intermediate variables to construct two types of models to deeply analyze the mechanism of the rear-end accidents occurred. The advantage of this approach is that it can intuitively observe the changes in the state of the vehicle when driving on a road with high accident risk, and these changes are important conditions that cause accidents. That is, when these changes are superimposed with some other factors (rain, snow, driver characteristics, traffic flow characteristics), it is more likely to cause an accident.
- The conclusions of this article lay the foundation for subsequent research on the relationship between vehicle microscopic force and driving safety on a larger scale, and thus form a complete "safe driving theory" based on vehicle microscopic force, which can provide a new perspective for the design and active safety control of intelligent vehicles in the future (e.g., trying to control the operation of the vehicle in a certain running state).

However, this study is not free of limitations; as it is a preliminary investigation, the current research has not yet formed a system. More data of microscopic kinetic parameters in dynamic environments such as rain, snow, and winds could be obtained in future studies. Also, simulation experiments using multi-body dynamics software such as ADAMS/Car should be further undertaken.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank National Natural Science Foundation of China (Grant No. 51778141, 52072069 and 71871078), Transportation Department of Henan Province (Grant No. 2018G7), Transportation Department of Zhejiang Province (sponsored by Project 2012H12), Scientific Research Foundation of Graduate School of Southeast University (YBPY2166), and Jiangsu Creative PHD student sponsored project (KYCX20_0138). Their assistance is gratefully acknowledged.

References

Abdel-aty, M. A., Hassan, H. M., Ahmed, M., & Al-ghamdi, A. S. (2012). Real-time prediction of visibility related crashes. *Transportation Research Part C: Emerging Technologies, 24*, 288–298. https://doi.org/10.1016/j.trc.2012.04.001.

Abdel-Aty, M. A., & Pemmanaboina, R. (2006). Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data. *IEEE Transactions on Intelligent Transportation Systems, 7*(2), 167–174. https://doi.org/10.1109/TITS.2006.874710.

Abrari Vajari, M., Aghabayk, K., Sadeghian, M., & Shiwakoti, N. (2020). A multinomial logit model of motorcycle crash severity at Australian intersections. *Journal of Safety Research, 73*, 17–24. https://doi.org/10.1016/j.jsr.2020.02.008.

Behnood, A., & Mannering, F. L. (2015). The temporal stability of factors affecting driver-injury severities in single-vehicle crashes : Some empirical evidence. *Analytic Methods in Accident Research, 8*, 7–32. https://doi.org/10.1016/j.amar.2015.08.001.

Beshah, T., & Hill, S. (2010). Mining road traffic accident data to improve safety: Role of road-related factors on accident severity in Ethiopia. *AAAI Spring Symposium - Technical Report*, 14–19.

Bian, Y. (2017). Research and Application of Stractual Equation Model. Dalian Maritime University.

Castro, B. L. De, Gracia, F. J., Tomás, I., & Peiró, J. M. (2017). The Safety Culture Enactment Questionnaire (SCEQ): Theoretical model and empirical validation. Accident Analysis and Prevention, 103, 44–55. https://doi.org/10.1016/j.aap.2017.03.018

Chang, F., Xu, P., Zhou, H., Chan, A. H. S., & Huang, H. (2019). Investigating injury severities of motorcycle riders: A two-step method integrating latent class cluster analysis and random parameters logit model. *Accident Analysis and Prevention, 131*, 316–326. https://doi.org/10.1016/j.aap.2019.07.012.

Chen, C., Zhang, G., Tarefder, R., Ma, J., Wei, H., & Guan, H. (2015). A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. *Accident Analysis & Prevention, 80*, 76–88. https://doi.org/10.1016/J.AAP.2015.03.036.

Cong, Z., Chen, H., Zheng, N., & Zhou, W. (2020). Surface Texture of Cement Concrete Pavementk:A Review. *Materials Reports, 34*(5), 110–116.

Dadashova, B., Ramírez, B. A., McWilliams, J. M., & Izquierdo, F. A. (2016). The Identification of Patterns of Interurban Road Accident Frequency and Severity Using Road Geometry and Traffic Indicators. *Transportation Research Procedia, 14*, 4122–4129. https://doi.org/10.1016/j.trpro.2016.05.383.
Dai, Y., Guo, Z., Ma, Y., & Ni, H. (2010). Safety evaluation index of driving environment of expressway tunnel. *journal of Tongji University: Natural Science, 38*(8), 1171–1176.
De Oña, J., Mujalli, R. O., & Calvo, F. J. (2011). Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks. *Accident Analysis and Prevention, 43*(1), 402–411. https://doi.org/10.1016/j.aap.2010.09.010.
Dingus, T. A., Klauer, S. ., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., & Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R. . (2006). The 100-Car Naturalistic Driving Study Phase II – Results of the 100-Car Field Experiment. In United States. Department of Transportation. National Highway Traffic Safety Administration. https://doi.org/10.1211/pj.2017.20203911
Dong, B., Ma, X., Chen, F., & Chen, S. (2018). Investigating the differences of single-vehicle and multivehicle accident probability using mixed logit model. *Journal of Advanced Transportation, 2018*, 1–9. https://doi.org/10.1155/2018/2702360.
Ellison, A. B., Greaves, S. P., & Bliemer, M. C. J. (2015). Driver behaviour profiles for road safety analysis. *Accident Analysis and Prevention, 76*, 118–132. https://doi.org/10.1016/j.aap.2015.01.009.
Fan, W., Kane, M. R., & Haile, E. (2015). Analyzing Severity of Vehicle Crashes at Highway-Rail Grade Crossings: Multinomial Logit Modeling. *Journal of the Transportation Research. Forum., 54*(2), 39–56. https://doi.org/10.5399/osu/jtrf.54.2.4282.
Garber, N. J., & Gadiraju, R. (1989). Factors affecting speed variance and its influence on accidents. *Transportation Research Record, 1213*, 64–71.
Geedipally, S. R., & Lord, D. (2010). Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson-gamma models. *Accident Analysis and Prevention, 42*(4), 1273–1282. https://doi.org/10.1016/j.aap.2010.02.004.
Ghasemzadeh, A., & Ahmed, M. M. (2018). Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches. *Transportation Research Part C: Emerging Technologies, 90*, 379–392. https://doi.org/10.1016/j.trc.2018.03.018.
Ghasemzadeh, A., & Ahmed, M. M. (2019). Quantifying regional heterogeneity effect on drivers' speeding behavior using SHRP2 naturalistic driving data: A multilevel modeling approach. *Transportation Research Part C: Emerging Technologies, 106*, 29–40. https://doi.org/10.1016/j.trc.2019.06.017.
Gitelman, V., Bekhor, S., Doveh, E., Pesahov, F., Carmel, R., & Morik, S. (2018). Exploring relationships between driving events identified by in-vehicle data recorders, infrastructure characteristics and road crashes. *Transportation Research Part C: Emerging Technologies, 91*, 156–175. https://doi.org/10.1016/j.trc.2018.04.003.
Haghighi, N., Liu, X. C., Zhang, G., & Porter, R. J. (2018). Impact of roadway geometric features on crash severity on rural two-lane highways. *Accident Analysis and Prevention, 111*, 34–42. https://doi.org/10.1016/j.aap.2017.11.014.
Haleem, K., & Gan, A. (2013). Effect of driver's age and side of impact on crash severity along urban freeways: A mixed logit approach. *Journal of Safety Research, 46*, 67–76. https://doi.org/10.1016/j.jsr.2013.04.002.
Han, Y., Hu, J., Cai, C., & Li, R. (2013). Numerical simulation on aerodynamic characteristics of vehicles and bridges under cross winds with the consideration of vehicle moving. *ENGINEERING MECHANICS, 30*(2), 318–325. https://doi.org/10.6052/j.issn.1000-4750.2011.07.0475.
Hao, H., Eric, Y., Medina, A., Gibbons, R. B., & Wang, L. (2020). Understanding crashes involving roadway objects with SHRP 2 naturalistic driving study data. *Journal of Safety Research, 73*, 199–209. https://doi.org/10.1016/j.jsr.2020.03.005.
Harvey, C., & Burnett, G. (2019). The influence of incentives and instructions on behaviour in driving simulator studies. *Transportation Research Part F: Traffic Psychology and Behaviour, 67*, 164–174. https://doi.org/10.1016/j.trf.2019.10.007.
He, J., Liu, X., Chen, Y., & Zhao, C. (2011). Influence of condition on running safety in atrocious weather. *Journal of Traffic and Transportation Engineering, 11*(1), 58–63.
Van Der Hoeven, F. (2011). A review of the impact of the RingRing study for multifunctional motorway tunnels in the Netherlands and beyond. *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research, 26*(2), 292–309. https://doi.org/10.1016/j.tust.2010.11.001.
Islam, M., & Mannering, F. (2020). A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. *Analytic Methods in Accident Research, 27*, 100128. https://doi.org/10.1016/j.amar.2020.100128.
Jiang, J., & Dong, L. (2012). Investigation on traffic flow characteristics around a type C weaving section based on cellular automaton model. *Chinese. Journal of Theoretical and Applied Mechanics, 44*(6), 996–1004.
Johnson, R., & Wichern, D. (1992). *Multivariate Statistical Analysis* (3rd ed.). Prentice-Hall.
Li, P., & He, J. (2016). Geometric design safety estimation based on tire-road side friction. *Transportation Research Part C: Emerging Technologies, 63*, 114–125. https://doi.org/10.1016/j.trc.2015.12.009.
Li, Z., Ci, Y., Chen, C., Zhang, G., Wu, Q., Qian, Z. (., ... Ma, D. T. (2019a). Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accident Analysis and Prevention, 124*, 219–229. https://doi.org/10.1016/j.aap.2018.12.020.
Li, Z., Wu, Q., Ci, Y., Chen, C., Chen, X., & Zhang, G. (2019b). Using latent class analysis and mixed logit model to explore risk factors on driver injury severity in single-vehicle crashes. *Accident Analysis and Prevention, 129*, 230–240. https://doi.org/10.1016/j.aap.2019.04.001.
Li, Z., Wang, W., Liu, P., Bigham, J. M., & Ragland, D. R. (2013). Using Geographically Weighted Poisson Regression for county-level crash modeling in California. *Safety Science, 58*, 89–97. https://doi.org/10.1016/j.ssci.2013.04.005.
Liu, P., & Fan, W. (2020). Exploring injury severity in head-on crashes using latent class clustering analysis and mixed logit model: A case study of North Carolina. *Accident Analysis and Prevention, 135*, 105388. https://doi.org/10.1016/j.aap.2019.105388.
Liu, Z., He, J., Zhang, C., Xing, L., & Zhou, B. (2020). The impact of road alignment characteristics on different types of traffic accidents. *Journal of Transportation Safety and Security, 12*(5), 697–726. https://doi.org/10.1080/19439962.2018.1538173.
Lyon, C., Oh, J., Persaud, B., Washington, S., & Bared, J. (2003). Empirical Investigation of Interactive Highway Safety Design Model Rural Intersections. *Transportation Research Record, 1840*, 78–86.
Mahmud, S. M. S., Ferreira, L., Hoque, M. S., & Tavassoli, A. (2019). Factors Affecting Driver Speed Choice Behaviour in a Two-Lane Two-Way Heterogeneous Traffic Environment: A Micro-level Analysis. *Journal of The Institution of Engineers (India): Series A, 100*(4), 633–647. https://doi.org/10.1007/s40030-019-00389-5.
Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22. https://doi.org/10.1016/j.amar.2013.09.001.
Medina, A. M. F., & Tarko, A. P. (2005). Speed factors on two-lane rural highways in free-flow conditions. *Transportation Research Record, 1912*, 39–46. https://doi.org/10.3141/1912-05.
Mehri, A., Sajedifar, J., Abbasi, M., Naimabadi, A., Akbar, A., Heidar, G., & Abolfazl, S. (2019). Safety evaluation of lighting at very long tunnels on the basis of visual adaptation. *Safety Science, 116*, 196–207. https://doi.org/10.1016/j.ssci.2019.03.018.
Miller, T. L., Davis, W. T., Reed, G. D., Doraiswamy, P., & Fu, J. S. (2003). Characteristics and Emissions of Heavy-Duty Vehicles in Tennessee Under the MOBILE6 Model. *Transportation Research Record, 1842*, 99–108.
Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis and Prevention, 40*(1), 260–266. https://doi.org/10.1016/j.aap.2007.06.006.
Molnar, L. J., Eby, D. W., Bogard, S. E., LeBlanc, D. J., & Zakrajsek, J. S. (2018). Using naturalistic driving data to better understand the driving exposure and patterns of older drivers. *Traffic Injury Prevention, 19*(sup1), S83–S88. https://doi.org/10.1080/15389588.2017.1379601.
Mothafer, G. I. M. A., Yamamoto, T., & Shankar, V. N. (2016). Evaluating crash type covariances and roadway geometric marginal effects using the multivariate Poisson gamma mixture model. *Analytic Methods in Accident Research, 9*, 16–26. https://doi.org/10.1016/j.amar.2015.11.001.
Osman, M., Paleti, R., Mishra, S., & Golias, M. M. (2016). Analysis of injury severity of large truck crashes in work zones. *Accident Analysis and Prevention, 97*, 261–273. https://doi.org/10.1016/j.aap.2016.10.020.
Park, E. S., & Lord, D. (2007). Multivariate poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record, 2019*(1), 1–6. https://doi.org/10.3141/2019-01.
Rudin-brown, C. M., Young, K. L., Patten, C., Lenné, M. G., & Ceci, R. (2013). Driver distraction in an unusual environment : Effects of text-messaging in tunnels. *Accident Analysis and Prevention, 50*, 122–129. https://doi.org/10.1016/j.aap.2012.04.002.
Sarrias, M. (2016). Discrete Choice Models with Random Parameters in R : The Rchoice Package. *Journal of Statistical Software, 74*(1), 1–31.
Seacrist, T., Douglas, E. C., Hannan, C., Rogers, R., Belwadi, A., & Loeb, H. (2020). Near crash characteristics among risky drivers using the SHRP2 naturalistic driving study. *Journal of Safety Research, 73*, 263–269. https://doi.org/10.1016/j.jsr.2020.03.012.
Simons-Morton, B. G., Gershon, P., O'Brien, F., Gensler, G., Klauer, S. G., Ehsani, J. P., ... Dingus, T. A. (2020). Crash rates over time among younger and older drivers in the SHRP 2 naturalistic driving study. *Journal of Safety Research, 73*, 245–251. https://doi.org/10.1016/j.jsr.2020.03.001.
Tay, R., Choi, J., Kattan, L., & Khan, A. (2011). A multinomial logit model of pedestrian-vehicle crash severity. *International Journal of Sustainable Transportation, 5*(4), 233–249. https://doi.org/10.1080/15568318.2010.497547.
Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for trannsportation data analysis (Second edi)*. Chapman and Hall/CRC.
Wei, W., Gan, H., Li, X., Zhong, S., Cao, Q., Li, W., & Li, Q. (2012). Analysis and testing the reliability, validity and resolving ability discrimination about evaluation index system of hospital technology capabilities in Sichuan Province. *Practical Journal of Clinical Medicine, 9*(3), 194–197.
Wu, K. F., & Jovanis, P. P. (2013). Defining and screening crash surrogate events using naturalistic driving data. *Accident Analysis and Prevention, 61*, 10–22. https://doi.org/10.1016/j.aap.2012.10.004.
Xing, L., He, J., Abdel-aty, M., Wu, Y., & Yuan, J. (2020). Time-varying Analysis of Tra ffi c Con fl icts at the Upstream Approach of Toll Plaza. *Accident Analysis and Prevention Journal, 141*, 1–14. https://doi.org/10.1016/j.aap.2020.105539.
Yang, L., Feng, Z., Zhao, X., Jiang, K., & Huang, Z. (2020). Analysis of the factors affecting drivers ' queue-jumping behaviors in China. *Transportation Research Part F: Psychology and Behaviour, 72*, 96–109. https://doi.org/10.1016/j.trf.2020.05.008.

Ye, X., Pendyala, R. M., Washington, S. P., Konduri, K., & Oh, J. (2009). A simultaneous equations model of crash frequency by collision type for rural intersections. *Safety Science, 47*(3), 443–452. https://doi.org/10.1016/j.ssci.2008.06.007.

Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science, 63*, 50–56. https://doi.org/10.1016/j.ssci.2013.10.012.

Zhou, Z., Ma, J., Lu, T., Li, G., Fang, S., & Tan, T. (2020). An Evaluation Method for Visual Search Stability in Urban Tunnel Entrance and Exit Sections Based on Markov Chain. *IEEE Access, 8*, 68559–68569. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2020.2986272.

Changjian Zhang is a Ph.D. Candidate in the School of Transportation, Southeast University, Nanjing China. His research interests include road traffic safety, traffic accident prevention, statistical modeling and vehicle kinetics analysis.

Jie He is a Professor in the School of Transportation, Southeast University, Nanjing China, and a Visiting Scholar at the Traffic Accident Research Center of Queensland University of Technology, Australia. His primary areas of research interests are the fields of traffic safety management and control, transportation planning and management, transportation and logistics, and traffic safety related to human and road environments. Dr. He worked as a Director of Institute of Vehicle Application Engineering and Transportation Laboratory, School of Transportation, Southeast University.

Xintong Yan is a Ph.D. Candidate in the School of Transportation, Southeast University, Nanjing China. Her research interests include road traffic safety, driver behavior and statistical modeling.

Liuzi Yang is a Ph.D. Candidate in the School of Transportation, Southeast University, Nanjing China. His research interests include traffic accidents analysis, complex node safety on urban roads and traffic safety simulation. He has applied Jiangsu Creative PHD student sponsored project.

Yikai Chen is an Associate Professor at School of Automobile and Traffic Engineering, Hefei University of Technology, Hefei China. He has served as a reviewer for several SCI and SSCI journals/conferences, such as Accident Analysis & Prevention, Traffic Injury Prevention、IEEE Intelligent Transportation Systems Transactions, Transportation Research Board, Journal of Central South University, Vehicle System Dynamics, Proc. ImechE, Part D Journal of Automobile Engineering, Mechanical Systems and Signal Processing and Strojniški vestnik - Journal of Mechanical Engineering. As a member of the Construction Machinery Professional Committee of Anhui Highway Society, his primary areas of research interests are the fields of road traffic safety, vehicle system dynamics and vehicle-road coupling and optimization.

Hao Zhang is a Ph.D. Candidate in the School of Transportation, Southeast University, Nanjing China. His research interests include optimization of highway signs and markings and road traffic safety.

Journal of Safety Research 78 (2021) 203–209

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Factors associated with concussion symptom knowledge and attitudes towards concussion care-seeking among parents of children aged 5–10 years

Juliet K. Haarbauer-Krupa [a,*], Johna K. Register-Mihalik [b,c,d], Aliza K. Nedimyer [c,d,e], Avinash Chandran [f], Melissa C. Kay [g], Paula Gildner [d], Zachary Y. Kerr [b,c,d]

[a] National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA, USA
[b] Department of Exercise and Sport Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[c] Matthew Gfeller Sport-Related Traumatic Brain Injury Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[d] Injury Prevention Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[e] Human Movement Science Curriculum, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[f] Datalys Center for Sports Injury Research and Prevention, Indianapolis, IN, USA
[g] School of Health Professions, University of Southern Mississippi, Hattiesburg, MS, USA

ARTICLE INFO

Article history:
Received 18 November 2020
Received in revised form 26 January 2021
Accepted 5 May 2021
Available online 20 May 2021

Keywords:
Concussion
Parent education
Elementary school
Injury risk
Young children

ABSTRACT

Background: Understanding parents' concussion-related knowledge and attitudes will contribute to the development of strategies that aim to improve concussion prevention and sport safety for elementary school children. This study investigated the association between parent- and child-related factors and concussion symptom knowledge and care-seeking attitudes among parents of elementary school children (aged 5–10 years). *Methods:* Four hundred parents of elementary school children completed an online questionnaire capturing parental and child characteristics; concussion symptom knowledge (25 items, range = 0–50; higher = better knowledge); and concussion care-seeking attitudes (five 7-point scale items, range = 5–35; higher = more positive attitudes). Multivariable ordinal logistic regression models identified predictors of higher score levels. Adjusted odds ratios (aOR) with 95% confidence intervals (CI) excluding 1.00 were deemed statistically significant. *Results:* Select parent and child characteristics were associated with higher score levels for both outcomes. For example, odds of better knowledge level in parents were higher with increased age (10-year increase aOR = 1.59; 95% CI = 1.10–2.28), among females (aOR = 3.90; 95% CI = 2.27–6.70), and among white/non-Hispanics (aOR = 1.79; 95%CI = 1.07–2.99). Odds of more positive concussion care-seeking attitude levels were higher among parents with a college degree (aOR = 1.98; 95%CI = 1.09–3.60). Child sports participation was not associated with higher score levels for either outcome. *Conclusions:* Certain elementary school parent characteristics were associated with parents' concussion symptom knowledge and care-seeking attitudes. While the findings suggest providing parents with culturally and demographically relevant concussion education might be helpful, they also emphasize the importance of ensuring education/prevention regardless of their children's sports participation. *Practical Applications:* Pediatric healthcare providers and elementary schools offer an optimal community-centered location to reach parents with this information within various communities.

1. Introduction

Concussion and mild traumatic brain injury (TBI) in children have received more attention in recent years as the rate of these injuries in children is increasing (Centers for Disease Control and Prevention, 2019; Haarbauer-Krupa et al., 2018; Sarmiento et al., 2019). Currently, much of the literature on concussion prevention

Abbreviations: TBI, traumatic brain injury; SSI, Survey Sampling International; IQR, interquartile ranges; aOR, adjusted odds ratio.

* Corresponding author at: Division of Injury Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, MS-S106-9, Atlanta, GA 30341, USA.

E-mail address: WUK1@cdc.gov (J.K. Haarbauer-Krupa).

https://doi.org/10.1016/j.jsr.2021.05.003
0022-4375/

and management focuses on high school age children (Nanos, Franco, Larson, Mara, & Laskowski, 2017; Sarmiento, Donnell, Bell, Tennant, & Hoffman, 2019; Suskauer et al., 2019; Wallace, Covassin, & Nogle, 2017). However, elementary school children can incur concussions from a variety of mechanisms, including sports participation (Haarbauer-Krupa et al., 2018; Master et al., 2020). When considering injury risk among younger children, especially those participating in sports, a focus on parents is essential as they often manage concussion identification and care for their children after an injury is sustained.

In consideration of sport safety and the prevention and management of concussion, it is important to assess parent concussion knowledge and attitudes as they have been shown to contribute to youth athletes' view of concussion and willingness to report a concussion (Register-Mihalik et al., 2018; Sarmiento et al., 2019). Parents' opinions towards seeking care after a concussion may be affected by factors such as socioeconomic status (Kroshus et al., 2018; Lin et al., 2015). Further, parents' beliefs about concussion reporting and seeking care can be influenced by their own concussion history (Kroshus et al., 2018), if their child has previously experienced a concussion (Kroshus, Stellino, Chrisman, & Rivara, 2017; Sarmiento et al., 2019), and if they are currently healthcare workers (Nanos et al., 2017). Beliefs and attitudes about their child's sports participation and performance also affect parents' opinions about concussion reporting and seeking medical care (Kroshus et al., 2017; Sarmiento et al., 2019). Parents may have concerns that their child will miss playing time or specialized sports achievement and these concerns can contribute to their communication with their children about reporting concussion (Kroshus et al., 2017; Sarmiento et al., 2019).

Although there is little evidence on concussion knowledge in children younger than age 10, there are reports from other domains related to risk taking that children in this age group have some understanding about their vulnerability and causality of risks (Cook, Peterson, & DiLillo, 2000; Kroshus, Gillard, Haarbauer-Krupa, Goldman, & Bickham, 2016; Morrongiello & Matheis, 2004; Morrongiello & Sedore, 2005). Children in this age group are more likely to rely on their parents for concussion education and care. The younger children are targeted with developmentally appropriate education about concussions, the less time they will be exposed to uncontested cultural messages about sport injury; however, it is important that children in this age group receive information based on rules to follow (Kroshus et al., 2016). Interventions for this stage of development about concussion prevention are likely to be different that those designed for adolescent athletes (Kroshus et al., 2016).

As such, the evidence suggests that parents play an integral role in the reporting and management of concussion in their children. While the previously mentioned research shows a robust association between parental opinions about concussion and subsequent attitudes and behaviors in their children, the bulk of this research has been conducted among parents of high school age children. Comparatively, little is known about whether this same relationship exists in elementary school age children (ages 5–10 years). Moreover, elementary school age years are a formative time for development of healthy attitudes around general well-being and safety. The aim of this study was to understand concussion symptom knowledge and concussion care-seeking attitudes among parents of children enrolled in elementary school.

2. Methods

The current study used a cross-sectional survey design. Our population of interest was parents of children enrolled in elementary school. A previous publication reported findings from this survey (for parents of middle school children) and describes the methodology in detail (Kerr et al., 2021). The study was approved by the Institutional Review Board at University of North Carolina at Chapel Hill.

2.1. Participants and recruitment

The study sample was recruited by Survey Sampling International (SSI), which used a pool of U.S. residents who agreed to participate in online survey research. These individuals provide demographic information from which SSI can identify those eligible for specific studies. SSI used certification processes such as digital fingerprinting, IP-verification, and built-in quality control questions to ensure data quality.

For the elementary aim of the larger study, SSI only targeted individuals who had self-reported as parents of children aged 5–10 years. Among this group of eligible participants, SSI randomly generated a sample that received an invitation to participate in this study. To avoid self-selection bias, specific study details were not included in the invitation; rather participants were simply invited to "take a survey," with study details provided upon accepting the invitation. Upon completion of a survey study, SSI reimbursed participants with "reward points" that can be redeemed for cash, gift cards, etc.

2.2. Data collection

Our online questionnaire was hosted on Qualtrics and was based off a modified version of a previously validated questionnaire. (Register-Mihalik et al., 2018; Kerr et al., 2021) Further, the questionnaire was refined for this study with input from injury epidemiologists, athletic trainers, sports medicine practitioners, and youth sport parents. The survey was piloted with a sample of five parents of young children and revised accordingly.

We provided the finalized survey via the URL of the online questionnaire to SSI, who integrated it into their survey platform. From September to October 2018, 475 randomly selected U.S. residents (aged ≥18 years) identifying as parents of children aged 5–10 years were invited and agreed to complete the online questionnaire. Of these 475 individuals, 400 respondents (81.2%) confirmed having children currently enrolled in elementary school at the time of responding (via a screening question within our questionnaire), completed all survey items, and were thus included in analyses. Information regarding items pertinent to the current study is provided below in results.

2.3. Measures

Our outcomes of interest were measurements of concussion symptom knowledge and care-seeking attitudes, modified for the current study population (Register-Mihalik et al., 2018; Kerr et al., 2021). These outcomes have been reported among a sample of middle school sport parents in a previous publication (Kerr et al., 2021). Table 2 includes measures on concussion symptom knowledge and concussion care-seeking attitudes. Concussion symptom knowledge included 25 symptoms (responses of yes/no/maybe). Correct answers scored 2 points; "maybe," 1 point; and incorrect answers, 0 points. The possible score range was 0 to 50, with higher scores indicating higher concussion symptom knowledge. Concussion care-seeking attitudes included five items on a 7-point scale. Items examined how a respondent would feel about seeking medical care for their children in elementary school if they had a concussion. The potential range of scores was 5 to 35, with higher scores indicating more positive attitudes toward seeking care.

Our explanatory variables of interest were parent and elementary school children characteristics. These variables were chosen because previous work in older athletes suggests potential connections between such variables and care-seeking outcomes (Register-Mihalik et al., 2013; Wallace, Covassin, Nogle, Gould, & Kovan, 2017). Parent characteristics included age (in years), sex, race/ethnicity (Register-Mihalik et al., 2013), education, concussion history, and competitiveness. The competitiveness (capturing an individual's desire to win in interpersonal situations) scale aimed to examine the extent of an individual's "desire to win in interpersonal situations" and included 20 statements on a 5-point Likert scale (Smither & Houston, 1992). The range of potential scores was 20–100, with higher scores indicating higher levels of competitiveness. Child characteristics (reported by the parent) included participation in organized sports within the past year and concussion history. Participating parents were instructed to provide characteristics only for their children currently enrolled in elementary school. If parents had multiple children in elementary school, they were asked to consider child characteristic questions collectively, and as a result we were unable to distinguish child characteristics on a child-by-child basis. Organized sports included sports played in elementary school or in youth club/recreation leagues.

Respondents whose children in elementary school played sports were asked to list all the sports that their children played from a pre-selected set (with a fill-in "other" option). We then classified sports according to contact level, based on the existing literature (Rice, 2008). Non-contact sports included: archery, cross country, dance, golf, swimming, tennis, and track and field. Limited contact sports included: baseball, fencing, flag football, racquetball, softball, and volleyball. Contact sports included: basketball, boxing, cheerleading, field hockey, gymnastics, ice hockey, lacrosse, martial arts, soccer, water polo, and wrestling. Although tackle football is classified as a contact sport, we opted to keep this as a separate category as previous research has found it to have higher concussion rates than other contact sports across multiple levels of play (Kerr et al., 2019; Kerr, Cortes, & Caswell, 2017; Rice, 2008; Tamimalam et al., 2018), If children played multiple sports, they were categorized according to the highest contact-level to which they were exposed (e.g., a child participating in ice hockey and tennis was classified in the "contact sports" category). As the count for non-contact sports was low, we merged non-contact and limited contact sports into one category.

2.4. Statistical analysis

Data were analyzed using SAS (version 9.4; SAS Institute Inc., Cary, NC). This analysis was similar to data from a middle school cohort that was part of the larger parent study (Kerr et al., 2021). Descriptive analyses were conducted for all measures of interest. For quantitative data, means and standard deviations were calculated when data followed normal distributions; medians and interquartile ranges (IQR) were calculated when data followed non-normal distributions. For categorical data, frequencies were calculated.

Multivariable ordinal logistic regression models identified predictors of higher score levels for each outcome (concussion symptom knowledge and care-seeking attitudes). Due to the discrete nature of the outcome measures, an *a priori* decision was made to categorize scores into 3 ordinal levels based on ~33% increments in the overall range of each score. Score cut-offs were selected to represent meaningful elevations in the levels of the outcomes. Thus, concussion symptom knowledge levels were 0–16, 17–33, and 34–50, while concussion care-seeking attitudes were 5–15, 16–25, and 26–35. Tests for the proportional odds assumption were conducted prior to fitting models.

In these models, parent characteristic-related adjusted odds ratios (aOR) were computed for: age (maintained as discrete variable, with the aOR examining the effect of 10-year increases); sex (female versus male); race/ethnicity (person of color versus white/non-Hispanic); education level (bachelor's degree or more versus less than a bachelor's degree); parent concussion history (yes versus no); and competitiveness (maintained as discrete variable, with the aOR examining the effect of 10% increases). Similarly, child characteristic-related aORs were computed for: concussion history (yes versus no) and sport participation (each contact level of sport participation versus no sports participation). All aORs with 95% confidence intervals (CI) excluding 1.00 were deemed statistically significant.

3. Results

3.1. Descriptive statistics

Overall, 400 parents completed the questionnaire, with most being female (70.0%), white/non-Hispanic (76.5%), without a college degree (52.3%), and with children in elementary school playing organized sports (72.0%). The mean parent age was 36 ± 8 years. The mean competitiveness score was 61 ± 11. In addition, 24.3% of parents reported a concussion history, while 15.5% of parents reported a concussion history for their children in elementary school.

We observed median scores of 39/50 (IQR = 32–44) for concussion symptom knowledge (Fig. 1) and 32/35 (IQR = 28–35) for concussion care-seeking attitudes (Fig. 2). Symptoms that were the most commonly answered correctly were: headache (79.3%), blurred vision (79.3%), confusion (77.3%), and nausea/vomiting (74.0%; Table 2). Parents were less likely to answer emotional symptoms correctly such as: sadness (30.0%), more emotional (33.3%), and feeling nervous or anxious (34.3%). For concussion care-seeking attitudes, the mean scores for each item were high, ranging from 5.62 to 6.25 (on a 7-point scale), with the lowest item mean being "extremely difficult...extremely easy" for care seeking. Data for both outcomes were skewed left, with 70.0% and 82.0% of parents in the highest-level groups for concussion symptom knowledge and care-seeking attitudes, respectively. In contrast, 4.5% and 2.8% of parents were in the lowest level groups for concussion symptom knowledge and care-seeking attitudes, respectively (Table 1).

3.2. Multivariable ordinal logistic regression models

In the multivariable model for concussion symptom knowledge, odds of greater knowledge level were higher with increased parental age (10-year increase aOR = 1.59; 95%CI = 1.10–2.28), with increased competitiveness (10% scale increase aOR = 1.25; 95% CI = 1.04–1.50), in female versus male parents (aOR = 3.90; 95% CI = 2.27–6.70), in white/non-Hispanic parents versus parents that were not white/Non-Hispanic (aOR = 1.79; 95%CI = 1.07–2.99), and in parents with a personal concussion history versus without a concussion history (aOR = 2.34; 95%CI = 1.25–4.36; Table 3). However, odds of greater knowledge level were lower in parents whose elementary school children did not have versus had a concussion history (aOR = 0.40; 95%CI = 0.21–0.78). In the multivariable model for concussion care-seeking attitudes, odds of more positive concussion care-seeking attitudes were only higher among parents with a college degree versus no college degree (aOR = 1.98; 95% CI = 1.09–3.60). In both models, higher levels of respective outcomes were not associated with whether parents' elementary school children aged 5–10 years played organized sports.

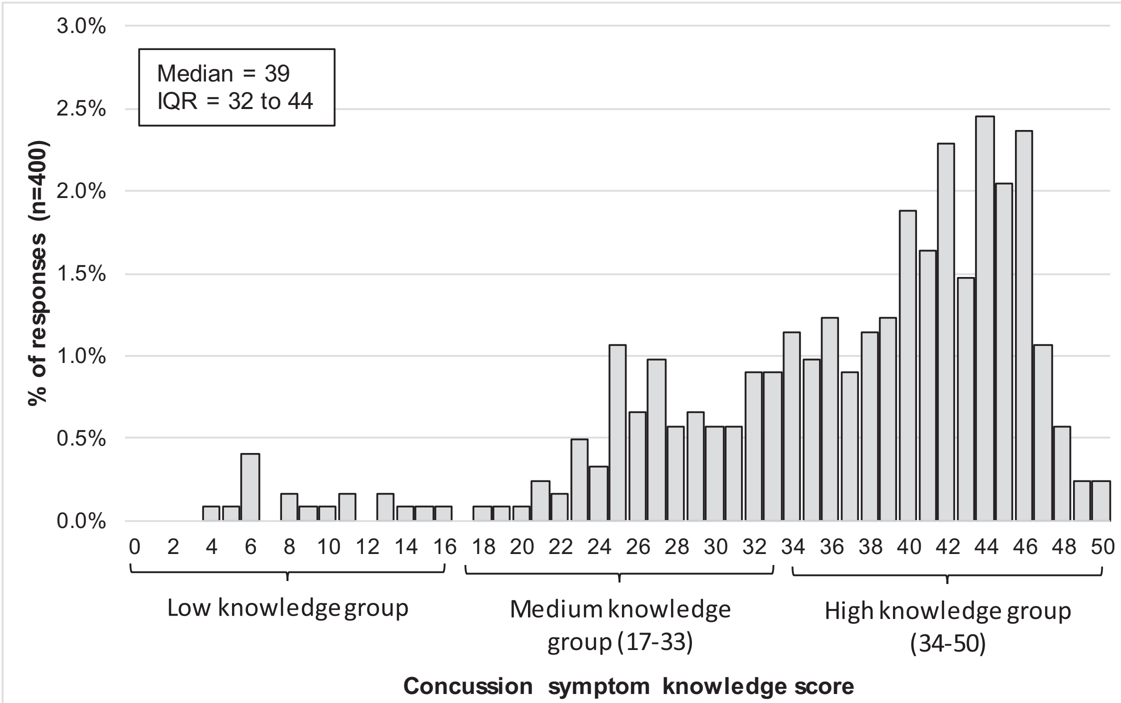

Fig. 1. Distribution of scores of concussion symptom knowledge among sample of 400 United States parents of elementary school children, September–October 2018.

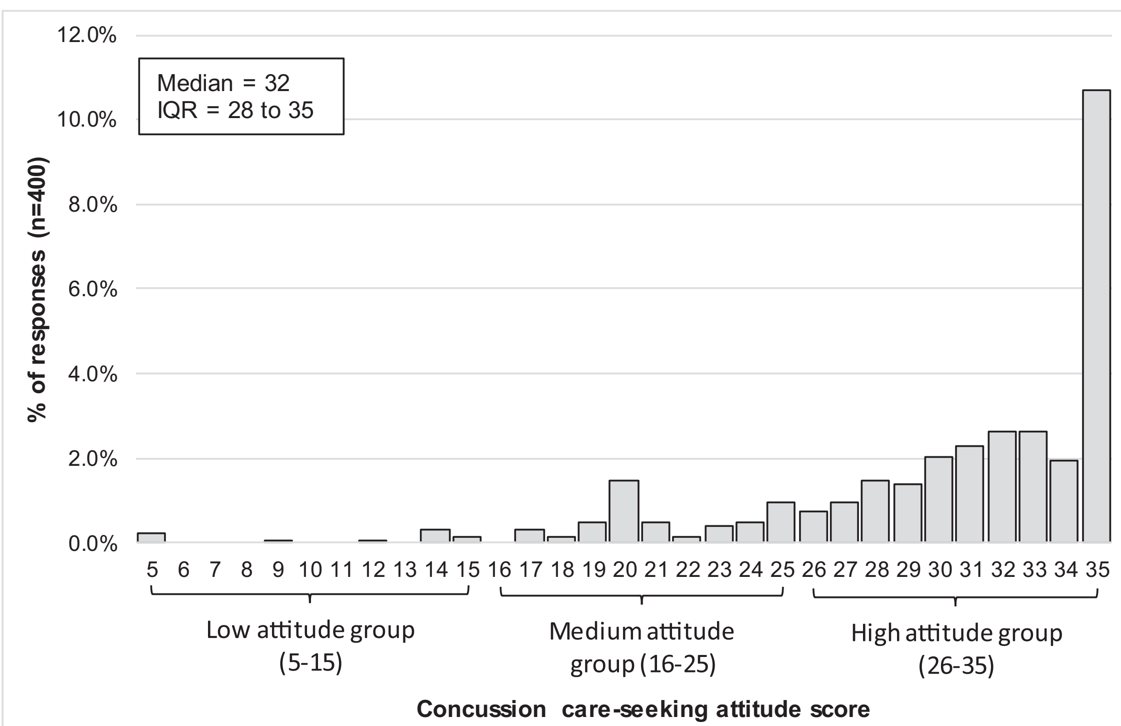

Fig. 2. Distribution of scores of concussion care-seeking attitudes among sample of 400 United States parents of elementary school children, September–October 2018.

4. Discussion

Elementary school age is a formative time for developing healthy and safe sport practices. It is important to understand how parents of elementary aged children perceive concussion safety because of their involvement with children's activities. This is one of the first studies utilizing a national sample to examine parent concussion knowledge and care seeking attitudes for elementary school children (aged 5–10 years). In combination with previous work in youth sports (Rice & Curtis Jun, 2019; Thomas et al., 2018; Waltzman & Sarmiento, 2019), these findings provide key data to inform parental education strategies to maximize health and safety for schools and other community programs that include this younger age group.

Table 1
Demographics of parents of elementary school children (n = 400).

Variable	n (%)
Parent characteristics	
Age in years	(Mean ± SD = 36 ± 8)
<30	78 (19.5)
30–39	231 (57.8)
40–49	67 (16.8)
>49	24 (6.0)
Gender	
Male	120 (30.0)
Female	280 (70.0)
Race/Ethnicity	
White/non-Hispanic	306 (76.5)
Person of color	94 (23.5)
Black/African-American	23 (5.8)
Asian/Pacific Islander	13 (3.3)
Latinx	46 (11.5)
Mixed race/other	12 (3.0)
Education	
Less than a bachelor's degree	209 (52.3)
Less than high school	9 (2.3)
High school graduate or GED	73 (18.3)
Some college; no degree	84 (21.0)
Associate's degree	43 (10.8)
Bachelor's degree and above	191 (47.8)
Bachelor's degree	116 (29.0)
Master's degree	51 (12.8)
Doctorate	18 (4.5)
Professional degree	6 (1.5)
Parent concussion history	
No	303 (75.8)
Yes	97 (24.3)
Competitiveness Index	(Mean ± SD = 61 ± 11)
20–39	17 (4.3)
40–59	164 (41.0)
60–79	198 (49.5)
80–100	21 (5.3)
Children characteristics	
Played organized sports within past year	
No sports	112 (28.0)
Yes, non-/limited contact sports	32 (8.0)
Yes, contact sports	227 (56.7)
Yes, football	29 (7.3)
Child concussion history	
No	338 (84.5)
Yes	62 (15.5)

Table 2
Concussion symptom knowledge descriptive statistics for parents of elementary school students (n = 400).

Concussion symptom knowledge (Options of Yes, Maybe, No)[a] *Question: Do you think the following are signs and symptoms of a concussion?*	Responses, n (%)		
	Yes	Maybe	No
Headache	317 (79.3)	58 (14.5)	25 (6.3)
"Pressure in the head"	273 (68.3)	98 (24.5)	29 (7.3)
Skin rash	47 (11.8)	93 (23.3)	260 (65.0)
Nausea or vomiting	296 (74.0)	81 (20.3)	23 (5.8)
Dizziness	326 (81.5)	45 (11.3)	29 (7.3)
Blurred vision	317 (79.3)	57 (14.3)	26 (6.5)
Balance problems	305 (76.3)	65 (16.3)	30 (7.5)
Sensitivity to light	243 (60.8)	119 (29.8)	38 (9.5)
Neck pain	224 (56.0)	139 (34.8)	37 (9.3)
Joint pain	89 (22.3)	185 (46.3)	126 (31.5)
Feeling slowed down	238 (59.5)	119 (29.8)	43 (10.8)
Feeling like "in a fog"	267 (66.8)	100 (25.0)	33 (8.3)
"Don't feel right"	284 (71.0)	81 (20.3)	35 (8.8)
Difficulty concentrating	282 (70.5)	87 (21.8)	31 (7.8)
Difficulty remembering	280 (70.0)	89 (22.3)	31 (7.8)
Fatigue or low energy	239 (59.8)	122 (30.5)	39 (9.8)
Confusion	309 (77.3)	63 (15.8)	28 (7.0)
Drowsiness	286 (71.5)	82 (20.5)	32 (8.0)
Sensitivity to noise	200 (50.0)	144 (36.0)	56 (14.0)
Trouble falling asleep	150 (37.5)	158 (39.5)	92 (23.0)
More emotional	134 (33.5)	182 (45.5)	84 (21.0)
Irritability	166 (41.5)	173 (43.3)	61 (15.3)
Sadness	120 (30.0)	166 (41.5)	114 (28.5)
Nervous or anxious	137 (34.3)	177 (44.3)	86 (21.5)
Teeth pain	99 (24.8)	178 (44.5)	123 (30.8)

[a] Yes = correct answer, except for "skin rash," "joint pain," and "teeth pain".

Generally, parents overall had a high level of knowledge about concussion and positive attitudes toward seeking care. However, our study noted factors associated with higher levels of each indicating the need to tailor educational efforts to specific demographics and cultural considerations. Older parents, females, white/non-Hispanic parents, those with a personal concussion history and whose children had a concussion history, and those with higher competitiveness scores displayed higher levels of concussion knowledge. Older parents potentially had more time to accrue concussion-related knowledge, however, specific reasons for this difference are unknown. As such, younger parents may benefit from additional concussion-related education. While it is also not fully known why female parents displayed higher concussion knowledge levels, this difference suggests varied education may be needed to provide foundational knowledge to both male and female parents. The findings of higher knowledge in those parents who identify as white/non-Hispanic illustrate potential disparities in concussion education. Such findings may be due to lack of availability of concussion materials in languages other than English (Krochus, Gonzalez, Chrisman, & Jimenez, 2019). Additionally, these findings highlight the need for more culturally relevant materials and diverse strategies for elementary school parents. Our competitiveness findings are novel, as the current study is one of the first to include competitiveness as a potential factor concerning concussion-related outcomes. The reasons between these associations are unknown; however, provide a foundation that varied parent characteristics, including competitive nature, may impact concussion knowledge.

Concussion knowledge in parents with a personal history of concussion was higher, implying that following a concussion parents may have more exposure to information that could improve knowledge levels. However, parental concussion knowledge was lower among those with a child with a history of concussion. This finding was unexpected. Our findings concerning concussion history differed from previous work in youth athletes (Kay, Register-Mihalik, Ford, Williams, & Valovich McLeod, 2017), where few associations were observed between parental and child experiences and concussion related knowledge and attitudes. Future studies should examine other factors that mediate parent concussion knowledge in the presence or absence of their child's experience with concussion; however, this may be tied into relationships with parents' own personal experiences.

Level of education (i.e., having at least a bachelor's degree) showed a positive relationship to care-seeking attitudes for concussion in elementary school children. This finding, as with previous findings, highlights potential disparities and access to information that may influence concussion perceptions. Future work should examine how to improve the gap in concussion perceptions that may be present among different levels of education and other factors closely associated with socioeconomic status.

In addition to understanding parent characteristics related to concussion education, elementary school aged children are able to understand about vulnerabilities for risk taking and how to follow rules that will contribute to their understanding of concussion and the importance or reporting this type of injury to their parents or a trusted adult. It is important to create developmentally appropriate materials on concussion symptoms, reporting, and prevention for this younger age group that parents can use to provide education to children.

Table 3
Concussion care-seeking attitudes descriptive statistics for parents of elementary school students (n = 400).

Concussion care-seeking attitudes (rated on a 1–7 scale with 7 being more favorable
Question: For the following, select the number closest to each word that describes how you feel about seeking medical care for your child(ren) in middle school when he/she may have a concussion.

	n (%)							Mean ± SD
	1	2	3	4	5	6	7	
Irresponsible. . .Responsible	6 (1.5)	2 (0.5)	5 (1.3)	32 (8.0)	31 (7.8)	77 (19.3)	247 (61.8)	6.25 ± 1.23
Harmful. . .Beneficial	8 (2.0)	7 (1.8)	7 (1.8)	28 (7.0)	48 (12.0)	74 (18.5)	228 (57.0)	6.09 ± 1.38
Extremely difficult. . .Extremely easy	5 (1.3)	6 (1.5)	17 (4.3)	70 (17.5)	70 (17.5)	73 (18.3)	159 (39.8)	5.62 ± 1.44
Bad. . .Good	7 (1.8)	5 (1.3)	9 (2.3)	30 (7.5)	44 (11.0)	84 (21.0)	221 (55.3)	6.09 ± 1.34
Unimportant. . .Important	6 (1.5)	6 (1.5)	8 (2.0)	33 (8.3)	32 (8.0)	74 (18.5)	241 (60.3)	6.16 ± 1.34

In summary, many of our findings align with previous work suggesting various parent and child characteristics are related to parental concussion knowledge in studies of older children (Kroshus et al., 2018; Lin et al., 2015; Sarmiento et al., 2019). Additionally, no differences in concussion knowledge were observed between parents with children who played organized sport and those who do not. Our combined findings highlight that the diversity of parent and child characteristics may also indicate the need to ensure that concussion prevention and sport safety information is tailored for specific audiences to aid dissemination among parents of elementary aged children. Furthermore, in combination with previous literature, findings illustrate the importance of devising approaches for concussion prevention and management that involve education of all elementary school parents, regardless of child sport participation. Efforts for this age group can include guidance for return to activities since children can experience concussion from a wide variety of injury mechanisms such as falls, playground injuries, in gym class and recess at school, dancing, and so forth (Haarbauer-Krupa et al., 2018). It is particularly important to begin concussion education for children at a young age to help shape future attitudes for concussion safety (Kroshus et al., 2016).

As many youth settings lack on-site access to athletic trainers or other medical professionals, children are increasingly seen in the emergency department for concussion care, especially following the passage of legislation in 50 states that requires medical clearance for return to play in organized sports (Tamimalam et al., 2018; Thomas et al., 2018). However, a recent study identified that parents may not always understand specific discharge advice or concussion symptoms and follow-up instructions provided in the healthcare setting (Thomas et al., 2018). Even for parents seeking care, they may not always know what to do to ensure safe return to activities and when to be seen for follow-up care, especially for persistent symptoms in young children. The current study findings provide the foundational considerations around parental and child factors that should be considered in educational initiatives to address such concussion-related information. Additionally, schools are important community resources to provide education and information through school health initiatives and parent groups (such as the Parent Teacher Association (PTA)). School professionals recognize the importance of care for a concussion (Romm et al., 2018) and can also offer further information to parents in their setting.

5. Limitations and future research

There are several limitations to this study. First, the questionnaire used was developed by investigators for this study and internal and external validity have not been confirmed. Responses obtained are from a cross-sectional sample that reflected respondents' knowledge and opinions at a particular time point. Although we used a nationwide sampling pool that SSI used to recruit participants for this study, we acknowledge the potential for different profiles of parents who completed the survey compared to those who did not. Further, although participant demographics varied across the sample, it is difficult to comprehend the representativeness of this sample in generalizing to the entire population. Second, parent responses to the survey may have been biased based on their level of education, which could have contributed to their understanding of the survey questions. Third, we acknowledge that additional parental-, child-, and community-level factors exist that we were not able to collect through the questionnaire used for this study. Third, race/ethnicity is often a proxy for SES and this study did not examine these areas specifically. Further prospective studies that can quantify aspects of the community (i.e., urban/rural communities, SES, and race/ethnicity), parents, and children as contributors to parent knowledge and willingness to seek care for concussions will be helpful for devising educational products and prevention efforts.

6. Conclusions

Parent demographics/characteristics were associated with concussion symptom knowledge and care-seeking attitudes for children aged 5–10 years. Such differences highlight the need for targeted strategies for parents regarding concussion prevention and management for their young children to achieve optimal health and safety. Further, the lack of associations with children's organized sports participation points to approaches offering messages to all parents and considering all causes of concussion beyond sport participation. Parent knowledge and care-seeking attitudes influence management of young children, which contributes to health and wellness in this age group.

7. Practical implications

The findings from this study offer key information to inform community safety initiatives concerning concussion among elementary school parents and their children.

8. Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Funding

This work was partly supported by awards (U01CE002885 and 1U01CE002880) from the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. The University of North Carolina Injury Prevention Research Center gratefully acknowledges the support of an Injury Control Research

Center award (R49/CE002479) from the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

References

Centers for Disease Control and Prevention (2019). *Surveillance report of traumatic brain injury-related emergency department visits, hospitalizations and deaths-United States, 2014*. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.
Cook, S., Peterson, L., & DiLillo, D. (2000). Fear and exhilaration in response to risk: An extension of a model of injury risk in a realworld context. *Behavior Therapy, 30*, 5–15.
Haarbauer-Krupa, J., Arbogast, K. B., Metzger, K. B., Al Kessler, R., Greenspan, R., Curry, A. E., Bell, J. M., DePadilla, L., Pfeiffer, M. R., Zonfrillo, M. R., & Master, C. L. (2018). Variations in mechanisms of injury for children with concussion. *Journal of Pediatrics, 197*, 241–248.
Kay, M. C., Register-Mihalik, J. K., Ford, C. B., Williams, R. M., & Valovich McLeod, T. C. (2017). Parents' and child's concussion history as predictors of parental attitudes and knowledge of concussion recognition and response. Orthopaedic Journal of Sports Medicine. 5(12): doi: 10.1177/2325967117742370.
Kerr, Z., Chandran, A., Nedimyer, A. K., Arakkal, B. S., Pierpoint, L. A., & Zuckerman, S. L. (2019). Concussion incidence and trends in 20 high school sports. *Pediatrics, 144*(5) e20192180.
Kerr, Z. Y., Cortes, N., Caswell, A. M., et al. (2017). Concussion rates in U.S. middle school athletes, 2015–2016 school year. *American Journal of Preventive Medicine, 53*(6), 914–918.
Krochus, E., Gonzalez, L. A., Chrisman, S. A., & Jimenez, N. (2019). Availability of concussion information in Spanish for parents of youth athletes. Health Promotion Practice. 20(3):372–380. doi: 10.1177/1524839918790231. Epub 2018 Aug 28.
Kroshus, E., Gillard, D., Haarbauer-Krupa, J., Goldman, R. E., & Bickham, D. S. (2016). Talking with young children about concussions: An exploratory study. *Child: Care, Health and Development, 43*(5), 758–767.
Kroshus, E., Qu, P., Chrisman, S. P. D., Schwien, C., Herring, S., & Rivara, F. (2018). Parent concern about concussion risk for their children. *Social Science & Medicine*. https://doi.org/10.1016/j.socscimed.2018.12.037.
Kroshus, E., Stellino, M. B., Chrisman, S. P. D., & Rivara, F. P. (2017). Threat, pressure, and communication about concussion safety: Implications for parent concussion education. *Health, Education and Behavior*, 1–8. https://doi.org/10.1177/10901981177115669.
Kerr ZY, Nedimyer AK, Kay MC, Chandra A, Gildner P, Byrd KH, Haarbauer-Krupa JK, Register-Mihalik JK. Factors associated with concussion symptom knowledge and attitudes towards concussion care-seeking among a national survey of parents of United States middle school children. *Journal of Sport and Health Science*, 10 (2), 2021, 113-121. https://doi.org/10.1016/j.jshs.2020.09.001.
Lin, A. C., Salzman, G. A., Bachman, S. L., Burke, R. V., Zaslow, T., Plasek, C. Z., et al. (2015). Assessment of parental knowledge and attitudes toward pediatric sports-related concussions. *Sports Health: A Multidisciplinary Approach*. https://doi.org/10.1177/1941738115571570.
Master, C. L., Curry, A. E., Pfeiffer, M. F., Metzger, K. B., Kessler, R. S., Haarbauer-Krupa, J., et al. (2020). Characteristics of concussion in school-aged children: Implications for clinical management. *Journal of Pediatrics*, 1–8.
Morrongiello, B. A., & Matheis, S. (2004). Determinants of children's risk-taking in different social–situational contexts: The role of cognitions and emotions in predicting children's decisions. *Journal of Applied Developmental Psychology, 25*, 303–326.
Morrongiello, B. A., & Sedore, L. (2005). The influence of child attributes and social–situational context on school-age children's risk taking behaviors that can lead to injury. *Journal of Applied Developmental Psychology, 26*, 347–361.
Nanos, K. N., Franco, J. M., Larson, D., Mara, K., & Laskowski, E. R. (2017). Youth sport-related concussions: Perceived and measured baseline knowledge of concussions among community coaches, athletes and parents. *Mayo Clinic Proceedings, 92*(12), 1782–1790.
Register-Mihalik, J. K., Guskiewicz, K. M., McLeod, T. C. V., Linnan, L. A., Mueller, F. O., & Marshall, S. W. (2013). Knowledge, attitude, and concussion-reporting behaviors among high school athletes: A preliminary study. *Journal of Athletic Training., 48*(5), 645–653.
Register-Mihalik, J. K., Williams, R. M., Marshall, S. W., Linnan, L. A., Mihalik, J. P., Guskiewicz, K. M., et al. (2018). Demographic, parental, and personal factors and youth athletes' concussion-related knowledge and beliefs. *Journal of Athletic Training, 53*(8), 768–775.
Rice, S. G. (2008). Medical conditions affecting sports participation. *Pediatrics, 121*(4), 841–848.
Rice, T., & Curtis, R. (2019 Jun). Parental knowledge of concussion: Evaluation of the CDC's "Heads up to parents" educational initiative. *Journal of Safety Research, 69*, 85–93. https://doi.org/10.1016/j.jsr.2019.02.007. Epub 2019 Mar 7 PMID: 31235239.
Romm, K. E., Ambegaonkar, J. P., Caswell, A. M., Parham, C., Cortes, N. E., Kerr, Z., et al. (2018). Schoolteachers' and administrators' perception of concussion management and implementation of return-to-learn guidelines. *Journal of School Health, 88*, 813–820.
Sarmiento, K., Donnell, Z., Bell, E., Tennant, B., & Hoffman, R. (2019). A qualitative study of barriers and opportunities for concussion management among parents of youth sports athletes. *Journal of Concussion, 3*, 107.
Sarmiento, K., Thomas, K. E., Daugherty, J., Waltzman, D., Haarbauer-Krupa, J., Peterson, A. B., et al. (2019). Emergency department visits for sports- and recreation-related traumatic brain injuries among children — United States, 2010–2016. *MMWR. Morbidity and Mortality Weekly Report, 68*, 237–242. https://doi.org/10.15585/mmwr.mm6810a2.
Smither, R. D., & Houston, J. M. (1992). The nature of competitiveness: The development and validation of the competitiveness index. *Educational and Psychological Measurement, 52*(2), 407–418.
Suskauer, S. J., Yeates, K. O., Sarmiento, K., Benzel, E. C., Breiding, M. J., Broom, C., et al. (2019). Strengthening the evidence base: recommendations for future research identified through the development of CDC's pediatric mild TBI guideline. *Journal of Head Trauma Rehabilitation [Internet], 34*(4), 215–223.
Tamimalam, A., Singichetti, B., Huang, H., Yi, L., Doerschun, R., Tiso, M., & Yang, J. (2018). Initial emergency department visit and follow-up care for concussions among children with Medicaid. *Journal of Pediatrics, 206*, 178–183.
Thomas, D. G., Bradley, L., Servi, A., Reilly, S., Apps, J. N., McCrea, M., et al. (2018). Parent knowledge and recall of concussion discharge instructions. *Journal of Emergency Nursing, 44*, 52–56.
Wallace, J., Covassin, T., Nogle, S., et al. (2017). Concussion knowledge and reporting behavior differences between high school athletes at urban and suburban high schools. *Journal of School Health, 87*(9), 665–674.
Wallace, J., Covassin, T., Nogle, S., Gould, D., & Kovan, J. (2017). Concussion knowledge and reporting behavior differences between high school athletes at urban and suburban high schools. *Journal of School Health, 87*, 665–674.
Waltzman, D., & Sarmiento, K. (2019). What the research says about concussion risk factors and prevention strategies for youth sports: A scoping review of six commonly played sports. Journal of Safety Research. 68:157–172. doi: 10.1016/j.jsr.2018.11.005. Epub 2018 Dec 12. PMID: 30876507; PMCID: PMC6615041.

Juliet Haarbauer-Krupa, PhD is a Senior Health Scientist on the Traumatic Brain Injury (TBI) Team in the Division of Injury Prevention (DUIP) at the National Center for Injury Prevention and Control (NCIPC). Dr. Haarbauer-Krupa has over 30 years of experience in the field and has authored multiple publications and presentations in the area of traumatic brain injury with specialties in TBI across the lifespan, rehabilitation and pediatric populations. She currently holds an adjunct position in the Department of Pediatrics, Emory University School of Medicine.

Johna K. Register-Mihalik, PhD is an associate professor in the Department of Exercise and Sport Science and the Matthew Gfeller Sport-Related Traumatic Brain Injury Research Center at UNC-Chapel Hill. Dr. Register-Mihalik also serves as a Co-Director for the STAR Heel Performance Laboratory and Core Faculty member at the Injury Prevention Research Center. Dr. Register-Mihalik's research interests include the negative consequences, prevention, education and clinical management of sport and military related traumatic brain injury (TBI). Her primary work centers on improved care, education and behavior change concerning TBI across the life span.

Aliza K. Nedimyer, MA, LAT, ATC is a doctoral student in the Human Movement Science Curriculum at the University of North Carolina at Chapel Hill. Her current research interests generally include the roles and impact of various stakeholders on injury prevention and sport safety.

Avinash Chandran, PhD, MS is the Director of the NCAA Injury Surveillance Program at the Datalys Center for Sports Injury Research and Prevention. Dr. Chandran is a quantitative epidemiologist, and his research interests are, broadly, in athlete health over the lifespan.

Melissa C. Kay, PhD is an Assistant Professor in the School of Health Professions at the University of Southern Mississippi in Hattiesburg, MS, USA. Her research specializes in utilizing mixed-methodological research approaches to examine issues of sport safety and cultural implications of injury, specifically regarding concussions.

Paula Gildner, MPH is a Project Director at the Injury Prevention Research Center at the University of North Carolina at Chapel Hill, NC, USA. She specializes in the management of studies related to injury prevention, specifically regarding traumatic brain injuries and sports-related concussions.

Zachary Y. Kerr, PhD, MPH is an Assistant Professor in the Department of Exercise and Sport Science at the University of North Carolina in Chapel Hill, NC, USA. His research specializes in the prevention of sport-related injuries across the lifespan.

Journal of Safety Research 78 (2021) 69–79

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Human resource management practices and organizational injury rates ☆

Nick Turner [a,*], Julian Barling [b], Jeremy F. Dawson [c], Connie Deng [a], Sharon K. Parker [d], Malcolm G. Patterson [c], Chris B. Stride [c]

[a] Haskayne School of Business, University of Calgary, Canada
[b] Smith School of Business, Queen's University, Canada
[c] Sheffield University Management School, University of Sheffield, UK
[d] Future of Work Institute, Curtin University, Australia

ARTICLE INFO

Article history:
Received 3 August 2020
Received in revised form 17 January 2021
Accepted 3 June 2021
Available online 16 June 2021

Keywords:
Human resource management
Injuries
Occupational safety

ABSTRACT

Introduction: This study investigated the extent to which five human resource management (HRM) practices—systematic selection, extensive training, performance appraisal, high relative compensation, and empowerment—simultaneously predicted later organizational-level injury rates. *Methods:* Specifically, the association between these HRM practices (assessed via on-site audits by independent observers) with organizational injury rates collected by a national regulatory agency one and two years later were modeled. *Results:* Results from 49 single-site UK organizations indicated that, after controlling for industry-level risk, organization size, and the other four HRM practices, only empowerment predicted lower subsequent organizational-level injury rates. *Practical Applications:* Findings from the current study have important implications for the design of HRM systems and for organizational-level policies and practices associated with better employee safety.

1. Introduction

The last three decades have seen considerable research interest in the effects of human resource management (HRM) systems on employee outcomes (e.g., Arthur, 1994; Becker & Huselid, 1998; Beijer, Peccei, van Veldhoven, & Paauwe, 2021; Boon, den Hartog, & Lepak, 2019; Delery & Doty, 1996; Huselid, 1995; MacDuffie, 1995; Toh, Morgeson, & Campion, 2008; West, Guthrie, Dawson, Borrill, & Carter, 2006; Youndt, Snell, Dean, & Lepak, 1996). A range of labels, such as 'high involvement management' (e.g., Forth & Millward, 2004), 'high commitment management' (e.g., Wood & de Menezes, 1998), and 'high performance work systems' (e.g., Huselid, 1995; Liao, Toya, Lepak, & Hong, 2009) have been used to describe various sets of organizational practices that aim to involve employees, generate employee commitment towards their work and the organization, and ultimately improve organizational performance.

Organizational practices that comprise HRM systems are "the specific methods and procedures that the organization adopts to implement the organization's principles and policies" (Posthuma, Campion, Masimova, & Campion, 2013, p. 1189). HRM systems comprise 'bundles' of these organizational practices that have complementary effects (Ogbonnaya, Daniels, Tregaskis, & Van Veldhoven, 2013), with each bundle of practices preferably "creating synergistic effects in which certain practices reinforce one another to increase organizational efficiency and effectiveness" (Posthuma et al., 2013, 1185). Many studies have focused on how these HRM systems are measured and how they affect performance (for reviews, see Boon et al., 2019; Godard, 2004; Wall & Wood, 2005; Wright, Gardner, Moynihan, & Allen, 2005). Most of these studies tend to concentrate on conventional financial and

☆ We gratefully acknowledge the Centre for Economic Performance at the London School of Economics, the Operations Unit at the Health and Safety Executive, and the Office of National Statistics for providing the data used in this study. Thanks to Terry Beehr, Louise Geraghty, Sarah Jane Tennant, Toby D. Wall, Michael A. West, Mo Wang, and Stephen J. Wood for their assistance and feedback at earlier stages of the project. We presented earlier versions of this paper at the 2006 SafetyNet/Canadian Association for Research on Work and Health conference, St. John's, Newfoundland, Canada, and the 2010 annual meetings of the Administrative Sciences Association of Canada, Regina, Saskatchewan, Canada. Financial support from the Canadian Centre for Advanced Leadership in Business (Haskayne School of Business), the Centre for Corporate Sustainability (Haskayne School of Business), ESRC Centre for Organization and Innovation, the Health and Safety Executive, the HROD-Strategy Research Grants program (Haskayne School of Business), the Industrial Accident and Prevention Association (Canada), and the Social Sciences and Humanities Research Council of Canada all helped to make this collaboration possible.
* Corresponding author at: Haskayne School of Business, University of Calgary, Calgary, Alberta T2N 1N4, Canada.
E-mail address: nicholas.turner@ucalgary.ca (N. Turner).

https://doi.org/10.1016/j.jsr.2021.06.003
0022-4375/

labor performance indices, to the neglect of more employee-centered criteria such as occupational health and safety (Delery & Gupta, 2016; Godard, 2004; Shaw & Delery, 2003). Yet, meta-analytic evidence suggests that several mechanisms by which HRM systems are posited to affect outcomes, such as by boosting employee engagement and organizational commitment, might also affect employee-centered criteria such as workplace injuries (Harter, Schmidt, & Hayes, 2002). Despite decades of HRM-performance research, we know relatively little about how HRM systems affect workplace injuries at the organizational level (Granger, Turner, & Grocutt, in press; Ogbonnaya et al., 2013; Turner & Dueck, 2015; Zacharatos & Barling, 2004), with greater understanding of the organizational-level determinants of workplace injuries needed more generally.

Drawing from reviews (e.g., Posthuma et al., 2013) of the prevalence of HRM practices in organizations, we identify five key HRM practices warranting investigation with organizational injuries: (1) systematic selection, (2) extensive training, (3) performance appraisal, (4) high relative compensation, and (5) empowerment. The current research examines the relationship among these key HRM practices and workplace injury rates at the organizational level of analysis. In doing so, we extend previous research by simultaneously assessing the association among these five HRM practices and subsequent organizational injury rates, thereby delineating organizational-level determinants of occupational safety. Furthermore, we respond to calls (e.g., Wright & Ulrich, 2017) in the recent literature for more rigorous prospective research designs, such as collecting multi-source data when assessing HRM-outcome linkages.

2. Theoretical background and hypotheses

After being virtually ignored within the organizational literature for many years (Barling & Frone, 2004; Campbell, Daft, & Hulin, 1982; Hofmann & Tetrick, 2003), workplace safety is increasingly the focus of theoretical and empirical attention at multiple levels of analysis (Clarke et al., 2016; Nunez & Prieto, 2019). Contemporary research draws on earlier studies that have examined the role of antecedents—such as high-quality leadership (e.g., Barling, Loughlin, & Kelloway, 2002), work design (e.g., Parker, Axtell, & Turner, 2001), job insecurity (e.g., Probst, 2004), and safety climate (e.g., Zohar, 2002)—of occupational safety-related outcomes such as workplace injuries. Such research is important given the worldwide rates of workplace injuries (Takala, 2019), with recent global estimates of lost-time injury rates as high as 11,096 per 100,000 persons in the workforce (Hämäläinen, Takala, & Tan, 2017).

Knowledge of the organizational-level determinants of safety, however, remains limited. Current evidence suggests a potential role for HRM practices designed to "enhance employee competencies, commitment, and productivity" (i.e., high performance work systems; Posthuma et al., 2013, p. 1843), but very little of this research has incorporated an organizational level of analysis. For example, Zacharatos, Barling, and Iverson (2005) showed that individuals' perceptions of high performance work systems were positively related to personal safety orientations and negatively related to occupational injuries, demonstrating how trust in management and perceptions of safety climate served as mechanisms by which HRM systems may exert effects. Additionally, Wallace, Popp, and Mondore (2006) examined the foundational climates (i.e., organizational support, management-employee relations) generated by HRM practices, demonstrating a positive relationship with workgroup safety climate and a negative relationship with workplace injuries. At the unit level, Lauver and Trank (2012) showed that organizations with higher levels of organizational decentralization and alignment of HR practices were less likely to suffer workplace injuries (as measured by regulator-collected logs of workplace incidents). Similarly, Newnam, Warmerdam, Sheppard, Griffin, and Stevenson (2017) showed in a sample of 83 organizations that high-performance work practices, particularly selection and work design, negatively related to work-related driving behaviors, but this effect was attenuated when upper management demonstrated commitment to safety.

While this evidence suggests that using practices that make up HRM systems is likely to reduce workplace injuries under some conditions, those data were collected at the individual-, workgroup-, and the unit-level (as reported by an upper-level manager in the organization) of analyses, and it cannot be assumed that the findings will apply at the organizational level (Chan, 1998). As scholars have argued, different processes might operate at different levels (Fulmer & Ostroff, 2016). Moreover, from the overall HRM constructs used in these aforementioned studies, it is not always clear which particular sets of practices might be negatively associated with workplace injury rates. It is important to tease out which or how many practices are important, in part to give insight into the mechanisms that might explain HRM-injury linkages, but also to provide practical guidance to organizations.

The aim of the study is to investigate the relationship between HRM systems and organizational injury rates. A further ambition is to examine multiple HRM practices concurrently, since research has often considered practices in isolation. For each of five HRM practice—systematic selection, extensive training, high relative compensation, performance appraisal, and empowerment—we generate hypotheses about their potential association with occupational injuries. We focused on these five practices because they characterize key elements of high performance work organization both historically (e.g., Huselid, 1995; Pfeffer, 1998) and more recently (e.g., Boon et al., 2019).

2.1. A taxonomy of human resource management practices

Posthuma et al. (2013) developed a taxonomy of high performance work practices based on a comprehensive review of the HRM-performance literature, examining the frequency with which HRM practices appear among peer-reviewed articles published between 1992 and 2011. They classified 61 specific HRM practices into nine categories and further synthesized them into five categories: recruiting and selection, training and development, compensation and benefits, performance management and appraisal, and job and work design. They then divided the practices within each category into three categories: core practices, practices that are frequent in the literature, and practices either maintaining or steadily growing within the literature. As an example, empowerment is central to the job and work design category, in that the core practice within job and work design includes decentralized and participative decisions. The five practices operationalized in the current study share similarities with Posthuma et al.'s taxonomy categories, reflecting the prevalence of HRM practices studied in the high performance work systems literature.

The current paper focuses on general HRM practices rather than criterion-focused practices (i.e., safety-specific HRM practices such safety training). General HRM practices focus on the specific methods and procedures that organizations adopt (Posthuma et al., 2013), rather than a specific criterion. For example, organizations may use selection practices to screen candidates (Posthuma et al., 2013), and a variety of selection tools to gather more information about a candidate (Youndt & Snell, 2004). Selection practices may include high-quality tools such as structured interviews (Posthuma et al., 2013), but may not necessarily focus on a criterion—for example, selecting for safety-specific competencies—but rather enable the organization to select individuals on a wide range of competencies. These general HRM practices reflect ways of improving employee capabilities, commitment, and productivity, which are also likely to have an impact on safety. In the following sections, we describe the conceptual reasons why each of these general HRM practices—and the set of them—may enhance organizational safety without a particular focus on the safety criterion.

2.1.1. Systematic selection practices and occupational safety

We propose that organizations with systematic selection practices will have lower levels of occupational injuries. Systematic selection practices involve organizations deciding in advance what the critical skills and attributes for success are in the organization, taking applicants through a systematic selection process, and, in their hiring decisions, focusing on skills, attitudes, and behaviors that are less amenable to change through training. The net effect of systematic selection processes should be organizational members that have a skill set consistent with job requirements and organizational aspirations, and therefore a reduction in on-the-job injuries. In addition, from a symbolic perspective, systematic selection processes signal to both current and future employees that management is committed to selecting the best possible organizational members (Pfeffer, 1998), with members wanting to reciprocate this commitment by doing their best work for the organization.

We propose that systematic selection strategies employed by organizations promote skill matching and facilitate skill development in their workforces, resulting in a link between systematic selection and lower organizational-level injury rates. Specifically, when organizations focus on matching skills with the requirements of the job, the selected workforce have the required competencies and experience to enable them to correctly carry out their work, and are thus more capable of completing tasks safely. Further, we expect that systematic selection systems will be related to lower organizational-level injury rates through other higher-level processes. For example, these systems may be used to create a highly skilled workforce (Takeuchi et al., 2007), which may promote more effective co-ordination within and across units, thereby enhancing safety at the organizational level. From an empirical perspective, there is evidence that organizations that use systematic selection procedures typically experience lower injury rates (e.g., Cohen, 1977; Smith, Cohen, Cohen, & Cleveland, 1978), although these studies do not provide evidence for the reasons why and reflect cross-sectional relationships.

Hypothesis 1: Systematic selection procedures will be negatively associated with injury rates.

2.1.2. Extensive training and occupational safety

Training provides employees with the opportunities to learn the competencies required for a given role (Posthuma et al., 2013). The degree to which extensive training is provided involves the intensity (e.g., duration) and the scope of training (e.g., breadth of training provided; Youndt & Snell, 2004). Extensive training may involve the amount of time spent training (e.g., Tharenou, Saks, & Moore, 2007), frequency and variety of training provided (e.g., Gong, Law, Chang, & Xin, 2009), and formalization of training programs (e.g., Delery & Doty, 1996). As such, we suggest that extensive training offered within an organization will influence occupational safety for several reasons. First, general workplace training can increase employees' problem-solving skills (Osterman, 1995) and commitment to the organization (Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991). Similarly, training for teams increases communication and information sharing (Baker, Day, & Salas, 2006). These skills may be useful for improving occupational safety: problem-solving skills are used to identify and find solutions to safety issues or communication skills may be used to clearly describe safety issues to other organizational members. Indeed, knowledge levels (e.g., Smith-Crowe, Burke, & Landis, 2003), commitment (e.g., Parker et al., 2001), and communication (Parker et al., 2001) have all been shown to be positively associated with safety outcomes.

Second, organizations that choose to introduce extensive training, beyond the mandatory training that is required by governments and unions, enhance the likelihood that employees have all the skills and knowledge needed to perform their job safely. By providing training that goes beyond the bare minimum also signals high commitment to employees, which we would expect employees to want to reciprocate through working safely. Kaminski (2001) finding that, amongst small manufacturing organizations, those offering more training hours were more likely to report lower lost-time injuries, is consistent with this explanation. More recently, Camuffo, De Stefano, and Paolino (2017) conducted a single firm, multi-plant study finding fewer lost-time injuries on average in units where front-line managers focused on developing subordinates' capabilities and skills through teaching and coaching.

Hypothesis 2: Extensive training will be negatively associated with injury rates.

2.1.3. Performance appraisals and occupational safety

Performance appraisal remains an integral part of HRM systems (Daley, Vasu, & Weinstein, 2002), and a critical component of performance management (DeNisi & Smith, 2014). One purpose of performance management is to focus on employee development, and the information gathered from such appraisals can be used to document performance and decisions concerning pay and promotion (DeNisi & Smith, 2014). High-quality performance management and performance appraisals generally include appraisals for development, appraisals based on objective results and behaviors, as well as frequent performance appraisal meetings (Posthuma et al., 2013). To our knowledge, there is an absence of research assessing the relationship between performance appraisal and occupational safety outcomes. However, the core components of performance appraisal—information sharing and feedback—suggest that an association between high-quality performance appraisal and occupational safety could exist, for several reasons.

First, feedback from performance appraisals can be used to identify employees' training needs (London & Smither, 2002), and as such lead to increases in the skills and behaviors that positively correlate with safety outcomes. Further, more frequent performance appraisal meetings can provide employees with the opportunity to review goals and adjust their training and developmental needs accordingly. Second, feedback from performance appraisals can enable learning from errors and near misses, both of which serve to enable safety improvement in the future (Littlejohn, Lukic, & Margaryan, 2014). Third, high-quality performance appraisals might also help to generate norms about the importance of information sharing and feedback, which in turn are likely to enhance organizational-level outcomes (Murphy & Cleveland, 1995). Further, organizations focusing more on seeking information and providing feedback, particularly with respect to safety incidents, can open up opportunities for learning (Dekker & Breakey, 2016), and may encourage important safety behavior such as speaking up. From a safety perspective, Cohen (1977), Smith et al. (1978), and Wallace et al. (2006) present evidence that more feedback between management and employees predicted lower injury rates. As a result, we hypothesize:

Hypothesis 3: Performance appraisals will be negatively associated with injury rates.

2.1.4. High relative compensation and occupational safety

We propose that high relative compensation in an organization—that is, higher pay relative to market norms—will be associated with lower injury rates. Compensation has consistently been considered an integral part of HRM systems (Pfeffer, 1998), with competitive pay, incentive compensation, and pay for performance as some of the core components of compensation practices generally found in a HRM systems (Posthuma et al., 2013). Research relating pay to safety has focused on performance incentives and intra-organizational pay dispersion. Consistent with the possible negative effects of performance-contingent pay (Dahl & Pierce, 2020; Parker, Bell, Gagné, Carey, & Hilpert, 2019), the existence of performance-based incentives was positively associated with injury rates in a sample of 86 manufacturing companies (Kaminski, 2001). Similarly, pay dispersion is negatively associated with satisfaction (Pfeffer & Langton, 1993), individual and team performance (Bloom, 1999), and positively related to turnover (Bloom & Michel, 2002) - presumably because it encourages employees to focus more intensively on relative individual worth (Pfeffer, 1998) and heightens perceptions of unfairness. For example, Shaw, Gupta, and Delery (2002) found that pay dispersion based on individual incentives for performance was a positive predictor of lost-time injuries in a sample of concrete production plants, over-and-above the effects of either pay dispersion or individual incentives.

In the same way that compensation fairness is an issue among employees within the same organization, we argue that employees in organizations who are paid above-market compensation relative to pay offered by similar organizations will perceive their work situation to be more than fair, and therefore exert more effort towards working safely. Werner, Kuiate, Noland, and Francia (2016) investigated the effect of supplemental retirement plans and safety behavior in the U.S. trucking industry, suggesting that as a part of a high performance work system, supplemental retirement plans act as a form of pay-above-market strategy. Offering supplemental retirement plans was negatively associated with driver insurance costs, indirectly indicating safer driver behavior through lower accidents, crashes, and driving violations. Pay-above-market strategies may also include additional benefits to employees that may have a positive influence on safety. For example, Weahrer, Miller, Hendrie, and Galvin (2016) investigated workplace injury rates across different sized organizations and industries finding that employee assistance programs (EAP) were negatively associated with workplace injury rates, particularly when EAP employees are on-site and when organizations offered telephone EAP services. Taken together, the symbolic advantage of paying employees above market rate implies a commitment-oriented approach to HRM, in which employees are valued and which previous research suggests is positively related to commitment and organizational performance (Tsui, Pearce, Porter, & Tripoli, 1997) and safety (Barling & Hutchinson, 2000).

Hypothesis 4: High relative compensation will be negatively associated with injury rates.

2.1.5. Empowerment and occupational safety

Central aspects of structural empowerment involve autonomy and participation (Seibert, Wang, & Courtright, 2011). Within HRM systems, structural empowerment practices involve the methods and procedures that enhance employees' opportunity to participate in decision-making, as well as employees' opportunity to exercise their discretion and use their skills (Posthuma et al., 2013). As such, these practices emphasize enhancing employees' opportunity to contribute and perform (Lepak, Liao, Chung, & Harden, 2006). Organizations implementing empowerment practices may seek to increase autonomous work, that is, work is designed to have employees participate in decision-making (e.g., self-managing teams, quality circles; Wall, Wood, & Leach, 2004). Of all HRM practices, research on the relationship between autonomous work and safety has received comparatively more research coverage than other HRM-safety links. Theoretically, enhancing autonomy and participation will reduce injuries for several reasons.

First, from a socio-technical systems perspective, when employees' jobs are designed in a way that maximizes job control and responsibility, they are able to manage the variances (i.e., changes in job demands) more quickly, encouraging a broader role orientation towards safety (Turner, Chmiel, & Walls, 2005) and potentially preventing injuries. Second, autonomy promotes learning (Wall, Jackson, & Davids, 1992) and the development of greater expertise (Wall et al., 2004), which again likely leads to safer working. Third, autonomy fosters intrinsic motivation and commitment (Parker, 2014), which should increase employees' motivation to work safely. Last, empowerment practices may signal to employees that speaking up and sharing constructive ideas intended to invoke positive organizational change or improvement are encouraged and valued by the organization (i.e., voice; Chamberlin, Newton, & LePine, 2018). Extending this to a safety perspective, empowerment practices may also signal to employees that speaking up about safety-related concerns (i.e., safety voice; Tucker & Turner, 2015) is encouraged, which in turn, may promote safer working conditions.

Consistent with these arguments, a number of studies and reviews at the individual- (e.g., Barling, Kelloway, & Iverson, 2003; Nahrgang, Morgeson, & Hofmann, 2011; Parker et al., 2001) and group-level of analyses (e.g., Hechanova-Alamay & Beehr, 2001; Simard & Marchand, 1997; Turner & Parker, 2004) show the benefits of more autonomous work on occupational safety. At the organizational level of analysis, existing research is less abundant and less systematic, although consistent with findings at lower levels of analysis. For example, Shannon et al. (1996) found that managers of companies with lower lost-time injury compensation claim rates were more likely to perceive employee involvement in organizational decision-making, and have a greater expectation that employees use their own initiative. Similarly, Yassi et al. (2004) found that hospital facilities offering staff greater discretion in conducting their work had, on average, lower staff injury rates than those facilities limiting staff discretion; Arocena, Nunez, and Villanueva (2008) showed same-year negative correlations between organizational-level empowerment and lost-time work injury rates in a sample of Spanish organizations; and Camuffo et al. (2017) demonstrated the negative association between empowerment and lost-time injury rates.

Hypothesis 5: Employee empowerment will be negatively associated with injury rates.

3. Present study

We hypothesize that the presence of each of the HRM practices will uniquely predict organizational injury rates. It is important to note that, while there is evidence for some of the practices when considered in isolation, no organizational-level study has considered the effect of multiple related HRM practices on occupational safety at the same time. Our approach provides a more comprehensive perspective on the association of HRM practices comprising HRM systems and occupational injuries for both conceptual and statistical reasons. Conceptually, HRM practices do not occur in isolation from one another. Statistically, examining these practices in isolation might well exaggerate their apparent effectiveness. We therefore test a model of the simultaneous associations of five HRM practices on organizational injury rates, seeking to understand if each practice makes a unique contribution.

To test this model, we used a prospective, multi-method approach with several advantages over existing research. First, previous research on HRM has been criticized for its sole reliance on singular and often untrained sources providing the data on the use of practices (Wall & Wood, 2005). Our assessment of HRM practices in each organization derives from a team of trained observers who were aware of the possible range of use of the separate practices, but unaware of the study hypotheses. This enabled consistent and informed ratings of the effectiveness of the five separate practices. Second, data on the five practices used in this study derive from multiple sources: interviews with managers and employees, site inspections, and written documentation. Multiple ratings can result in a more reliable composite (Horowitz, Inouye, & Siegelman, 1979), and avoids the potential threat of mono-source biases. Third, the dependent variable was collected by the Health and Safety Executive in the United Kingdom, a regulatory body that is involved in inspecting and collecting data on organizational safety performance. Fourth, we used a sample of single-site companies to ensure that data on HRM practices pertain to that site, and the injury data could not be confused with that of another site of a multi-site organization. One of the challenges of conducting research on the relationship between organizational practices and variables such as organizational injury rates is "to have reliable and compatible data" (Askenazy, 2001, p. 493) on both sides of the equation. This study meets that criterion.

Fifth, the relationships tested here are predictive insofar as the HRM practice data precede in time the injury data. Existing research (e.g., Shannon et al., 1996; Kaminski, 2001) on organizational practices and safety conducted at the organizational level of analysis has been based on data collected at the same time. Our prospective design is an improvement over this approach because we collected data on the dependent variable subsequent to the independent variable. Finally, we implement necessary statistical controls to reflect industry differences in risk and organization size. Taken together, these methodological features heighten the extent to which strong inferences can be drawn.

4. Method

4.1. Sample and data collection

We collected data from 58 single-site manufacturing companies throughout the United Kingdom, as part of a wider study on organizational practices, employee attitudes, and economic performance.[1] Organizations reflected a number of sectors: mechanical engineering ($n = 21$), plastics and rubber manufacturing ($n = 20$), electronics ($n = 3$), and other miscellaneous sectors ($n = 14$). These sectors were chosen because they were the most populous in terms of number of firms, and number of employees, in the United Kingdom. Company size ranged from 50 to 900 employees, reflecting small- and medium-sized enterprises.

To assess the use of HRM practices, a team of researchers conducted a three-stage audit of each company, drawing on a range of sources of information. First, detailed structured interviews with senior managers responsible for each practice were conducted on site. The total time spent interviewing in each company was approximately three hours, with an average of three different managers. Second, the audit team toured the facilities and interviewed shop-floor employees, enabling them to observe the practices-in-use (rather than the espoused practices) and hear opinions from the workforce directly affected by these practices. Third, the research team reviewed written documentation (e.g., training schedules, quality documents) related to the practices. Taking all this information together with the comparative experience of auditing the other companies in the sample, the audit team then made a series of ratings of the sophistication of each of these practices. We provide more detail about the ratings in the Measures section.

A key criterion for selecting these companies was the fact that they were single-site organizations. This has two benefits for the present study. First, the interviews with site managers focused specifically on the HRM practices at that site, rather than the use of these practices across multiple sites. This meant that respondents provided answers about the site they knew best, and the subsequent rating of the site practices by the audit team was based on information provided during the in-depth interviews and documents pertaining to practices in that site only. Second, obtaining injury data on this type of organization from the Health and Safety Executive (HSE) archives minimized the possibility of confusion with another site of the same organization. This way, we ensured that the level of analysis used to measure the practices corresponded directly to the workplace injury data (Askenazy, 2001).

We were able to obtain data on the number of injuries for one and two years following the practice audit for 49 of these 58 single-site organizations. Of these 49 companies, 18 were in the engineering sector, 18 in rubber/plastics, 3 in electronics, and 10 in other miscellaneous manufacturing areas. All 49 were small or medium sized companies, ranging from 63 to 900 employees (*M* employees = 174, *Mdn* employees = 126).

4.2. Measures

Five HRM variables were assessed, namely systematic selection, extensive training, performance appraisal, high relative compensation, and empowerment. Systematic selection, extensive training, and performance appraisal were derived from interviewer ratings, whereas high relative compensation and empowerment were formed directly from responses given by interviewees. In 27 of the 49 organizations, there were two interviewers who rated the HRM practices separately, allowing inter-rater reliability [ICC (2, k); Shrout & Fleiss, 1979] to be established.

4.2.1. Systematic selection

The interview included detailed questions about what selection methods were used for each staff type (i.e., shopfloor, clerical/administrative staff, professional/technical staff, and management), and which of the 10 selection procedures (ranging from application forms to assessment centers) were used for each staff type. After assessing answers to all the previous questions on selection, interviewers then rated the overall approach to selection used by the company for each of the four staff types, on a scale ranging from 1 = "Non-existent" to 5 = "Excellent with careful planning." These four ratings then formed a scale, with Cronbach's alpha = 0.88. The ICC(2, k) was 0.92.

4.2.2. Extensive training

A large number of open and closed-ended questions were asked about training in the organization. These included whether: (a) there was an overall training strategy (if there was, the documentation was requested); (b) the average annual hours of formal training for a typical employee of each staff type; (c) a series of questions about Investors in People™ status (a sought-after stan-

[1] Previous papers resulting from these data include: Neal et al. (2005); Patterson, Warr, and West (2004); Patterson et al. (2005); Patterson, West, and Wall (2004); Shipton et al. (2002); Shipton et al. (2006a); and Shipton et al. (2006b).

dard, awarded to UK organizations that meet a series of criteria relating to the management and development of their staff); (d) questions about systems for assessing training needs; and (e) general questions about the type of training that occurred. Interviewers rated the extent of training for shopfloor employees, supervisors and management, on a scale ranging from 1 = "Very limited" to 5 = "Very extensive." These formed a scale, with Cronbach's alpha = 0.91. Inter-rater reliability as measured by ICC(2, k) was 0.97.

4.2.3. Performance appraisal

The interview included questions on whether there was a formal appraisal system, and if so: (a) how long it had been in operation for each of the four types of staff; (b) whether and how often these types of staff were appraised; (c) whether the appraisal was linked to remuneration; and (d) whether appraisers received any formal training. Interviewees were also asked a series of open questions about the appraisal scheme, allowing them to describe the details of the scheme more fully. Interviewers then rated the sophistication of the scheme for each of the four types of staff, on a scale ranging from 1 = "Nonexistent" to 5 = "Highly sophisticated." These formed a scale, with Cronbach's alpha = 0.96. Inter-rater reliability measured by ICC (2, k) was unity.

4.2.4. High relative compensation

Interviewees were asked how compensation for shopfloor staff, supervisors, and management compared with local companies or competitors' rates. Responses were given on a scale ranging from 1 = "Well below average" to 5 = "Well above average." These were added together to form a scale (Cronbach's alpha = 0.74), but weighted so that shopfloor employees' pay counted for four times as much as the other groups. This reflected the approximate number of each type of staff in the organizations used in the final sample.

4.2.5. Empowerment

Interviewees were asked to what extent shopfloor operators were responsible for or involved in eight tasks: a significant quality problem, material supply problem, machine repair following minor breakdown, routine maintenance of machines, setting up machines for changeover of product, setting up machines for a new product, when to take breaks, and the order in which they do their work (Wall, Jackson, & Mullarkey, 1995). Responses were given on a scale from 1 = "Not at all" to 4 = "Very much." Cronbach's alpha was 0.75.

4.2.6. Workplace injuries

Data on the number of injuries reported at each company were collected from the UK Health and Safety Executive, a government body responsible for overseeing safety in the workplace. Data were collected for both the year following the interviews and the subsequent year. Injuries were classed as fatal, major, or minor. However, there were no fatal and very few major injuries (14% of all injuries). The total number of injuries reported for the two years combined across 49 companies was 252, ranging from 0 (in 22 companies in year 1, and 20 companies in year 2) to 20 (in one company in year 2), and with the majority falling in between. As such, we used the total number of injuries across both years for each company.

4.2.7. Control variables

In the analyses, we included data on organization size and industry-level average injury rate as calculated by Office of National Statistics for companies in the Standard Industry Code to which each organization in the current sample belonged.

5. Results

5.1. Analytic strategy and descriptive findings

The means (or medians), standard deviations (or interquartile ranges), and Spearman's Rank (i.e., non-parametric) correlations of all study variables appear in Table 1. Analysis of the dependent variable (number of injuries) reveals that its distribution is severely non-normal, being positively skewed with its peak and lower limit at zero, as is typical for counts of rare events. A goodness-of-fit test showed that the data differed significantly from such a standard Poisson distribution (in which the mean is equal to the variance), being over-dispersed (i.e., with higher variance, and hence a longer tail), and hence more similar to a negative binomial distribution. This concurs with McCullagh and Nelder (1989, p. 199), who suggest that the number of incidents in an organization may be the sum of individual Poisson variables, forming a negative binomial distribution. Consequently, we chose to analyze the injury data by fitting a negative binomial regression model with a logarithmic link function (i.e., transformation of the dependent variable). Given that organizations had differing numbers of employees, it was appropriate to model injuries per employee as opposed to total injuries: as such, we included the logarithm of the number of employees in each organization as an offset term in our model, therefore effectively modelling injury rate per employee. We also controlled for industrial sector by entering the sector-average injury rate. The analysis was conducted in SPSS.

5.2. Hypothesis testing

The hypotheses address the collective and separate effects of each of five HRM practices on injury rates. The models presented in Table 2 represent the following pattern: Model 1 is a baseline model controlling for the log of organization size and industry sector injury rate, and Model 2 includes the HRM practices as an omnibus test of the hypotheses.

As a block, the HRM practices added significant explanatory power to the baseline model, $\Delta\chi^2$ (5, $N = 49$) = 73.05, $p < 0.001$. This suggests that there is an overall effect of the HRM variables on organizational injury rate. However, the only HRM practice to have a significant unique effect on injury rates is empowerment. The coefficient of -0.78 (see Table 2) is equivalent to an incidence rate ratio of $\exp(-0.78) = 0.46$: an increase of one point on the empowerment scale is associated with a reduction in the injury rate by a factor of 0.46, or a 54% reduction, all else being equal.

In the above analysis, all five practices were entered together as a set. Supplementary analyses, in which the HRM variables were entered individually into separate models, resulted in the same conclusion (i.e., empowerment was the only significant predictor of injury, whether assessed alongside other HRM practices or alone).

6. Discussion

We set out to contribute to human resource management and occupational safety research by investigating the relative effects of particular HRM practices on safety performance. Specifically, we tested simultaneously five practices—systematic selection, extensive training, performance appraisal, high relative compensation, and empowerment—as predictors of organizational injury rate, controlling for company size and industrial sector injury rate. The present data show that, in this sample, higher empowerment is related to lower injury rates. The significant finding for empowerment is consistent with previous findings at the

Table 1
Means, standard deviations, and correlations among study variables.

	M/Mdn	*SD/IQR*	1	2	3	4	5	6	7
1. Injury rate (1-year lag)[a]	0.63	0–1.67							
2. Injury rate (2-year lag)[a]	1.04	0–2.22	0.42**						
3. Organizational size[b]	174.24	178.90	0.09	0.23					
4. Performance appraisal	2.31	1.10	−0.21	−0.11	0.27				
5. Systematic selection	3.19	0.59	0.06	0.03	0.24	0.31*			
6. High relative compensation	3.46	0.66	0.18	0.18	0.12	0.02	0.07		
7. Extensive training	2.95	0.92	−0.06	−0.17	0.15	0.58***	0.58***	0.00	
8. Empowerment	2.40	0.68	−0.14	−0.22	0.02	0.23	0.08	0.45**	0.16

Note. Correlations involving injury rate variables are non-parametric (Spearman) correlations. Injury rate measured as number of injuries per 100 employees. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. [a]The median (*Mdn*) and interquartile range (*IQR*) are reported these variables, due to a large skew. [b]The mean and standard deviation reported here are for the raw variable, even though the log of this variable is used in the inferential analysis.

Table 2
Regression of organizational injury rates on HRM practice variables.

	Model 1	Model 2
Intercept	*B* = 0.504 (0.270)	*B* = 0.159 (0.233)
Organizational size (log)	−0.350 (0.188)	−0.494 (0.145)
Sector-average injury rate	0.450 (0.221)	0.339 (0.190)
Systematic selection	–	0.529 (0.214)
Extensive training	–	−0.232 (0.262)
Performance appraisal	–	−0.167 (0.246)
High relative compensation	–	0.551 (0.192)
Empowerment	–	−0.780*(0.192)
Model χ^2	175.21	102.15
df	46	41
$\Delta\chi^2$, Δdf	–	73.06, 5, $p < 0.005$

Note. Figures in central section of table are regression coefficients (standard errors in brackets). *N* = 49. $^{*}p < 0.05$.

organizational- (e.g., Camuffo et al., 2017; Shannon et al., 1996; Yassi et al., 2004), group- (e.g., Hechanova-Alamay & Beehr, 2001), and individual-level of analysis (e.g., Parker et al., 2001) that indicate more autonomous working is related to better safety performance. Our study extends these findings using multiple sources of data and predicting organizational injury rates in the future to show that the relationship between empowerment and safety operates at the organizational-level analysis. Like previous organizational-level analyses, it shows associations between organizational-level constructs and organizational-level outcomes that appear stronger than in individual-level research (Ostroff & Bowen, 2000). Our study also extends prior research because it shows the value of empowerment over-and-above other intercorrelated HRM practices.

From a practical perspective, the current findings suggest that to reduce workplace injury rates, designing work to provide greater opportunity for autonomous work is one way organizations might achieve this. This could also be achieved by enriching jobs (e.g., job enrichment programs) or developing leadership skills among supervisors that value psychological empowerment (Parker & Wall, 1998). This finding is interesting in that it is the same practice (empowerment) that is most strongly associated with organizational productivity (Birdi et al., 2008). Thus, it appears that a key initiative likely to promote safety is consistent with, rather than at odds with, the basic economic need to enhance performance.

Contrary to our hypotheses, however, none of systematic selection, extensive training, performance appraisal, nor high relative compensation were associated with organizational injury rates. One factor that might account for these null effects is the relatively small size of the sample (49 companies with complete HRM practice and organizational injury data) may have insufficient power to find true relationships among the study variables. One specific consequence of this is that any effects are underestimated in the form of non-significant regression coefficients, increasing the likelihood of a Type II error; any results indicating null effects need to be considered as tentative. This possibility especially applies to the practices of performance appraisal and extensive training, which both had negative correlations with injury rates. These effects might have been significant had the power in the study been greater. In contrast, high relative compensation had a positive correlation with injuries, and selection practices had a negligible association, so, irrespective of sample size, these practices may seem less likely to be important. However, it is more likely that a complex relationship among HRM practices and organizational safety exists. Specifically, we cannot reasonably determine, due to the small size of the current sample, whether the interaction of certain practices (e.g., extensive training and systematic selection), or "bundles" of multiple practices, explain additional variance in organizational injury rates over-and-above the main effects of the five practices together. Thus, while there may not be main effects for systematic selection, extensive training, performance appraisal, and high relative compensation, their effects may still interact with other HRM practices.

Finally, despite the fact that measures used in the current study were not safety-specific and instead more general HRM practices, findings suggest the importance of safety-oriented HRM practices and safety outcomes. Similarly, there is a separate stream of literature focused on safety-specific management practices (i.e., occupational health and safety management systems (OHSMS), Fernández-Muñiz, Montes-Peón, & Vázquez-Ordás, 2007; Li & Guldenmund, 2018; Yorio & Wachter, 2014), which emphasize the integration of safety into all organizational capabilities (Fernández-Muñiz, Montes-Peón, & Vázquez-Ordás, 2009). The existing evidence suggests safety-specific practices such as safety-specific selection criteria (e.g., Vredenburgh, 2002), safety-specific training (e.g., Burke, Holman, & Birdi, 2006), safety-related compensation (e.g., safety incentive programs; Lauver, 2007), and safety-specific empowerment (e.g., employee involvement in safety-related activities; Yorio & Wachter, 2014) might predict organizational injury rates, although there is also evidence on the contrary that some safety-specific management practices may not contribute to the reduction of organizational injury (Lauver, 2007; Vredenburgh, 2002). Thus, the inconsistent findings from previous literature and the current findings emphasize the need to bridge HRM research and OHSMS research. In doing so, we may enhance our understanding of the relative importance of general HRM practices and safety-specific management practices in predicting organizational injury rates (cf. Robinson & Smallman, 2006).

6.1. Study limitations and future research

Like all studies, our study has a number of limitations. First, an important methodological concern is the generalizability of the final model. Despite the methodological strengths of our study (i.e., multi-source data and prospective design), we cannot be sure that some unmeasured third factor, such as a climate of trust, does not lead to both empowerment and lower injury rates. Second, although this study tested the hypotheses in a prospective design and found that empowerment was associated with organizational injury rates collected at a later point, this does not rule out the possibility of a reverse causal explanation. As the research design is not a 'true' longitudinal design in that it did not assess both empowerment and injuries at multiple time points (as recommended by Zapf, Dormann, and Frese (1996), the possibility exists that lower injury rates in some way led to the implementation of HRM practices such as empowerment.

Third, there might be a degree of error in the injury rates reported, with employees underreporting injuries and underestimation of organizational-level injury rates (Probst, Brubaker, & Barsotti, 2008). However, this is unlikely to be related systematically to empowerment or the other HRM practices, thus it would attenuate any observed relationship. A fourth possible limitation stems from the measurement of some of the key variables, such as systematic selection and extensive training, which were based on managers' ratings. Obtaining data from several different sources, as was the case in this study, has the advantage of minimizing concerns about common method variance. Nevertheless, it does raise the question of whether managers or other non-incumbents have a more accurate sense of how these HRM practices are implemented. Halo effects or other demand characteristics may bias respondent judgments (e.g., Semmer, Zapf, & Greif, 1996), and even asking trained raters to make judgments of HRM practices across a range of jobs may bias the relationships depending on the raters' point of view. Again, however, unless such biases are related to the number of injuries recorded, which is unlikely in that raters were unaware of those injury rates, the association would be to attenuate rather than exaggerate relationships.

Another possible limitation is that the current study did not include a measure of efficacy for the HRM practices. Mendelson, Turner, and Barling (2011) suggest that measuring the presence of a practice does not indicate whether the practice is actually efficacious. Inclusion of efficacy measures could help validate the measures used in the current study, and provide an indication of whether these practices were having an impact on intended outcomes. For example, extensive training practices may be validated by whether there are observed behavioral changes in the workplace. Additionally, turnover rates may be used as an indicator of whether systematic selection practices are effectively selecting people with the required knowledge, skills, and abilities (e.g., determining fit of candidate; Seldon & Sowa, 2015). As such, it is important for future research to consider more than the mere presence of HRM practices.

A final limitation is that we have not addressed the mechanisms by which HRM practices may exert their effects. It would be beneficial to examine potential intermediate linkages to better understand the role HRM practices have in shaping workplace safety outcomes (Granger et al., in press). There are several possible ways in which HRM practices might lead to lower injuries. First, at the organizational level of analysis, one potential mechanism is safety climate. Safety climate refers to employees' shared sense of the policies, practices, and procedures that reflect the extent to which safety is valued and rewarded (Zohar, 2014). HRM systems— comprised of practices that are designed to enhance employees' abilities, motivation, and opportunities to perform (Applebaum, Bailey, Berg, & Kalleberg, 2000)—may signal to employees that the organization encourages workplace safety. For example, empowerment practices aimed to enhance employees' autonomy (e.g., employees might proactively address safety issues with having control over their work methods; Turner et al., 2005) and encourage employee participation in organizational issues (e.g., speaking up about safety concerns; Tucker & Turner, 2015) may promote a positive safety climate. Furthermore, context-specific HRM practices may send stronger signals to employees about desirable behaviors and attitudes (Bowen & Ostroff, 2004). As we suggested previously, safety-specific HRM practices may send stronger signals about the relative importance of safety (e.g., safety-specific training) than general HRM practices do.

Second, in addition to an organization's HRM practices exerting effects on organizational-level injury rates, they may also be linked with employee-level injuries, especially if there are systematic differences between employees in different companies; this could be explored by cross-level analysis. A cross-level model consisting of HRM practices might explain between-organization variance in an employee-level mediating variable; this in turn can explain additional employee-level variance in an employee-level outcomes, while controlling for other employee-level factors.

Two possible paths through which this may occur are a mutual gains perspective and the conflicting outcomes perspective (Ogbonnaya et al., 2013; Van De Voorde, Paauwe, & Van Veldhoven, 2012). A mutual gains pathway suggests that both the organization and employee benefit from implementing HRM practices (Van De Voorde et al., 2012). Specifically, HRM practices may be negatively associated with employee-level injuries by fostering positive employee attitudes such as satisfaction, commitment, and trust (Ogbonnaya et al., 2013). For example, job autonomy might enhance employees' commitment to the organization, strengthening their motivation to meet organizational goals such as safety (Parker et al., 2001). This latter explanation would be consistent with findings that employee engagement might mediate the association between work practices and injury rates (Harter et al., 2002; Nahrgang et al., 2011). A second path suggests a trade-off between organizational outcomes and employee outcomes. In this view, organizations may reap the benefits of HRM practices, but HRM practices may not be beneficial, and may even be detrimental to employee outcomes (Van De Voorde et al., 2012). This critical perspective suggests that employee injuries may be positively associated with HRM practices through work intensification (Ogbonnaya et al., 2013). For example, increased perceived job demands is one way work intensification may manifest itself (Boxall & Macky, 2014), with evidence indicating that extended and overtime hours are related to increased risk of injury (Dembe, Erickson, Delbos, & Banks, 2005). In addition, more general job demands within the context of workplace safety (i.e., risks and hazards, physical demands, and complexity) are positively related to worsened safety outcomes through increased burnout (Nahrgang et al., 2011). Taken together, future research should explore the intermediate and cross-level linkages between HRM practices and injuries to enable a greater understanding of the conditions that promote safety

7. Conclusion

In summary, this study advances our understanding of organizational-level workplace safety. The results support the idea that organizations that promote empowered working also have lower injury rates, and that there is this association in the presence of other HRM practices. Future research should test the robustness of the model in other samples. Meanwhile, a clear policy implication from these findings is that there is merit in going beyond traditional occupational health and safety management systems to understand how more general HRM practices may help to improve workplace safety.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Applebaum, E., Bailey, T., Berg, P., & Kalleberg, A. L. (2000). *Manufacturing advantage: Why high-performance work systems pay off*. Cornell University Press.

Arocena, P., Nunez, I., & Villanueva, M. (2008). The impact of prevention measures and organizational factors on occupational injuries. *Safety Science, 46*, 1369–1384.

Arthur, J. B. (1994). Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal, 37*, 670–687.

Askenazy, P. (2001). Innovative workplace practices and occupational injuries and illnesses in the United States. *Economic and Industry Democracy, 22*, 485–516.

Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Services Research, 41*, 1576–1598.

Barling, J., & Frone, M. R. (Eds.). (2004). *The psychology of workplace safety*. Washington, DC: American Psychological Association.

Barling, J., & Hutchinson, I. (2000). Commitment vs. control-oriented safety practices, safety reputation and perceived safety climate. *Canadian Journal of Administrative Sciences, 17*, 76–84.

Barling, J., Kelloway, E. K., & Iverson, R. D. (2003). High-quality work, job satisfaction, and occupational injuries. *Journal of Applied Psychology, 88*, 276–283.

Barling, J., Loughlin, C., & Kelloway, E. K. (2002). Development and test of a model linking safety-specific transformational leadership and occupational safety. *Journal of Applied Psychology, 87*, 488–496.

Becker, B. E., & Huselid, M. A. (1998). High performance work systems and firm performance: A synthesis of research and managerial implications. In G. R. Ferris (Ed.), *Research in personnel and human resource management* (pp. 53–101). Greenwich, CT: JAI Press.

Beijer, S., Peccei, R., van Veldhoven, M., & Paauwe, J. (2021). The turn to employees in the measurement of human resource practices: A critical review and proposed way forward. *Human Resource Management Journal, 31*, 1–17.

Birdi, K., Clegg, C., Patterson, M., Robinson, A., Stride, C., Wall, T., & Wood, S. (2008). The impact of human resource and operational management practices on company productivity: A longitudinal study. *Personnel Psychology, 61*, 467–501.

Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal, 42*, 25–40.

Bloom, M., & Michel, M. G. (2002). The relationships among organizational context, pay dispersion, and managerial turnover. *Academy of Management Journal, 45*, 33–42.

Boon, C., den Hartog, D. N., & Lepak, D. P. (2019). A systematic review of human resource management and their measurement. *Journal of Management, 45*, 2498–2537.

Bowen, D. E., & Ostroff, C. (2004). Understanding HRM-firm performance linkages: The role of the "strength" of the HRM system. *Academy of Management Review, 29*(2), 203–221.

Boxall, P., & Macky, K. (2014). High-involvement work processes, work intensification and employee well-being. *Work, Employment and Society, 28*(6), 963–984.

Burke, M. J., Holman, D. J., & Birdi, K. (2006). A walk on the safe side: The implications of learning theory for developing effective safety and health training. In G. P. Hodgkinson & J. K. Ford (Eds.). *International review of industrial and organizational psychology* (Vol. 21). Chichester: Wiley.

Campbell, J. P., Daft, R. L., & Hulin, C. L. (1982). *What to study: Generating and developing research questions*. Thousand Oaks, CA: Sage.

Camuffo, A., De Stefano, F., & Paolino, C. (2017). Safety reloaded: Lean operations and high involvement work practices for sustainable workplaces. *Journal of Business Ethics, 143*, 245–259.

Chamberlin, M., Newton, D. W., & LePine, J. A. (2018). A meta-analysis of empowerment and voice as transmitters of high-performance managerial practices to job performance. *Journal of Organizational Behavior, 39*, 1296–1313.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83*, 234–246.

Clarke, S., Probst, T. M., Guldenmund, F., & Passmore, J. (Eds.). (2016). *The Wiley Blackwell Handbook of the Psychology of Occupational Safety and Workplace Health*. Chichester: John Wiley & Sons.

Cohen, A. (1977). Factors in successful occupational safety programs. *Journal of Safety Research, 9*, 168–178.

Dahl, M. S., & Pierce, L. (2020). Pay-for-performance and employee mental health: Large sample evidence using employee prescription drug usage. *Academy of Management Discoveries, 6*, 12–38.

Daley, D. M., Vasu, M. L., & Weinstein, M. B. (2002). Strategic human resource management: Perception among North Carolina Country Social Services professionals. *Public Personnel Management., 31*, 343–359.

Dekker, S. W., & Breakey, H. (2016). Just culture: Improving safety by achieving substantive, procedural and restorative justice. *Safety Science, 85*, 187–193.

Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal, 39*, 802–835.

Delery, J., & Gupta, N. (2016). Human resource management practices and organizational effectiveness: Internal fit matters. *Journal of Organizational Effectiveness, 3*(2), 139–163.

Dembe, A. E., Erickson, J. B., Delbos, R. G., & Banks, S. M. (2005). The impact of overtime and long work hours on occupational injuries and illnesses: New evidence from the United States. *Occupational and Environmental Medicine, 62*(9), 588–597.

DeNisi, A., & Smith, C. E. (2014). Performance appraisal, performance management, and firm-level performance: A review, a proposed model, and new directions for future research. *Academy of Management Annals, 8*(1), 127–179.

Fernández-Muñiz, B., Montes-Peón, J. M., & Vázquez-Ordás, C. J. (2007). Safety management system: Development and validation of a multidimensional scale. *Journal of Loss Prevention in the Process Industries, 20*, 52–68.

Fernández-Muñiz, B., Montes-Peón, J. M., & Vázquez-Ordás, C. J. (2009). Relation between occupational safety management and firm performance. *Safety Science, 47*(7), 980–991.

Forth, J., & Millward, N. (2004). High-involvement management and pay in Britain. *Industrial Relations, 43*, 98–119.

Fulmer, C. A., & Ostroff, C. (2016). Convergence and emergence in organizations: An integrative framework and review. *Journal of Organizational Behavior, 37*, S122–S145.

Godard, J. (2004). A critical assessment of the high-performance paradigm. *British Journal of Industrial Relations, 42*, 349–378.

Gong, Y., Law, K. S., Chang, S., & Xin, K. R. (2009). Human resources management and firm performance: The differential role of managerial affective and continuance commitment. *Journal of Applied Psychology, 94*, 263–275.

Granger, S., Turner, N., & Grocutt, A. (in press). The next best safety dollar: Using evidence to decide how to invest in workplace safety. Organizational Dynamics.

Hämäläinen, P., Takala, J., & Tan, B.K. (2017). Global estimates of occupational injuries and work-related illnesses 2017. Workplace Safety and Health Institute. Retrieved from https://www.wsh-institute.sg/~/media/wshi/past%20publications/2017/global%20estimate%20of%20occupational%20injuries%20and%20workrelated%20illnesses%202017.pdf.

Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology, 87*, 268–279.

Hechanova-Alamay, R., & Beehr, T. A. (2001). Empowerment, span of control, and safety performance in work teams after workforce reduction. *Journal of Occupational Health Psychology, 6*, 275–281.

Hofmann, D. A., & Tetrick, L. E. (Eds.). (2003). *Health and safety in organizations: A multi-level perspective*. San Francisco: Wiley.

Horowitz, L. M., Inouye, D., & Siegelman, E. T. (1979). On averaging judges' ratings to increase their correlation with an external criterion. *Journal of Consulting and Clinical Psychology, 47*, 453–458.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal, 38*, 635–672.

Kaminski, M. (2001). Unintended consequences: Organizational practices and their impact on workplace safety and productivity. *Journal of Occupational Health Psychology, 6*, 127–138.

Lauver, K. J. (2007). Human resource safety practices and employee injuries. *Journal of Managerial Issues, 19*(3), 397–413.

Lauver, K. J., & Trank, C. Q. (2012). Safety and organizational design factors: Decentralization and alignment. *Journal of Business and Management, 18*(1), 61–80.

Lepak, D. P., Liao, H., Chung, Y., & Harden, E. E. (2006). A conceptual review of human resource management systems in strategic human resource management research. In J. J. Martocchio (Ed.). *Research in personnel and human resources management* (Vol. 25, pp. 217–271). Elsevier Science/JAI Press.

Li, Y., & Guldenmund, F. W. (2018). Safety management systems: A broad overview of the literature. *Safety Science, 103*, 94–123.

Liao, H., Toya, K., Lepak, D., & Hong, Y. (2009). Do they see eye to eye? Management and employee perspectives of high-performance work systems and influence processes on service quality. *Journal of Applied Psychology, 94*, 371–391.

Littlejohn, A., Lukic, D., & Margaryan, A. (2014). Comparing safety culture and learning culture. *Risk Management, 16*, 272–293.

London, M., & Smither, J. W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12*, 81–100.

MacDuffie, J. P. (1995). Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial and Labor Relations Review, 48*, 197–221.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. Boca Raton, FL: Chapman and Hall.

Mendelson, M. B., Turner, N., & Barling, J. (2011). Perceptions of the presence and effectiveness of high involvement work systems and their relationship to employee attitudes. *Personnel Review, 40*(1), 45–69.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.

Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology, 96*, 71–94.

Neal, A., West, M. A., & Patterson, M. G. (2005). Do organizational climate and competitive strategy moderate the relationship between human resource management and productivity. *Journal of Management, 31*, 492–512.
Newnam, S., Warmerdam, A., Sheppard, D., Griffin, M., & Stevenson, M. (2017). Do management practices support or constrain safe driving behaviour? A multi-level investigation in a sample of occupational drivers. *Accident Analysis and Prevention, 102*, 101–109.
Nunez, I., & Prieto, M. (2019). The effect of human capital on occupational health and safety investment: An empirical analysis of Spanish firms. *Human Resource Management Journal, 29*, 131–146.
Ogbonnaya, C., Daniels, K., Tregaskis, O., & Van Veldhoven, M. (2013). Using HPWP to drive towards growth: The impact of occupational health and safety leadership. In G. Saridakis & C. L. Cooper (Eds.), *How HR can drive growth* (pp. 228–252). UK: Edward Elgar Publishing Limited.
Osterman, P. (1995). Skill, training and work organization in American establishments. *Industrial Relations, 34*, 125–147.
Ostroff, C., & Bowen, D. E. (2000). Moving HRM to a higher level: HRM practices and organizational effectiveness. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.
Parker, S. K. (2014). Beyond motivation: Job and work design for development, health, ambidexterity, and more. *Annual Review of Psychology, 65*, 661–691.
Parker, S. K., Axtell, C. M., & Turner, N. (2001). Designing a safer workplace: Importance of job autonomy, communication quality, and supportive supervisors. *Journal of Occupational Health Psychology, 6*, 211–228.
Parker, S. L., Bell, K., Gagné, M., Carey, K., & Hilpert, T. (2019). Collateral damage associated with performance-based pay: The role of stress appraisals. *European Journal of Work and Organizational Psychology, 28*, 691–707.
Parker, S. K., & Wall, T. D. (1998). *Job and work design: Organizing work to promote well-being and effectiveness*. London: Sage.
Patterson, M. G., Warr, P., & West, M. A. (2004). Organizational climate and company productivity: The role of employee affect and employee level. *Journal of Occupational and Organizational Psychology, 77*, 193–216.
Patterson, M. G., West, M. A., Shackleton, V. J., Dawson, J. F., Lawthom, R., Maitlis, S., Robinson, D. L., & Wallace, A. M. (2005). Validating the Organizational Climate Measure: Links to managerial practices, productivity and innovation. *Journal of Organizational Behavior, 26*, 379–408.
Patterson, M. G., West, M. A., & Wall, T. D. (2004). Integrated manufacturing, empowerment and company performance. *Journal of Organizational Behavior, 25*, 641–665.
Pfeffer, J. (1998). *The human equation: Building profits by putting people first*. Boston: Harvard Business School Press.
Pfeffer, J., & Langton, N. (1993). The effect of wage dispersion on satisfaction, productivity, and working collaboratively: Evidence from college and university faculty. *Administrative Science Quarterly, 38*, 382–407.
Posthuma, R. A., Campion, M. C., Masimova, M., & Campion, M. A. (2013). A high performance work practices taxonomy: Integrating the literature and directing future research. *Journal of Management, 39*(5), 1184–1220.
Probst, T. M. (2004). Safety and insecurity: Exploring the moderating effect of organizational safety climate. *Journal of Occupational Health Psychology, 9*, 3–10.
Probst, T. M., Brubaker, T. L., & Barsotti, A. (2008). Organizational injury rate underreporting: The moderating effect of organizational safety climate. *Journal of Applied Psychology, 93*, 1147–1154.
Robinson, A. M., & Smallman, C. (2006). The contemporary British workplace: A safer and healthier place? *Work, Employment and Society, 20*, 87–107.
Seibert, S. E., Wang, G., & Courtright, S. H. (2011). Antecedents and consequences of psychological and team empowerment in organizations: A meta-analytic review. *Journal of Applied Psychology, 96*, 981–1003.
Seldon, S. C., & Sowa, J. E. (2015). Voluntary turnover in nonprofit human service organizations: The impact of high performance work practices. *Human Service Organizations: Management, Leadership & Governance, 39*(3), 182–207.
Semmer, N., Zapf, D., & Greif, S. (1996). 'Shared job strain': A new approach for assessing the validity of job stress measurements. *Journal of Occupational & Organizational Psychology, 69*, 293–310.
Shannon, H. S., Walters, V., Lewchuk, W., Richardson, J., Moran, L., Haines, T., & Verma, D. (1996). Workplace organization correlates of lost time accident rates in manufacturing. *American Journal of Industrial Medicine, 29*, 258–268.
Shaw, J. W., & Delery, J. E. (2003). Strategic HRM and organizational health. In D. A. Hofmann & L. E. Tetrick (Eds.), *Health and safety in organizations: A multi-level perspective* (pp. 233–360). San Francisco: Wiley.
Shaw, J. D., Gupta, N., & Delery, J. E. (2002). Pay dispersion and workforce performance: Moderating effects of incentives and interdependence. *Strategic Management Journal, 23*, 491–512.
Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–428.
Shipton, H. J., Dawson, J. F., West, M. A., & Patterson, M. G. (2002). Learning in manufacturing organizations: What predicts effectiveness? *Human Resource Development International, 5*, 55–72.
Shipton, H. J., West, M. A., Dawson, J. F., Patterson, M. G., & Birdi, K. (2006). Human resource management as a predictor of innovation. *Human Resource Management Journal, 16*, 3–27.
Shipton, H. J., West, M. A., Parkes, C. L., Dawson, J. F., & Patterson, M. G. (2006). When promoting positive feeling pays: Aggregate job satisfaction, work design features and organizational innovation in manufacturing organizations. *European Journal of Work and Organizational Psychology, 15*, 404–430.
Simard, M., & Marchand, A. (1997). Workgroups' propensity to comply with safety rules: The influence of micro-macro organizational factors. *Ergonomics, 40*, 172–188.
Smith, M. J., Cohen, H. H., Cohen, A., & Cleveland, R. J. (1978). Characteristics of successful safety programs. *Journal of Safety Research, 10*, 5–15.
Smith-Crowe, K., Burke, M. J., & Landis, R. (2003). Organizational climate as a moderator of safety knowledge-safety performance relationships. *Journal of Organizational Behavior, 24*, 861–876.
Takala, J. (2019). Burden of injury due to occupational exposures. In U. Bültmann & J. Siegriest (Eds.), *Handbook of disability, work, & health* (pp. 1–22). Cham: Springer.
Takeuchi, R., Lepak, D. P., Wang, H., & Takeuchi, K. (2007). An empirical examination of the mechanisms mediating between high-performance work systems and the performance of Japanese organizations. *Journal of Applied Psychology, 92*(4), 1069–1083.
Tannenbaum, S. I., Mathieu, J. E., Salas, E., & Cannon-Bowers, J. A. (1991). Meeting trainees' expectations: The influence of training fulfillment on the development of commitment, self-efficacy, and motivation. *Journal of Applied Psychology, 76*, 759–769.
Tharenou, P., Saks, A. M., & Moore, C. (2007). A review and critique of research on training and organizational-level outcomes. *Human Resource Management Review, 17*, 251–273.
Toh, S. M., Morgeson, F. P., & Campion, M. A. (2008). Human resource configurations: Investigating fit with the organizational context. *Journal of Applied Psychology, 93*, 864–882.
Tsui, A. S., Pearce, J. L., Porter, L. W., & Tripoli, A. M. (1997). Alternative approaches to employee-organization relationships: Does "mutual investment" provide any advantage for organizations? *Academy of Management Journal, 40*, 1089–1121.
Tucker, S., & Turner, N. (2015). Sometimes it hurts when supervisors don't listen: Antecedents and consequences of safety voice among young workers. *Journal of Occupational Health Psychology, 20*, 72–81.
Turner, N., & Dueck, P.M. (2015). Health and safety. In D.E. Guest & D. Needle (Eds.), Wiley Encyclopedia of Management (3rd. ed., Vol. 5: Human Resource Management, pp. 111-114). Wiley: London.
Turner, N., Chmiel, N., & Walls, M. (2005). Railing for safety: Job demands, job control, and safety citizenship role definition. *Journal of Occupational Health Psychology, 10*, 504–512.
Turner, N., & Parker, S. K. (2004). The effect of teamwork on safety processes and outcomes. In J. Barling & M. R. Frone (Eds.), *The psychology of workplace safety* (pp. 35–62). Washington, DC: American Psychological Association.
Van De Voorde, K., Paauwe, J., & Van Veldhoven, M. (2012). Employee well-being and the HRM-Organizational performance relationship: A review of quantitative studies. *International Journal of Management Reviews, 14*, 391–407.
Vredenburgh, A. G. (2002). Organizational safety: Which management practices are most effective in reducing employee injury rates? *Journal of Safety Research, 33*, 259–276.
Wall, T. D., Jackson, P. R., & Davids, K. (1992). Operator work design and robotics system performance: A serendipitous field study. *Journal of Applied Psychology, 77*, 353–362.
Wall, T. D., Jackson, P. R., & Mullarkey, S. (1995). Further evidence on some new measures of job control, cognitive demand and production responsibility. *Journal of Organizational Behavior, 16*, 431–455.
Wall, T. D., & Wood, S. J. (2005). The romance of human resource management and business performance, and the case for big science. *Human Relations, 58*, 429–462.
Wall, T. D., Wood, S. J., & Leach, D. J. (2004). Empowerment and performance. In C. Cooper & I. T. Robertson (Eds.). *International review of industrial and organizational psychology* (Vol. 19, pp. 1–46). London: Wiley.
Wallace, J. C., Popp, E., & Mondore, S. (2006). Safety climate as a mediator between foundation climates and occupational accidents: A group-level investigation. *Journal of Applied Psychology, 91*, 681–688.
Weahrer, G. M., Miller, T. R., Hendrie, D., & Galvin, D. M. (2016). Employee assistance programs, drug testing, and workplace injury. *Journal of Safety Research, 57*, 53–60.
Werner, S., Kuiate, C. S., Noland, T. R., & Francia, A. J. (2016). Benefits and strategic outcomes: Are supplemental retirement plans and safer driving related in the U. S trucking industry? *Human Resource Management, 55*(5), 885–900.
West, M. A., Guthrie, J. P., Dawson, J. F., Borrill, C. S., & Carter, M. R. (2006). Reducing patient mortality in hospitals: The role of human resource management. *Journal of Organizational Behavior, 27*, 983–1002.
Wood, S., & de Menezes, L. (1998). High commitment management in the UK: Evidence from the Workplace Industrial Relations Survey, and Employers' Manpower and Skills Practices Survey. *Human Relations, 51*, 485–515.
Wright, P. M., Gardner, T. M., Moynihan, L. M., & Allen, M. R. (2005). The relationship between HRM practice and firm performance: Examining causal order. *Personnel Psychology, 58*, 409–446.
Wright, P. M., & Ulrich, M. D. (2017). A road well traveled: The past, present and future journey of strategic human resource management. *Annual Review of Organizational Psychology and Organizational Behavior, 4*, 45–65.
Yassi, A., Cohen, M., Cvitkovich, Y., Park, I. H., Ratner, P. A., Ostry, A. S., ... Pollak, N. (2004). Factors associated with staff injuries in intermediate care facilities in British Columbia. *Nursing Research, 53*, 87–97.
Yorio, P. L., & Wachter, J. K. (2014). The impact of human performance focused safety and health management practices on injury and illness rates: Do size and industry matter? *Safety Science, 62*, 157–167.

Youndt, M. A., & Snell, S. A. (2004). Human resource configurations, intellectual capital, and organizational performance. *Journal of Managerial Issues, 16*(3), 337–360.
Youndt, M. A., Snell, S. A., Dean, J. W., & Lepak, D. P. (1996). Human resource management, manufacturing strategy, and firm performance. *Academy of Management Journal, 39*, 836–866.
Zacharatos, A., & Barling, J. (2004). High-performance work systems and occupational safety. In J. Barling & M. R. Frone (Eds.), *The psychology of workplace safety* (pp. 203–222). Washington, DC: American Psychological Association.
Zacharatos, A., Barling, J., & Iverson, R. D. (2005). High-performance work systems and occupational safety. *Journal of Applied Psychology, 90*, 77–93.
Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies of organizational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology, 1*, 145–169.
Zohar, D. (2002). The effects of leadership dimensions, safety climate, and assigned priorities on minor injuries in work groups. *Journal of Organizational Behavior, 23*, 75–92.
Zohar, D. (2014). Safety climate: Conceptualization, measurement, and improvement. In B. Schneider & K. M. Barbera (Eds.), *The Oxford handbook of organizational climate and culture* (pp. 317–334). Oxford University Press.

Nick Turner, PhD (Sheffield) is Professor and Distinguished Research Chair in Advanced Leadership at the Haskayne School of Business, University of Calgary, Canada. His research focuses on work design, transformational leadership, and occupational health and safety.

Julian Barling, PhD (Witwatersrand) is Professor, Distinguished University Professor, and Borden Chair in Leadership at Smith School of Business, Queen's University, Canada. His research focuses on the predictors and consequences of leadership and well-being.

Jeremy F. Dawson, PhD (Aston) is Professor of Health Management at Sheffield University Management School, University of Sheffield. UK. His research focuses on management of health care organizations, team working, and statistics.

Connie Deng, MSc (Sheffield) is Research Associate at the Canadian Centre for Advanced Leadership in Business, Haskayne School of Business, University of Calgary, Canada. Her research focuses on workplace safety.

Sharon K. Parker, PhD (Sheffield) is Australian Research Council Laureate Fellow, John Curtin Distinguished Professor, and Director of the Centre for Transformative Work Design at Curtin University, Perth, Australia. Her research focuses on work design, proactive behavior, and well-being.

Malcolm G. Patterson, PhD (Sheffield) is Senior Research Fellow, Sheffield University Management School, University of Sheffield, UK. His research focuses on organizational culture and climate, human resource management and performance, and work and depression.

Chris B. Stride, PhD (Warwick) is Senior Lecturer, Sheffield University Management School, University of Sheffield, UK. His research focuses on applied statistics, cheating in sport, and statues of sportsmen and sportswomen.

Journal of Safety Research 78 (2021) 189–202

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Identifying the impact of the COVID-19 pandemic on driving behavior using naturalistic driving data and time series forecasting

Christos Katrakazas *, Eva Michelaraki, Marios Sekadakis, Apostolos Ziakopoulos, Armira Kontaxi, George Yannis

National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou str., GR-15773 Athens, Greece

A R T I C L E I N F O

Article history:
Received 4 November 2020
Received in revised form 1 February 2021
Accepted 27 April 2021
Available online 6 May 2021

Keywords:
COVID-19
Driving behavior
Time-series forecasting
SARIMA
XGBoost

A B S T R A C T

Introduction: COVID-19 has disrupted daily life and societal flow globally since December 2019; it introduced measures such as lockdown and suspension of all non-essential movements. As a result, driving activity was also significantly affected. Still, to-date, a quantitative assessment of the effect of COVID-19 on driving behavior during the lockdown is yet to be provided. This gap forms the motivation for this paper, which aims at comparing observed values concerning three indicators (average speed, speeding, and harsh braking), with forecasts based on their corresponding observations before the lockdown in Greece. *Method:* Time series of the three indicators were extracted using a specially developed smartphone application and transmitted to a back-end platform between 01/01/2020 and 09/05/2020, a time period containing normal operations, COVID-19 spreading, and the full lockdown period in Greece. Based on the collected data, XGBoost was employed to identify the most influential COVID-19 indicators, and Seasonal AutoRegressive Integrated Moving Average (SARIMA) models were developed for obtaining forecasts on driving behavior. *Results:* Results revealed the intensity of the impact of COVID-19 on driving, especially on average speed, speeding, and harsh braking per 100 km. More specifically, speeds were found to increase by 2.27 km/h on average compared to the forecasted evolution, while harsh braking/100 km increased to almost 1.51 on average. On the bright side, road crashes in Greece were reduced by 49% during the months of COVID-19 compared to the non-COVID-19 period.

1. Introduction

The first cases of COVID-19 (also reported as SARS-CoV-2 or simply Coronavirus) were reported in the city of Wuhan in China in December 2019 (Cheng & Shan, 2020; Lau et al., 2020; Wu et al., 2020). After a significant rise in the new cases across the globe, it was declared a pandemic in March 2020 (WHO, 2020). At present, confirmed cases of COVID-19 are more than 93.1 million, while COVID-19-induced casualties are more than 1.98 million (WHO, 2020).

In an effort to restrict the spread of the virus among susceptible population groups, a "lockdown" restricting all non-essential activities was imposed by the majority of governments worldwide. Citizens were also instructed to practice "social distancing" by means of keeping at least 2 meters away from each other. Confinement and "social distancing" aimed to slow down the spread of the disease. Moreover, in combination with the aforementioned measures, schools, theaters, cinemas, restaurants, fitness centers, and shops were closed to avoid crowding. As a result, financial, environmental, and social impacts were observed (Anderson et al., 2020; Hendrickson & Rilett, 2020; Zhang et al., 2020a, 2020b).

Driving behavior also changed radically. Road traffic volume, public transport users, and overall mobility activity reduced significantly (Apple, 2020; Google LLC, 2020; Moovit, 2020). For example, a study in the city of Santander, Spain, analyzed the impact of COVID-19 confinement and demonstrated that overall activity decreased by 67% and public transport use decreased by 93% (Aloi et al., 2020), while forecasts on travel demand revealed less traffic, public transport usage, and congestion or flow levels (Aloi et al., 2020; De Vos, 2020). In the same context, nearly 80% of people in the Netherlands reduced their activities outdoors, and subsequently elderly people had a greater decrease (de Haas et al., 2020). A behavioral change in mobility as a result of COVID-19 could also be on track, as according to de Haas et al. (2020) 27% of Dutch people stated that they will work from home more frequently, while 20% expressed the willingness to cycle and walk more in the future. Road traffic crashes were found to be reduced as road traffic and pedestrian volume decreased (Aloi et al., 2020). Furthermore, data provided by TomTom showed that traffic volumes decreased

* Corresponding author.
E-mail address: ckatrakazas@mail.ntua.gr (C. Katrakazas).

https://doi.org/10.1016/j.jsr.2021.04.007
0022-4375/

by 70–85% in the majority of European cities (ETSC, 2020; TomTom, 2020). Despite a dramatic decline in traffic volumes due to COVID-19 restrictions, in urban areas there was a 35% increase in speeding and an almost 200% increase in stunt driving offences in the March 15–31, 2020 timeframe, compared to the same period last year (City of Toronto, 2020).

Nevertheless, to date, the impact of COVID-19 on transportation can only be assessed through individual reports (e.g., Molloy, 2020) or web applications of data companies such as Google (Google LLC, 2020), Apple (2020), and TomTom (2020), which have recorded mobility activities during the lockdown phase. The impact on driving behavior still remains relatively unknown. This fact is the motivation for the current paper, which aims at quantifying the effect of the COVID-19 lockdown on driving behavior through a naturalistic driving dataset captured through a novel mobile phone application. More specifically data on driving speed, speeding, and harsh braking/100 km are recorded before, during, and slightly after the imposition of a lockdown state in Greece. Time series forecasts of driving indicators based on the normal (pre-lockdown) phase are used to compare observed driving behavior with a normal evolution of itself, so as to quantify the change during lockdown.

A variety of published studies and reports were examined concerning road collisions, injuries, and fatalities. Road crashes were reduced in the majority of countries as road traffic and pedestrian volume decreased (Aloi et al., 2020). Road crashes in Germany decreased by approximately 23% during a quarantine month, injured people by 27%, and fatalities by 32% compared to the same period last year (DW, 2020). The same impact was observed in the Netherlands, where 50% less collisions were reported. Italy, France, and Spain displayed a drop in road deaths of 40–70%, however, in Australia the reported deaths had not declined despite the overall reduced traffic (ETSC, 2020). Barnes et al. (2020) revealed that the total number of crashes decreased; but unfortunately, crashes involved individuals (especially males) from age 25 to 64. Lin et al. (2020) highlighted that, although the number of nonfatal crashes reduced, the severe and fatal cases of road crashes were not changed during the pandemic. The overall number of road crashes as well as crash fatalities reported across United States was also reduced (Wagner et al., 2020). Although the number of road crashes was in general positively correlated with the amount of traffic volumes, the number of fatalities, surprisingly, was observed to experience an increase at some states during COVID-19 period (Vingilis et al., 2020).

This study is structured as follows: (a) an introduction to the subject of the paper (see above); (b) a brief literature review on driving behavior in relation with the effect of COVID-19 or other pandemics on transportation; (c) a description of the methodological approach and the utilized data; (d) XGBoost analyses, which are used to explore the importance of contributor variables, are then conducted; (e) the main part of the paper is dedicated to time-series forecasting and the comparison between observed driving behavior indicators and the forecasted ones; and (f) conclusions on the impact of COVID-19 on driving behavior are drawn and a discussion on how policy-makers and researchers should take advantage of the analysis is provided.

2. Literature review

The literature research aimed to link driving behavior, mobility, and transportation with the COVID-19 pandemic. The literature search was conducted in popular scientific databases such as Scopus, Science Direct and Google Scholar. The Boolean terms used to search these databases were *"COVID-19" or "Corona Virus" or "SARS Cov 2" and "road" or "traffic" or "safety" or "accidents" or "collisions" or "mobility" or "transport" or "transportation" or "behavior" or "behavior."* The searches were limited to engineering and social sciences, and the results included approximately 18,500 studies (at present: 5/1/2021). These studies were screened concerning their titles and abstracts, and the most relevant papers to the investigating topic are included in this review.

Initially, De Vos (2020) analyzed the effect of COVID-19 in terms of the impact of social distancing on travel behavior, while Budd and Ison (2020) introduced a new theoretical concept of responsible transport that tries to reconsider transport policy due to behavioral change of passenger during the pandemic. Moreover, Vingilis et al. (2020) investigated the COVID-19 disease and its effects on road safety and it was revealed that travel decreased and drivers were exposed to a lower risk of collisions. Inada et al. (2020) indicated that empty roads triggered some speed-related traffic law violations among drivers, such as speeding, failing to stop at a stop sign, red light running, and failing to yield to pedestrians. In addition, Neuburger and Egger (2020) revealed an increase in risk perception of COVID-19, travel risk perception, and travel behavior over a short period of time. However, the aforementioned studies were limited to discussions over the impact of COVID-19 and did not provide quantifiable results on the impact of the pandemic on driving behavior.

Apart from studies discussing the impact of COVID-19 on travel behavior, particular emphasis was given on descriptive statistics regarding average speed, speeding, harsh events, mobile phone use, and driving distance per trip during the pandemic. For example, Aloi et al. (2020) conducted an empirical study and concentrated exclusively on urban mobility and COVID-19. Only descriptive results were included in that study and the authors demonstrated the change of mode choice, purpose of trip, number of trips, macroscopic traffic flow, public transport trips travel time and demands, and general trip features during lockdown. It was also revealed that, in Tokyo, speeding increased by 52% in March 2020 compared to March 2019, while the police officially enhanced enforcement of fines and penalties for speeding during the following months (Inada et al., 2020). Similarly, Katrakazas et al. (2020) provided descriptive evidence from Greece and Saudi Arabia on the deterioration of road safety levels during the period of the lockdown. In particular, it was shown that reduced traffic volumes due to lockdown led to a slight increase in average driving speed by 6–11%, but more importantly, to more frequent harsh accelerations and harsh braking per 100 km (up to 12%). Additionally, during March and April 2020, which were the months where COVID-19 spread was at its peak, mobile phone use while driving increased by 42%, while driving within the risky hours (00:00am–05:00am) dropped by up to 81%. Furthermore, spatial patterns of speeding pre (2019) and post (2020) the COVID-19 outbreak were visualized and compared in order to examine if the spatial extent of speeding increased (Lee et al., 2020).

Regarding studies employing questionnaires, a large-scale sample survey was conducted by de Haas et al. (2020) with questions concentrating mostly on mobility behavior, population, or demographic characteristics. Their findings concern the mobility behavior change since the COVID-19 outbreak. They investigated the change concerning the purpose of traveling, number of trips travel modes, stated opinion for future outdoors activities, remotely working, or education aspects. In the same context, Mogaji (2020) conducted an online survey to evaluate the impact of COVID-19 on transportation. More explicitly, the examined variables were mode choice, public transportation choice, and reduction of social, religious, and economic activities during the COVID-19.

To date, only a few studies have conducted statistical analyses on the effect of COVID-19 on driving behavior. One of them is Stavrinos et al. (2020), which utilized multi-level modelling to investigate driving behavior of adolescents in the United States before and during the pandemic period. Their results indicated that after the appearance of COVID-19 pandemic and its corresponding restrictions, driving days per week decreased by 37%, while vehicle miles driven dropped by 35. Nevertheless, the data used were concerned with self-reported driving behavior, and as a result a bias existed. Within-subjects general linear models were used in Roe et al. (2020) to investigate driving behavior of older adults during COVID-19. Indicators used included mean length in miles, the average speed in miles per hour of each trip, along with the mean number of three types of aggressive behaviors (i.e., harsh braking, harsh accelerations, and speeding) per mile per trip. It was clearly highlighted that participants reduced the proportion of days driven during the pandemic compared with the same period the year before. At the same time, trips per day showed a similar decline. Participants also took shorter trips, drove slower, had fewer speeding incidents, and had different trip destinations.

As can be understood from the previous paragraph, no study has statistically analyzed the impact of the COVID-19 pandemic on driving behavior and road safety, nor has any study taken into account time patterns in corresponding data. As a result, a gap in the literature exists that the current paper hopes to fill by performing time-series analysis in driving behavior data during the COVID-19 pandemic. In order to quantify the daily impact of COVID-19 on driving behavior, time-series analysis is deemed the most appropriate method and, as a result, a review of the literature was also conducted on methodological issues. Several published papers have used the corresponding variables to estimate the driving behavior.

3. Methodology

In order to quantify the daily impact of the COVID-19 pandemic on driving behavior indicators, a statistical relationship between COVID-19 and observed driving indicators had to be established. Therefore, a feature importance algorithm was used to evaluate the significance of variables on forecasting speed, speeding and harsh braking/100 km. After the initial explanatory analysis, in order to assess how driving behavior changed over time during the pandemic, time-series forecasting was exploited. For each of the three indicators (i.e., speed, speeding, harsh braking/100 km), the daily time-series was extracted as well as the time-series describing the evolution of COVID-19 cases and casualties. For the time-series analysis using ARIMA models, the following steps were followed according to Bisgaard and Kulahci, (2011), Box and Jenkins, (1976) and Essi, (2018):

- Seasonal decomposition to identify the trend, seasonality and residual variance
- Stationarity check using the augmented Dickey-Fuller test (Dickey & Fuller, 1979)
- Consideration of a general ARIMA Model
- Autocorrelation and Partial Autocorrelation plots to explore the relationship between time point and individual lags and find a tentative model
- Determination of the model using a parameter search
- Split into training and test dataset
- Forecasting and evaluation of the predictions

The aforementioned methods and steps are further elaborated in the following paragraphs.

3.1. XGBoost algorithms

As a preliminary step, Extreme Gradient Boosting (XGBoost) algorithms were implemented so that the importance of the collected variables, including the COVID-19 related variables, could be assessed and quantified in regards to the examined driving behavior indicators (i.e., speed, speeding, and harsh braking/100 km). XGBoost is a potent machine learning (ML) technique, encompassing multiple Classification And Regression Trees (CART), also known as tree ensemble. Additionally, XGBoost belongs to the family of supervised ML techniques, meaning that it uses labeled training data, the structure of which is defined by the researcher. In practice, this means that the independent/dependent variable division is known and present in the examined variables, and the outcome is a mapping function to the effect of $y = f(x)$.

XGBoost algorithms apply the gradient boosting decision tree algorithm, also known as multiple additive regression trees, stochastic gradient boosting, or gradient boosting machines. The learning process of the algorithm is iterative and includes correction of previous errors in future iterations of the algorithm. A detailed presentation of the algorithm is described in the seminal study by Chen and Guestrin (2016). XGBoost has been demonstrated to outclass other ML methods such as Random Forests and Support Vector Machines in performance both in road safety (Ting et al., 2020) and in other fields (Nielsen, 2016).

Furthermore, XGBoost algorithms have functions that can calculate the importance of each predictor variable. This is known as Gini feature importance, or, equivalently, Mean Decrease in Impurity (MDI), and was proposed in a seminal study by Breiman (2001). One definition for Gini Importance for tree-based algorithms is the following: Gini Importance is the value obtained as the sum over the number of splits that include the feature across all trees, optionally divided by the number of samples it splits. This allows for powerful and accurate models to be created by utilizing only the most important predictor variables from a given dataset.

In XGBoost, three particular variable importance metrics are observed (XGBoost developer team, 2019):

- Gain, describing the improvement in accuracy added by a feature to the branches it is on.
- Cover, describing the relative quantity of observations (or number of samples) concerned by a feature.
- Frequency, describing the number of times a feature is used in all generated trees.

These variable importance metrics used by the XGBoost algorithms were calculated in the analysis and examined to reveal that variables are informative to describe the examined driving behavior indicators.

3.2. Time-series forecasting

Autoregressive Integrated Moving Average (ARIMA) type models are considered the most popular time-series models, and are extensively used in the transportation research field. Their popularity can be explained due to their well-defined theoretical background and their quite straightforward calculations (Karlaftis & Vlahogianni, 2009). Thus, ARIMA models were deemed the most appropriate to model the impact of COVID-19 on daily driving behavior. An ARIMA model is a generalization of an Autoregressive Moving Average (ARMA) model and are generally denoted as:

$$ARIMA(p, d, q) \tag{1}$$

where: p denotes the autoregressive order (i.e., number of time lags), d denotes the differencing (i.e., the number of differencing transformations required by the time series in order to become stationary.), q denotes the non-seasonal moving average order (i.e., the lag of the error component, which is the part of the time series not explained by trend or seasonality).

Then, the model can be written more formally as (Wang et al., 2020):

$$\Phi_p(B)\left(1-B\right)^d y_t = \theta_q(B)\epsilon_t \quad (2)$$

where: $\Phi \in \mathbb{R}^p$ is a vector of coefficients for the AR terms, $\theta \in \mathbb{R}^q$ is a vector of coefficients for the MA terms, y_t is the outcome variable measured at time t, B is a vector used equivalently to indicate the lag operator, ϵ_t is random error (white noise, residual) associated with measurement t with $\epsilon_t \sim N(0,\sigma^2)$

On the other hand, Seasonal ARIMA models are used when the time series exhibits seasonality. These models are similar to ARIMA models and they are usually denoted as:

$$ARIMA(p,d,q)(P,D,Q)m \quad (3)$$

where: p denotes the non-seasonal autoregressive order, d denotes the non-seasonal differencing, q denotes the non-seasonal moving average order, P denotes the seasonal autoregressive order, D denotes seasonal differencing, Q denotes seasonal moving average order, m is the number of periods in each season and the seasonal ARIMA model can be generalized as:

$$\Phi_p\left(B^S\right)\Phi_p(B)\left(1-B\right)^d\left(1-B^S\right)y_t = \theta_q\left(B^S\right)\theta_q(B)\epsilon_t \quad (4)$$

where: p denotes the non-seasonal autoregressive order, S is the period at which the seasonal trend occurs, B is a vector used equivalently to indicate the lag operator, Φ is a vector of coefficients for the AR terms, d denotes the non-seasonal differencing and y_t is the outcome variable measured at time t.

3.2.1. Seasonal decomposition and stationarity

In cases of evident seasonality, ARIMA models can be extended to seasonal ARIMA (SARIMA) models. SARIMA models are considered as a straightforward extension of the non-seasonal ARIMA (Hipel & McLeod, 1994). With regards to SARIMA models, related studies were found to perform better than models of random walk (Clark et al., 2003; Ghosh et al., 2005; Williams, 2003), support vector regression (SVR) (Lippi et al., 2013), historical average (Chung & Rosalion, 2001; Williams, 2003) as well as regular ARIMA (Lippi et al., 2013; Clark et al. in Williams, 2003). Another study reported that the seasonal ARIMA models predicted more accurately, compared to the best performing k-NN (k-nearest neighbors algorithm) forecast models (Smith et al., 2002 in Kumar & Vanajakshi, 2015).

As a first step for the model identification and the interpretation of time-series data, the decomposition of the time series of the observed variable was required in order to identify its fundamental (and unobserved) parts: trend, seasonality, and residuals. A time series decomposition was used to measure the strength of trend and seasonality in a time series (Wang et al., 2006). The manner in which the decomposition is performed depends on whether time-series data are multiplicative or additive (Hyndman & Athanasopoulos, 2018). The decomposition can be written as:

$$y_t = T_t + S_t + R_t \quad (5)$$

where: y_t is the outcome variable measured at time t, T_t is the smoothed trend component, S_t is the seasonal component, R_t is a remainder component.

It was also essential to make sure that the utilized time series were stationary (Hyndman & Athanasopoulos, 2018). In order to make a time series stationary, a transformation was applied to the data, using the method of differencing. The latter removed the changes in the level of a time series, eliminating trend and seasonality and consequently stabilizing the mean of the time series. In order to check a time series for stationarity the Augmented Dickey-Fuller test was utilized (Dickey & Fuller, 1979). The Augmented Dickey-Fuller (ADF) test is checking if $\varphi = 0$ in models of the form:

$$\Delta y_t = a + \beta t + \varphi y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots \varepsilon_t \quad (6)$$

where: y_t is the observed time series data, a is a constant, β is the coefficient of the time trend, ε_t is a zero-mean error term. Using the ADF test, if $\varphi = 0$, then a unit root does not exist for the observed time series and the time series is non-stationary. In the different case that $\varphi < 0$, the time series is stationary.

3.2.2. Autocorrelation and partial autocorrelation

In order to identify an initial ARIMA model, the plots of the Autocorrelation (ACF) and Partial Autocorrelation functions (PACF) were used.

Correlation between two random variables X and Y can be defined as (Dettling, 2018):

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \quad (7)$$

where: $Corr(X,Y)$ is the correlation between the two variables, $Cov(X,Y)$ is the covariance of the two variables, and $Var(X), Var(Y)$ are their individual variances.

For stationary time series, autocorrelation (i.e., the correlation of a specific variable with its earlier iteration) can be defined as a function of the lag k (Dettling, 2018):

$$\rho(k) = Corr(X_{t+k}, X_t) \quad (8)$$

where: X_t is the measurement at time t and X_{t+k} is the measurement at lag k. It can be understood that the ACF defines the correlation of an observation X_{t+k} with previous measurements X_t of the same variable.

Similarly, partial autocorrelation at lag k can be defined as:

$$\pi(k) = Corr(X_{t+k}, X_t | X_{t+1} = x_{t+1}, \cdots, X_{t+k-1} = x_{t+k-1}) \quad (9)$$

which denotes the association between X_{t+k} and X_t, given that the linear dependence between X_{t+1} and X_{t+k-1} is removed.

By plotting both ACF and PACF it was easier to identify the correlation between more recent observations of the variable and simultaneously the existence of either actual lagged autocorrelations or autocorrelations caused by other measurements.

3.2.3. Model identification

In order to decide the parameters p, d, q for the ARIMA model as mentioned in equation (3) and the corresponding parameters for a potential SARIMA model, an automatic search of the best parameters according to the Akaike information criterion (AIC) or Bayesian Information Criterion (BIC) was used. The automatic search was based on popular packages in R and Python programming languages (Hyndman & Khandakar, 2007; Smith, et al., 2017), which have been found to be implemented successfully in recent publications (Ma et al., 2018, 2020). The best fitting model was selected based on the smallest AIC and BIC.

The Ljung-Box test (Ljung & Box, 1978) a popular diagnostic tool to test model fitness was also utilized. The Ljung-Box test is defined as:

- H_0 :The model does not exhibit lack of fit
- H_1 :The model exhibits lack of fit

given the test statistic:

$$Q = n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k^2}{n-k} \quad (10)$$

where: n is the length of the time series, $\hat{r}_k$ is the estimated autocorrelation of the time series at lag k and m is the number of lags being tested.

The test rejects the null hypothesis if:

$$Q > \chi^2_{1-\alpha,h} \quad (11)$$

where: $\chi^2_{1-\alpha,h}$ is the chi-square distribution table value with h degrees of freedom and signicance level α. The degrees of freedoms should be equal to $m-p-q$, where m is the number of residual autocorrelations that need to be checked, and p, q are the autoregressive and moving average ARIMA parameters, respectively.

3.2.4. Choosing the training and testing samples

As the purpose of this paper was to quantify the effect of COVID-19 on three driving behavior indicators (i.e., driving speed, speeding, and harsh braking/100 km), the ARIMA models were trained using a representative dataset of normal operations (i.e., prior to COVID-19) and tested on the early stages of COVID-19 spread in Greece when no countermeasures were taken. Following the development of training and testing procedures for the algorithm, then forecasts of these normal operations-based models during the lockdown time period would give a picture of how these traffic indicators would normally evolve and could enable comparisons between the actual observations during the lockdown phase and the forecasted ones. In order to assist comparisons, time series models were trained using data from the months of January and February (i.e., when no COVID-19 case was reported in Greece), were tested on the period before the lockdown and were validated on the time period concerning mid-March until early May when the lockdown status was lifted. It should be noted that the training, test, and validation set was the same for all the examined variables (i.e., average speed, speeding, harsh braking/100 km). Fig. 1 depicts an example of training, test, and validation set.

3.2.5. Evaluation of predictions

After developing the ARIMA models on the testing and validation sets, forecasts were evaluated using popular forecasting evaluation metrics such as:

- Mean Error (ME), which gives the mean of the forecasting error:

$$ME = \frac{1}{N}\sum e_t \tag{12}$$

- Mean Absolute Error (MAE), which gives the mean of the absolute forecasting error:

$$MAE = \frac{1}{N}\sum |e_t| \tag{13}$$

- Mean Percentage Error (MPE), which gives the mean of the forecasting error in percentage:

$$MPE = \frac{1}{N}\sum \frac{e_t}{observed_t} \cdot 100 \tag{14}$$

- Mean Absolute Percentage Error (MAPE), which depicts the mean error in percentage terms:

$$MAPE = \frac{1}{N}\sum \frac{|e_t|}{observed_t} \tag{15}$$

- Root Mean Squared Error (RMSE), which is the square root of the average squared error:

$$RMSE = \sqrt{\frac{1}{N}\sum e_t^2} \tag{16}$$

where: N is the number of forecasted points, and e_t is the error (i.e. $observed_t - forecasted_t$)

Finally, statistical significance of the non-seasonal and seasonal components of the ARIMA models was checked.

4. Data overview

For the purposes of this study, a large naturalistic dataset of daily driving trips was used. The datasets correspond to a complete 5-month timeframe spanning from 01/01/2020 to 09/05/2020 in Greece. The timeframe was chosen so that sufficient periods are available both before the spread of COVID-19 to represent normal operations and during the COVID-19 pandemic to quantify the effect of the lockdown measures.

The first case of COVID-19 in Greece was diagnosed on 26/02/2020. The first reactive measure that was enforced in Greece, after the initial diagnosis of coronavirus, was the nationwide suspension of the operation of educational institutions of all levels on 10/03/2020. This was followed by the decision to close

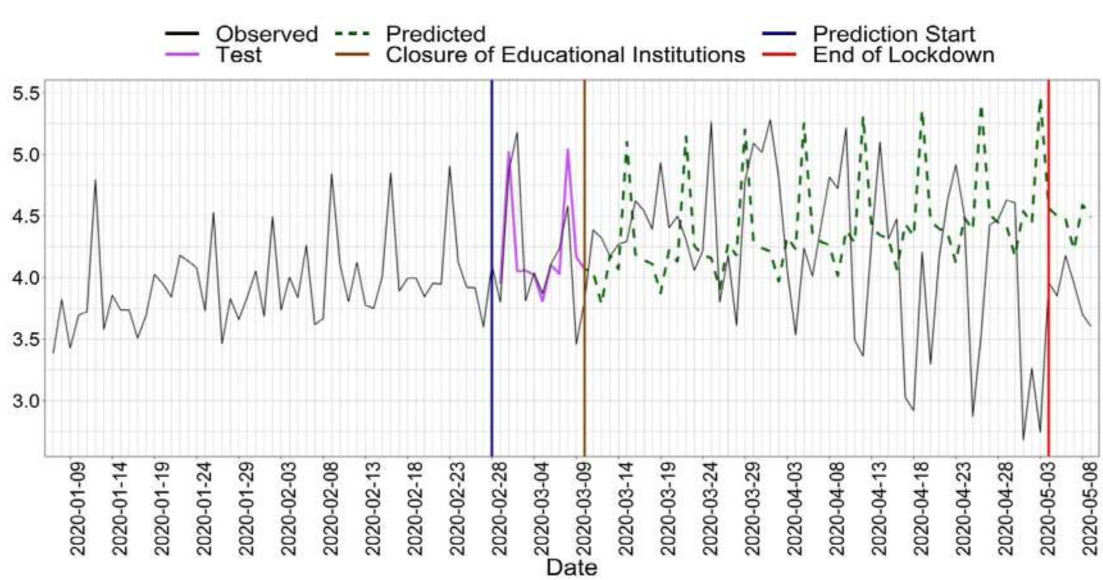

Fig. 1. An overview of training, test and validation sets.

Table 1
Driving performance indicators along with their corresponding description *(Source: OSeven, Data processing: NTUA)*.

Variables	Unit	Description
Average speed	km/h	Average speed during driving with stops excluded from the duration of the trip
Average total speed	km/h	Average speed during the total duration of the trip
Speeding	km/h	Average speed over the speed limit
Duration of speeding	sec	Total duration of speeding in a trip
Harsh accelerations/100 km	–	Number of harsh acceleration events per distance (100 km)
Harsh braking/100 km	–	Number of harsh braking events per distance (100 km)
Total distance	km	Total trip distance
Total duration	sec	Total trip duration
Driving duration	sec	Total duration of driving, i.e. duration of stops has been excluded
Risky hours driving	km	Distance driven during risky hours (00:00–05:00)
Duration of mobile phone use	sec	Total duration of mobile usage
GR-Total Cases	–	Total number of confirmed cases due to COVID-19 pandemic in Greece

down all cafes, restaurants, bars, shopping centers, sports facilities, museums, and all services in the areas of religious worship of any religion and finally, a restriction on all non-essential movement was put in place on the 23/03/2020. The lockdown was lifted on the 04/05/2020 after 42 days.

For the purpose of the analyses, microscopic trip data and a representative subset of 122,275 trips was provided by OSeven Telematics. It should be noted that the microscopic trip data used referred to the users of OSeven smartphone application and not the entire population of Greece. Both male and female drivers aged 18–65 participated and a large database of thousands of trips was obtained through the OSeven application. The sample utilized in this research was also counterbalanced regarding age group and gender, in order to be as representative as possible. The raw driving behavior data from the mobile phone sensors (i.e., GPS, accelerometers, or gyroscope) was collected through driving behavioral analytics platforms, state-of-the-art technologies and smart algorithms, reliable metrics and novel gamification schemes, developed by OSeven. Several published studies have used naturalistic driving data from smartphone sensors provided by OSeven Telematics for investigating driving parameters such as driving behavior (Papadimitriou et al., 2019; Stavrakaki et al., 2020; Tselentis et al., 2019). Regarding data collection, data from smartphone sensors have been shown to allow for continuous and rapid data collection and seamless storage and analysis. Since smartphones are programmable, flexible implementation possibilities become available. However, there are increased demands in data storage and analysis, and considerable upfront costs during development of the data handling infrastructure, with much lower costs as time progresses and participant numbers increase (Ziakopoulos et al., 2020).

It should be mentioned that as privacy and security consist two of the platform's main design principles, all data are stored and specific measures are taken to protect them based on encryption standards for data in transit and at rest. The above procedure is done using the latest technologies that comply with the national regulation in EU as well as with the General Data Protection Regulation (GDPR). As a result, all data has been provided by OSeven Telematics in a completely anonymized format. Readers are also referred to the studies provided in section 2.1 for a more detailed description of the OSeven application and platform in the scientific literature.

The collected trip data contained information on driving performance regarding average driving speed, average total speed, average speed limit exceedance (speeding), harsh events (i.e., harsh accelerations, harsh braking), other trip characteristics (i.e., total distance and total duration), as well as mobile phone use or driving during risky hours (00:00am–05:00am). The descriptive statistics of the aforementioned indicators during the collection period (i.e., between 01/01/2020 and 09/05/2020) are depicted in Table 1.

Within this paper, three variables were selected and analyzed in detail:

- average speed (km/h)
- speeding (km/h); namely average excess speed over the limit
- harsh braking per distance (100 km)

Furthermore, the enforcement of quarantine measures during the critical period for Greece is treated as a binary quantity: the value 1 is assigned for trips during the period from 23/03/2020 to 05/05/2020, and the value 0 is assigned for all other examined trips.

Table 2 illustrates descriptive statistics (i.e., mean, standard deviation, maximum value, minimum value) with regards to the examined variables for the complete subset of trips (122,275 trips), while Table 3 depicts descriptive statistics for the examined variables for the months of COVID-19 in Greece (46,614 trips). It should be also noted that the sample size for each variable in Table 2 is $N = 130$, while for the examined variables for the months of COVID-19, the sample size is denoted by $n = 56$ values.

5. Exploratory feature analysis with XGBoost

An initial exploration of feature importance as yielded by the implementation of XGBoost methodology is conducted in this section. All variables are positive continuous variables, and are

Table 2
Descriptive statistics for the examined variables for the complete subset of trips in Greece (from 01/01/2020 to 09/05/2020).

Variable	Mean	Standard deviation	Maximum value	Minimum value
Average speed (km/h)	43.16	2.65	49.68	38.82
Speeding (km/h)	4.09	0.53	5.28	2.68
Harsh braking/100 km	13.07	1.91	20.33	7.06
Total trips	**122,275**			
Sample size of each variable (N)	**130**			

Table 3
Descriptive statistics for the examined variables for the months of COVID-19 in Greece (from 26/02/2020 to 09/05/2020).

Variable	Mean	Standard deviation	Maximum value	Minimum value
Average speed (km/h)	44.34	2.68	49.68	38.82
Speeding (km/h)	4.17	0.61	5.28	2.68
Harsh braking/100 km	13.26	2.08	20.33	7.70
Total trips	**46,614**			
Sample size of each variable (n)	**56**			

therefore examined with a regression with squared loss function. In order to calibrate the XGBoost tree ensemble, a uniform split was applied in the described data: 75% was randomly designated as the training subset, while the remaining 25% was randomly designated as the test subset. Furthermore, a number of hyperparameters can be optimized for each XGBoost ensemble, such as learning rate (eta), gamma, maximum tree depth, minimum child weight, number of rounds and mean squared error. The selection of the optimal values is conducted by examining large numbers of hyperparameter combinations, as described by Bischl et al. (2016); in this research, a grid search of 5000 hyperparameter combinations was conducted for each analysis. All XGBoost analyses were conducted in R-studio (R Core Team, 2019).

5.1. Average speed (km/h)

The examined range and obtained parameters from the XGBoost tuning for average speed are provided in Table 4.

The predictive power was provided by the application of the XGBoost tree ensemble on the test subset, and yielded RMSE = 1.106 and MAPE = 0.024. The respective obtained feature importance is provided in Table 5.

5.2. Speeding (km/h)

The examined range and obtained parameters from the XGBoost tuning for speeding is provided in Table 6.

The predictive power was provided by the application of the XGBoost tree ensemble on the test subset, and yielded RMSE = 0.318 and MAPE = 0.062. The respective obtained feature importance is provided in Table 7.

5.3. Harsh braking/100 km

The examined range and obtained parameters from the XGBoost tuning for harsh braking/100 km is provided in Table 8.

Table 4
Examined and optimized hyperparameters for average speed XGBoost algorithms.

Hyperparameter	Examined range	Optimized Value
Learning rate	0.000–1.000	0.38
Gamma	0–100	4.17
Maximum tree depth	1–50	9
Minimum child weight	1–10	2
Number of rounds	1–1000	42
Mean Squared Error	as low as possible	1.256

Table 5
Feature importance of average speed XGBoost algorithms.

	Feature	Gain	Cover	Frequency
1	GR-Total Cases	0.574	0.306	0.308
2	Total distance	0.387	0.489	0.500
3	Trip duration	0.035	0.154	0.154
4	Risky hours	0.005	0.051	0.038

Table 6
Examined and optimized hyperparameters for speeding XGBoost algorithms.

Hyperparameter	Examined range	Optimized Value
Learning rate	0.000–1.000	0.06
Gamma	0–100	0.34
Maximum tree depth	1–50	2
Minimum child weight	1–10	4
Number of rounds	1–1000	250
Mean Squared Error	as low as possible	0.177

Table 7
Feature importance of speeding XGBoost algorithms.

	Feature	Gain	Cover	Frequency
1	Total distance	0.551	0.558	0.467
2	GR-Total Cases	0.212	0.224	0.228
3	Trip duration	0.130	0.142	0.152
4	Duration of mobile use	0.054	0.053	0.076
5	Quarantine	0.027	0.018	0.033
6	Risky hours	0.026	0.006	0.043

Table 8
Examined and optimized hyperparameters for harsh braking/100 km XGBoost algorithms.

Hyperparameter	Examined range	Optimized Value
Learning rate	0.001–0.6	0.374
Gamma	0.001–10	1.37
Maximum tree depth	2–10	6
Minimum child weight	1–10	1
Number of rounds	1–250	242
Mean Squared Error	as low as possible	0.018

The predictive power was provided by the application of the XGBoost tree ensemble on the test subset, and yielded RMSE = 1.279 and MAPE = 0.08. The respective obtained feature importance is provided on Table 9.

In summary, the COVID-19-related parameter of total cases in Greece seems to exert a considerable influence in allowing the prediction of average speed, speeding, and harsh braking/100 km, as expressed by the gain scores of each XGBoost tree ensemble. This applies for the presence and enforcement of quarantine measures for average speeding as well. It is apparent that the exposure variables of total trip distance and duration also affect all examined quantities, and a small contribution is also provided by driving during risky night-time hours.

6. Time-series modelling assessment

6.1. Model specification

Following the identification of the influence of COVID-19-related parameters on driving behavior indicators, Seasonal Autoregressive Integrated Moving Average (SARIMA) modelling was followed to quantify the impact of the pandemic. The three components (i.e., trend, seasonality, and residuals) for the time-series of the considered indicators were analyzed. It was observed that the seasonal component for all three indicators changed over time, and similar patterns were observed for consecutive months. However, later observations displayed greater difference. With regards to average speed and speeding, there was an overall increasing trend through the months, which means that there was a significant rise in average speed and speeding during the period of COVID-19 pandemic. Taking into consideration harsh braking/100 km, a smaller seasonal trend was evident. With regards to the non-seasonal trend, values reached a maximum during mid-March, but started to decrease thereafter. Lastly, the contribution of random noise was negligible for all the examined variables.

Table 9
Feature importance of harsh braking/100 km XGBoost algorithms.

	Feature	Gain	Cover	Frequency
1	Duration	0.541	0.396	0.327
2	Totaldist	0.278	0.211	0.173
3	GRTotalCases	0.075	0.043	0.135
4	Risky_hours	0.058	0.206	0.192
5	Time_mobile_usage	0.049	0.144	0.173

As described in the methodology, the next step in the time series pipeline was to check for stationarity. For that purpose, the ADF test was performed for the 1st difference (Y_t-Y_{t-1}) of average speed, speeding, and harsh braking/100 km. In order to eliminate the seasonal effect from the time series observations, a seasonal first differencing was utilized for all considered time-series. The first difference was used because all the original time series were not stationary. Such a transformation also assisted to consolidate the variance of a time series. Moreover, differencing can help stabilize the mean of a time series by removing changes in the level of a time series, and therefore eliminating, or reducing, trend and seasonality.

Table 10 illustrates the ADF test for the original variables, while Table 11 depicts the ADF test for the 1st difference of each variable. In addition, Table 12 illustrates the Box-Ljung Test (white noise check) for the examined variables, which was performed on the 1st difference of average speed, speeding and harsh braking/100 km.

ACF and PACF plots for the 1st difference of driver behavior indicators (i.e., average speed, speeding and harsh braking/100 km) were performed that indicated the levels at which the autocorrelation is significant and determined the order of the autoregressive term. It was found that both ACF and PACF dropped to zero relatively quickly. For all the candidate SARIMA models, p-values of the autoregressive and seasonal autoregressive terms were found to be less or equal to 0.05, which indicates their statistical significance. Table 13 shows the final selected candidate models along with their corresponding ARIMA terms. The specifications of the best model per driving behavior indicator with regards to AIC and BIC are also demonstrated.

Following the observation of the ACF and PACF plots, the best models were obtained through the forecast package in R (Hyndman et al., 2020), as described in Table 14. The number in brackets (i.e. [7]) represents the exponential decay in weekly lags. As described in section 3.2, the corresponding values for the best SARIMA models denote the non-seasonal and seasonal autoregressive order, differencing as well as moving average order, respectively.

In order to further validate the models, their residuals were also checked. As depicted in Fig. 2, the residuals appear to be randomly scattered, and no evidence of the error terms being correlated with each other exists. Consequently, the residuals or errors can be conceived as independently and identically distributed (i.i.d.) sequences with a constant variance and a zero mean. Therefore, the developed SARIMA models appeared to be well-fitted and were chosen to be used for prediction.

Table 15 illustrates the results of the estimated SARIMA models for each of the three variables. With regards to the error terms of RMSE and MAE, the values of RMSE were proven to be larger, which means that all the errors are not of the same magnitude; actually, the greater difference between them, the greater the variance in the individual errors.

Regarding RMSE, the best performance is observed for speeding with 0.45, while the worst performance is observed using the average speed time series. Looking, however, at the MAPE indicator, it is distinguishable that the speed time series resulted in the best forecasting performance, with only 3.46% difference from the observed measurements. This is further resembled in the MPE indicator with speed having a 0.23% difference from the observed values. Lastly, with regards to the first-order autocorrelation coefficient (ACF1), all three SARIMA models perform well, with the speeding time series having the best performance. It should be noted that as the autocorrelation function can provide the correlation among different points separated by various time lags, ACF1 is a measure of how much is the current value influenced by the previous values in a time series.

Table 10
Augmented Dickey-Fuller Test for the considered variables.

	Augmented Dickey-Fuller Test		
Original	Test statistics	Lag order	*p*-value
Average speed	−2.49	5	0.37
Speeding	−3.59	5	0.04
Harsh braking/100 km	−3.53	5	0.04

Table 11
Augmented Dickey-Fuller Test for the 1st difference of each variable.

	Augmented Dickey-Fuller Test		
Variable	Test statistics	Lag order	*p*-value
diff(*Average speed*)	−8.94	5	0.01
diff(*Speeding*)	−10.09	5	0.01
diff(*Harsh braking/100 km*)	−8.39	5	0.01

Table 12
Box-Ljung Test for the 1st difference of each variable.

	Box-Ljung Test		
Variable	X^2	df	*p*-value
diff(*Average speed*)	78.1	24	1.218e-07
diff(*Speeding*)	55.64	24	2.563 e-04
diff(*Harsh braking/100 km*)	44.75	24	6.22 e-03

Table 14
SARIMA models for the 1st difference of each variable.

Variable	SARIMA Model
diff(*Average speed*)	(0,1,1)x(0,1,1)[7]
diff(*Speeding*)	(1,1,0)x(1,1,0)[7]
diff(*Harsh braking/100 km*)	(0,1,2)x(1,0,2)[7]

Table 13
Summary of estimated candidate SARIMA models for the 1st difference of each variable.

Variable	Candidate model	Estimate	Std. Error	z value	Pr(>|z|)	AIC	BIC
diff(*Average speed*)	ma1	−0.95	0.13	−7.12	1.07e-12 ***	116.94	122.36
	sma1	−0.74	0.36	−2.01	0.042 *		
diff(*Speeding*)	ar1	−0.60	0.12	−4.99	6.15e-07 ***	25.11	30.53
	sar1	−0.49	0.14	−3.65	2.61e-04 ***		
diff(*Harsh braking/100 km*)	ma1	−0.45	0.13	−3.43	6.09e-04 ***	194.11	205.70
	ma2	−0.34	0.14	−2.49	0.013 *		
	sar1	0.66	0.22	2.93	3.337e-03 **		
	sma1	−0.92	0.31	−2.95	3.174e-03 **		
	sma2	0.69	0.37	1.87	0.061 .		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

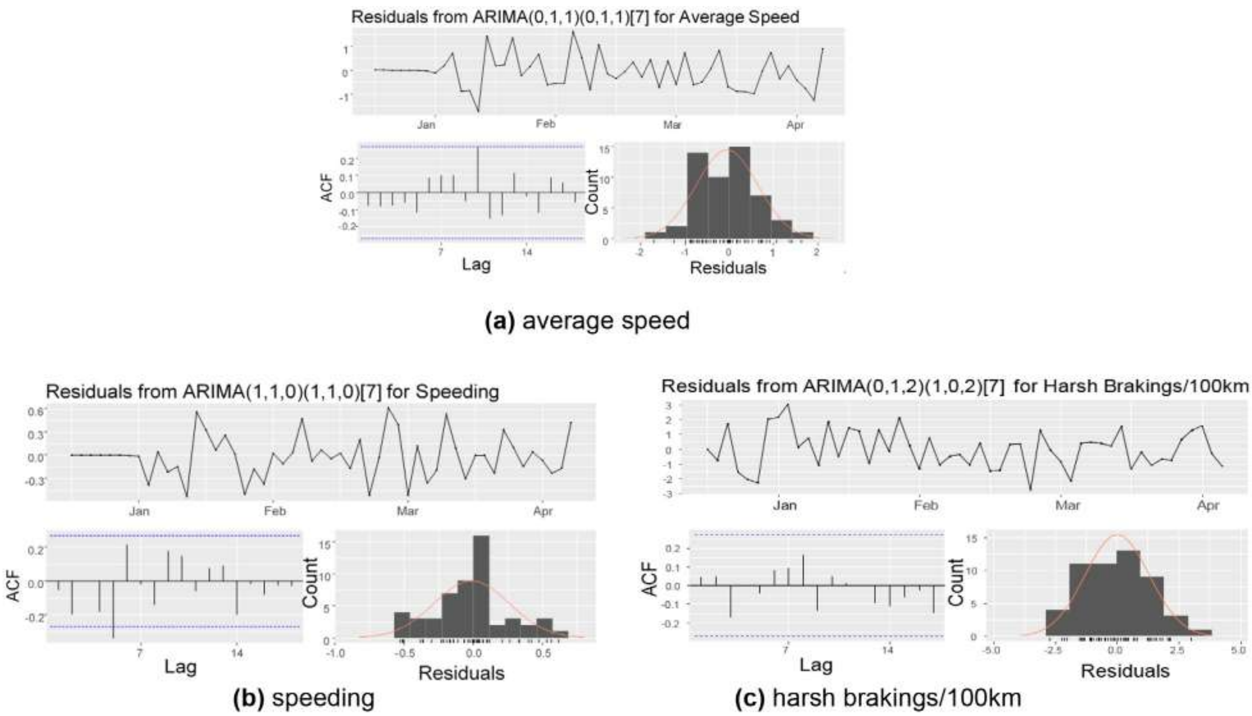

Fig. 2. Residual Plots for the 1st difference of each variable: (a) average speed, (b) speeding, (c) harsh braking/100 km.

Table 15
Performance metrics for the estimated SARIMA model on the test set.

Variable	ME	RMSE	MAE	MPE	MAPE	ACF1
diff(*Average speed*)	0.19	2.23	1.51	0.23	3.46	0.17
diff(*Speeding*)	−0.05	0.45	0.31	−1.98	7.25	−0.01
diff(*Harsh braking/100 km*)	−0.33	1.98	1.61	−5.60	15.04	0.34

6.2. Forecasting

After validating the performance of the developed model, focus was given on the forecasting performance. The results of the models with regards to the COVID-19 lockdown period is described in the following sections.

6.2.1. Average speed (km/h)

With regards to average speed, the forecasted values were based on the (0,1,1)x(0,1,1) SARIMA model. It can be observed that speed significantly increased over the COVID-19 lockdown with large fluctuations, while in normal conditions speed would not increase above 46 km/h. To further illustrate the effect of COVID-19 lockdown on average speed, the differences between forecasted and observed values, as well as the RMSE, MAPE and ACF1 were estimated. Fig. 3 depicts the SARIMA model for average speed for the prediction time along with the differences between average speed observed and predicted values.

6.2.2. Speeding (km/h)

Concerning speeding, the forecasts were based on the (1,1,0)x(1,1,0) best-fit SARIMA model. It is evident that speeding was forecasted to be increased during the months of March and April, but actually demonstrated a downwards trend during the pandemic. Regarding the difference between the observed and forecasted values, Fig. 4 illustrates that in the beginning of March and until the beginning of April, the actual values for speeding were higher than the potential normal forecasted values, while within April speeding gradually decreased.

6.2.3. Harsh braking/100 km

It was found that observed values differ a lot from the forecasts. In more detail, harsh braking/100 km was forecasted to have a frequency of around 12 and 13 events/100 km, but observed values are largely higher than the forecasts, reaching a maximum of 21 and minimum of 9 harsh braking/100 km. From Fig. 5, it is further validated that during the COVID-19 pandemic, values for harsh braking/100 km were much higher than the forecasted values.

6.2.4. Overall evaluation

In order to have an overall picture of the difference between forecasted and observed values, the MAPE, RMSE, ACF1 errors as well as the minimum and maximum and average of the three indicators were obtained and are described in Table 16.

From Table 16, it can be observed that in terms of RMSE, speed is performing worse than the rest of the three indicators, but better in terms of MAPE. As a result, forecasts for speed tend to be more accurate than forecasts for speeding and harsh braking/100 km. With regards to the average difference between observed and forecasted values, similar to RMSE speed provided larger errors but this is due to the fact that speed units (i.e. km/h) are larger than the measurement units of speeding or harsh braking per distance. Finally, observing the dates for minimum and maximum values of forecasts, the minimum difference between observed and forecasted values was identified on the 3rd of May, the last day of the lockdown, for speed and speeding, while the minimum difference for harsh braking/100 km was found on the 9th of March. In addition, the maximum difference for the average speed and speeding time series was observed on the 25th of March, a traditional Greek holiday (i.e., Greek Independence Day), and with regards to harsh braking/100 km, the maximum difference was identified in the 11th of April.

7. Discussion

This study aimed to quantify the effect of the COVID-19 pandemic on driving behavior by forecasting the evolution of time-

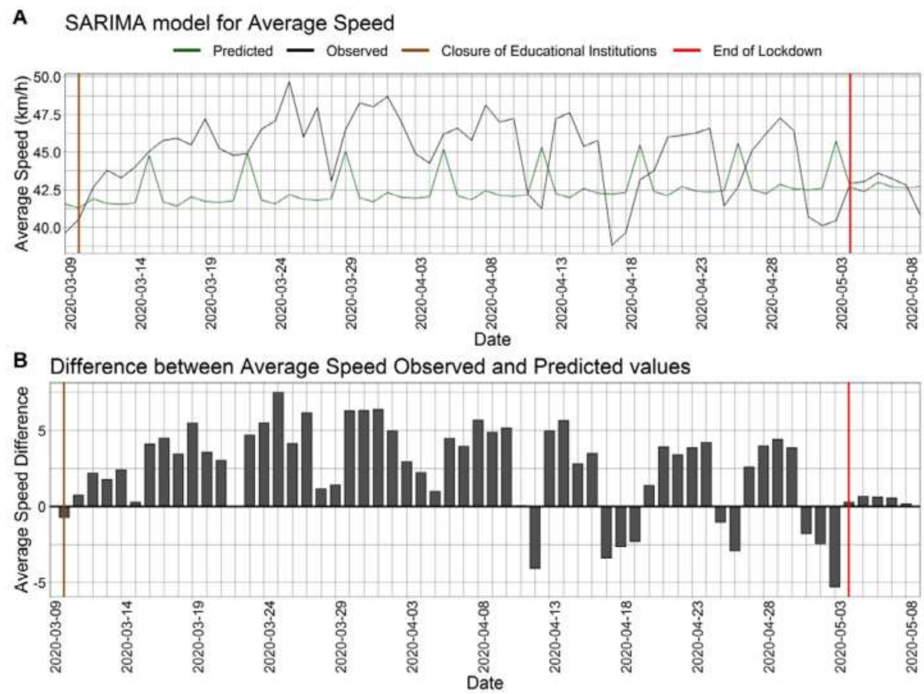

Fig. 3. SARIMA model forecasts for average speed and Differences between observed and predicted average speed values.

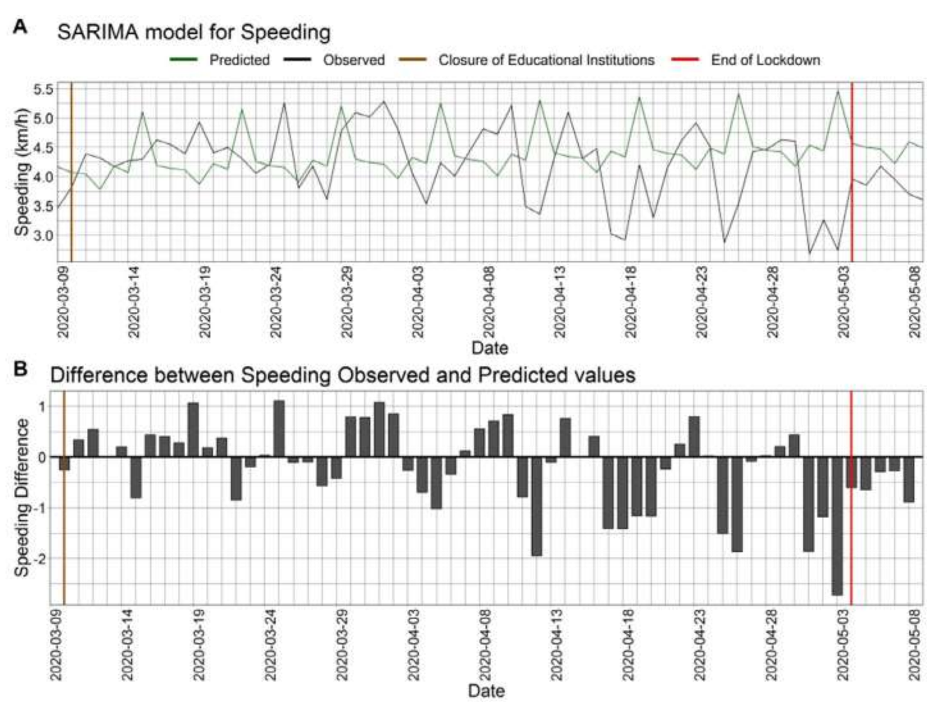

Fig. 4. SARIMA model forecasts for speeding and differences between observed and predicted speeding values.

series data based on values before the pandemic and comparing forecasts with actual values during the COVID-19 lockdown in Greece. Initially, the importance of COVID-19 indications (e.g., cases, casualties, lockdown countermeasures) was investigated using the feature importance extracted from XGBoost algorithms. The number of total cases was one of the two most important factors for three out of the three examined indicators (i.e. speed, speeding, and harsh braking/100 km). As a result, it can be derived that the spread of the virus had a significant effect on driving behavior. Total distance and trip duration were also among the most influential factors for all examined indicators. This can be explained by the significant decrease in trip duration and distance driven during the lockdown phase as seen in Fig. 6.

The effect of the lockdown initiation was not found to have a significant effect on driving behavior indicators, as was indicated in section 5 of the current paper. This is probably explained by

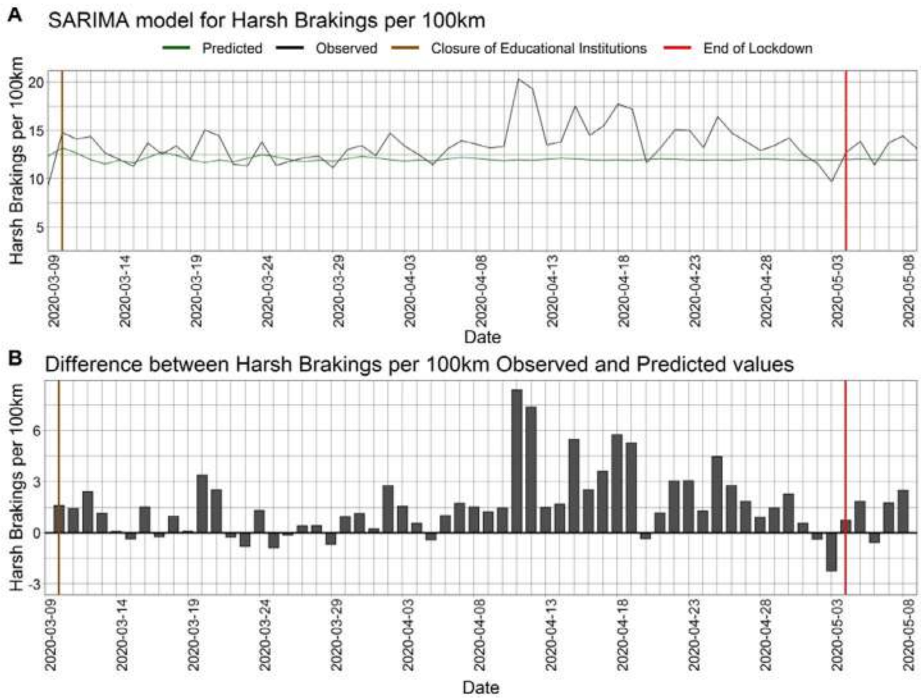

Fig. 5. SARIMA model for harsh brakings /100 km and differences between observed and predicted harsh braking values.

Table 16
MAPE, RMSE, ACF1, min, max, average difference for observed and forecasted values.

Variables	MAPE	RMSE	ACF1	Minimum Difference (Date of occurrence)	Maximum Difference (Date of occurrence)	Average
Average speed	7.12	3.76	0.50	−5.30 (3/5/2020)	7.51 (25/3/2020)	2.27
Speeding	17.76	0.87	0.44	−2.72 (3/5/2020)	1.11 (25/3/2020)	−0.22
Harsh brakings/100 km	12.27	2.51	0.43	−3.06 (9/3/2020)	8.37 (11/4/2020)	1.51

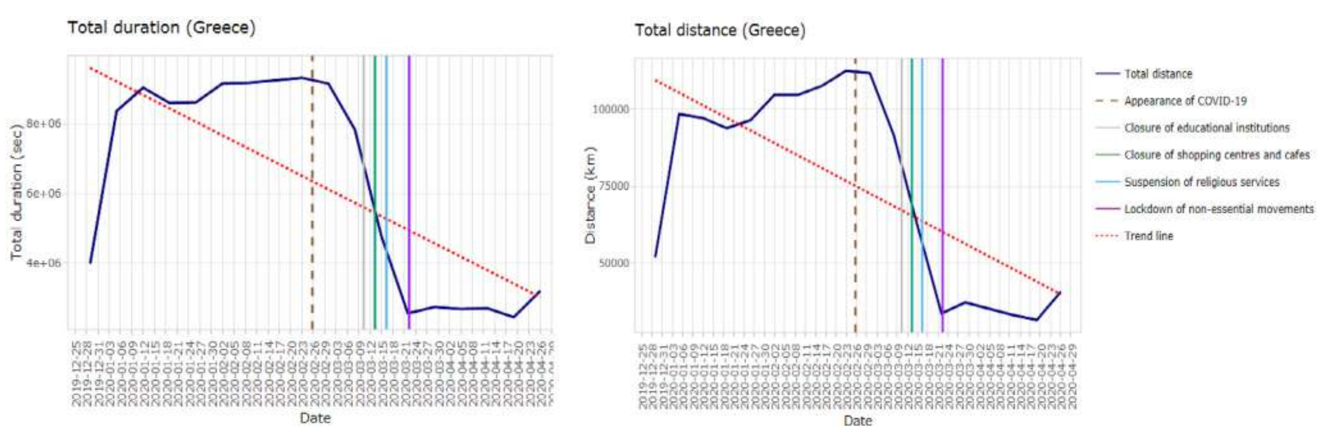

Fig. 6. Total duration and distance of trips during the COVID-19 period in Greece.

the fact that Greek drivers did not change their behavior due to the lockdown but rather because of the spread of the virus that led them to minimize trips and car driving.

The authors recognize that there is no immediate causal connection between total COVID-19 cases and/or quarantine presence and road safety, in the form of an established road safety risk factor or measure. The findings of the exploratory analysis through XGBoost, however, suggest a correlational value of the parameter of number of cases at least temporally. Therefore, they provided additional incentive for the time-series modelling and examination of trends that impact the three examined road safety indicators, as the pandemic progresses.

The effect of COVID-19 on driving behavior in terms of average speed, speeding, harsh braking/100 km both during the COVID-19 pandemic and the time period before the first case of the disease in Greece was quantified through seasonal time series modeling approach. With regards to the forecasting of the "normal evolution" (i.e., the potential evolution if COVID-19 had not spread and no lockdown measures were applied) it was observed that the best model was obtained for average speed with only 3.46%

Table 17
Road crashes and persons injured from January to April 2020 in Greece (*Hellenic Statistical Authority (2021)*).

	January	February	March	April	Change (March–April) – (January–February)
Road crashes	788	858	507	326	−**49%**
Fatalities	49	47	24	21	−**53%**
Severe injuries	46	31	26	21	−**39%**
Slightly injuries	909	965	545	332	−**53%**

of MAPE. In general, no large errors were observed with harsh braking/100 km (i.e., the variable with the larger MAPE) being able to forecast with 84.96% accuracy on the test dataset. By applying the developed models on the validation dataset that described the lockdown period in Greece and looking at the difference between observed and forecasted values, the effect of COVID-19 could be evaluated for all three indicators.

With respect to average speed, it was revealed that the observed values were higher than the forecasted ones, which means that a significant increase in average speed was identified during the COVID-19 lockdown. The maximum difference (7.51) between observed and predicted values of average speed was identified in 25/03/2020, during the lockdown period due to COVID-19, which seems reasonable as the 25th of March is a public holiday. Conversely, the minimum difference (-5.30) between actual and forecasted values of average speed was found in 03/05/2020, a day before the gradually lift restrictions in Greece, when drivers started to restart their business activity and return to their daily routines. On average, it was demonstrated that observed speeds are 2.27 km/h higher than the forecasted ones, but as seen from the maximum and minimum values a lot of variance existed. This finding can be explained by the fact, that with emptier streets and much lower volumes, average vehicle speed tends to be increased. This finding can be supported by Inada et al. (2020) who indicated that the empty roads possibly triggered speed-related violations among drivers.

Speeding was forecasted to be increased during the months of March and April; however, a downward trend during the pandemic was demonstrated, but the models failed to predict it. Especially in March, it was demonstrated that actual values for speeding were higher than forecasted, while within April speeding gradually decreased, with an overall average difference of 0.22 fewer events between observed and predicted values. The demonstrated average reduction is contradicting with the increase of speed that was observed, but can be explained by the fact that the forecasting model for speeding was the worst in terms of RMSE and MAPE. As a result, the forecasting ability of the model cannot capture successfully the evolution of speeding occurrence and results should be interpreted with caution.

The forecasting results on harsh braking/100 km demonstrated that if no lockdown was imposed, the average number of harsh brakings/100 km would be lower for the majority of the lockdown days. Increases of harsh braking/100 km with lower traffic and higher speeds are compliant with recent research using similar data (i.e., from smartphones), where it is stated that with higher speeds more harsh braking events occur (Petraki et al., 2020).

Lastly, a more comprehensive picture of the effects of COVID-19 pandemic on road safety can be drawn from the high quality data on total number of road crashes along with the corresponding fatalities, severe and slight injuries. Table 17 illustrates the difference in the total number of road crashes and persons injured from January to April 2020 in Greece. In particular, a 49% reduction in the total number of road crashes was observed during March-April 2020 (i.e., months of COVID-19) compared to January–February 2020 (i.e., when no COVID-19 case was reported in Greece). Furthermore, during March–April 2020, the total number of fatalities decreased by 53%, severe injuries were reduced by 39%, while slight injuries were reduced by 53% compared to January–February 2020.

Despite the fact that provisional data for road crashes occurred in 2020 (Hellenic Statistical Authority, 2021) showed that there was a decrease in absolute numbers of crashes, fatalities, and injuries, driving performance was found to be more careless and more risky overall during the lockdown period. This finding can be supported by previous studies in which it was found that less vehicle traffic volumes and empty roads led to higher speeds and harsh events (Carter, 2020). Results from the current research are also consistent with findings reported by Wagner et al. (2020), who analyzed U.S. data from the second quarter of 2020 compared to the first quarter. It was revealed that the total number of road crashes and fatalities reported across states was reduced, while drivers were more willing to take risks that included speeding, driving while impaired, and not using their seat belts. These drivers, along with a potential reduction in law enforcement and safety messaging, were identified as possible factors that created an environment favoring risky driving. The same finding was recently reported by Brodeur et al. (2020) who used difference-in-differences in order to evaluate the impact of Stay-at-Home orders on road crashes for five states in the United States and a 50% reduction in road crashes was identified. However, some of the conclusions delivered from Lin et al. (2020) were found to be different compared to the present ones. In particular, the impact of COVID-19 on road traffic safety in Los Angeles and New York was examined. Results indicated that the pandemic has disproportionately affected certain age groups and that nonfatal road crashes decreased, while the number of fatal crash cases remained the same during the pandemic (Lin et al., 2020).

8. Conclusions

This paper presented an investigative approach to quantify the impact of the COVID-19 pandemic on driving behavior using naturalistic driving data obtained from smartphone sensors and time series forecasting in Greece. The evaluation of the impact of COVID-19 was based on the comparison between observed values for three driving indicators (i.e., speed, speeding, and harsh braking) and forecasts based on the period before the coronavirus spread. Methodologically, the influence of COVID-19 was initially evaluated with explanatory XGBoost feature importance and was primarily modelled using seasonal ARIMA models, which have been a popular choice for transportation-related forecasting.

Results demonstrated the magnitude of the impact due to the COVID-19 lockdown, as it was observed that the "natural evolution" of the three aforementioned indicators was forecasted with major differences compared to the actual observations. Measurements regarding speed were the ones demonstrating the larger difference. The most reliable forecasting model for speed demonstrated that speeds increased by 2.27 km/h on average and up to 7.5 km/h on a national holiday day during the lockdown. Furthermore, the increase in speeds also assisted in manifesting an increase in harsh braking/100 km, which is supported by recent literature. The number of road crashes and road traffic fatalities and injuries decreased during the COVID-19 period (i.e., from March to

April 2020), compared to non-COVID-19 period (i.e., from January to February 2020).

Nevertheless, this paper is not without shortcomings. The developed models for speeding have limitations and do not seem to capture the trend and seasonality of the original time series effectively. More sophisticated models, such as deep neural networks (e.g., Convolutional Neural Networks (CNNs) or Long Short-Term Memory Networks (LSTMs)) could have a better fit on the time series data and provide better forecasts. Furthermore, rates for harsh acceleration events per km were not found statistically significant in this work, but the aforementioned sophisticated models could succeed in forecasting using these variables as well.

Future research should initially concentrate on comparing COVID-19 driving indicators from different countries so as to compare and contrast different effects. Furthermore, the development of more sophisticated models (as those mentioned in the previous paragraph), as well as multivariate forecasting models using Vector AutoRegression (VAR) in order to capture the interdependencies between time series should provide more insights on the impact of COVID-19 on driving behavior. Finally, more driving behavior indicator time series, such as the use of mobile phone during driving or aggressiveness levels, would also assist in quantifying the effects of lockdown on driving.

Acknowledgements

The authors would like to thank OSeven Telematics, London, UK for providing all necessary data exploited to accomplish this study.

References

Aloi, A., Alonso, B., Benavente, J., Cordera, R., Echániz, E., González, F., et al. (2020). Effects of the COVID-19 lockdown on urban mobility: Empirical evidence from the City of Santander (Spain). *Sustainability, 12*(9), 3870. https://doi.org/10.3390/su12093870.

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet, 395*(10228), 931–934. https://doi.org/10.1016/S0140-6736(20)30567-5.

Apple. (2020). COVID 19 - Mobility Trends Reports - Apple [WWW Document]. Accessed 11/06/20. https://www.apple.com/covid19/mobility.

Barnes, S. R., Beland, L. P., Huh, J., & Kim, D. (2020). The effect of COVID-19 lockdown on mobility and traffic accidents: Evidence from Louisiana (No. 616). GLO Discussion Paper.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., et al. (2016). mlr: Machine Learning in R. *The Journal of Machine Learning Research, 17*(1), 5938–5942.

Bisgaard, S., & Kulahci, M. (2011). *Time series analysis and forecasting by example*. Wiley & Sons.

Box, G. E., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control San Francisco. 1970, 1989.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Brodeur, A., Cook, N., & Wright, T. (2020). On the effects of COVID-19 safer-at-home policies on social distancing, Car Crashes and Pollution.

Budd, L., & Ison, S. (2020b). Responsible transport: A post-COVID agenda for transport policy and practice. *Transportation Research Interdisciplinary Perspectives, 6*, 100151. https://doi.org/10.1016/j.trip.2020.100151.

Carter, D. (2020) Effects of COVID-19 shutdown on crashes and travel in North Carolina. North Carolina Department of Transportation. Transportation Research Board: Traffic Trends and Safety in a COVID-19 World webinar: https://vimeo. com/425250264 Accessed 05/01/21.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Cheng, Z. J., & Shan, J. (2020). 2019 Novel coronavirus: Where we are and what we know. *Infection, 48*(2), 155–163. https://doi.org/10.1007/s15010-020-01401-y.

Chung, E., & Rosalion, N. (2001). Short term traffic flow prediction. Proceedings of the 24th Australian Transportation Research Forum, Hobart, Tasmania.

City of Toronto, 2020. City of Toronto Urges Drivers to Obey Rules of the Road. 2020 November 3, https://www.toronto.ca/news/city-of-toronto-urges-drivers-to-obey-rules-of-the-road/.

de Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives, 6*, 100150. https://doi.org/10.1016/j.trip.2020.100150.

De Vos, J. (2020). The effect of COVID-19 and subsequent social distancing on travel behaviour. *Transportation Research Interdisciplinary Perspectives, 5*. https://doi.org/10.1016/j.trip.2020.100121 100121.

Dettling, M. (2018). Applied Time Series Analysis - Lecture notes SS 2018.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a Unit Root. *Journal of the American Statistical Association, 74* (366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531.

DW. (2020). Germany: Traffic deaths fall to lowest point since reunification. https://www.dw.com/en/germany-traffic-deaths-fall-to-lowest-point-since-reunification/a-53556709.

Essi, I. D. (2018). Seasonal autoregressive integrated moving average (SARIMA) modelling and forecasting of inflation rates in Nigerian (2003-2016). 4(1), 1–14.

ETSC. (2020). COVID-19: Huge drop in traffic in Europe, but impact on road deaths unclear. https://etsc.eu/covid-19-huge-drop-in-traffic-in-europe-but-impact-on-road-deaths-unclear/.

Ghosh, B., Basu, B., & O'Mahony, M. (2005). Time-series modelling for forecasting vehicular traffic flow in Dublin. 85th Annual Meeting of the . . ., 12(00353), 1–22. https://doi.org/10.1109/JPHOTOV.2015.2434597.

Google LLC. (2020). Google COVID-19 Community Mobility Reports [WWW Document]. Accessed 11/06/20. https://www.google.com/covid19/mobility/.

Hellenic Statistical Authority (2021). Road traffic accidents - Monthly/October 2020 Accessed 05/01/21. Retrieved from https://www.statistics.gr/en/statistics/-/publication/SDT03/-.

Hendrickson, C., & Rilett, L. R. (2020). The COVID-19 pandemic and transportation engineering. *Journal of Transportation Engineering Part A: Systems, 146*(7), 1–2. https://doi.org/10.1061/JTEPBS.0000418.

Hipel, K., & McLeod, A. (1994). Chapter 12 seasonal autoregressive integrated moving average models. Developments in Water Science, 45.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

Hyndman, R. J., & Khandakar, Y. (2007). Automatic time series forecasting : the forecast package for R Automatic time series forecasting : the forecast package for R. June.

Hyndman, R. J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., & Wang, E. (2020). Package 'forecast'. Online] https://cran. r-project. org/web/packages/forecast/forecast. pdf.

Inada, H., Ashraf, L., & Campbell, S. (2020). COVID-19 lockdown and fatal motor vehicle collisions due to speed-related traffic violations in Japan: A time-series study. *Injury Prevention, 27*(1), 98–100.

Karlaftis, M. G., & Vlahogianni, E. I. (2009). Memory properties and fractional integration in transportation time-series. 17, 444–453. https://doi.org/10.1016/j.trc.2009.03.001.

Katrakazas, C., Michelaraki, E., Sekadakis, M., & Yannis, G. (2020). A descriptive analysis of the effect of the COVID-19 pandemic on driving behavior and road safety. *Transportation Research Interdisciplinary Perspectives, 7*, 100186. https://doi.org/10.1016/j.trip.2020.100186.

Kumar, S. V., & Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review, 7*(3), 1–9. https://doi.org/10.1007/s12544-015-0170-8.

Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Schubert, J., Bania, J., et al. (2020). The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of Travel Medicine, March*, 1–7. https://doi.org/10.1093/jtm/taaa037.

Lee, J., Porr, A., & Miller, H. (2020). Evidence of Increased Vehicle Speeding in Ohio's Major Cities during the COVID-19 Pandemic.

Lin, L., Shi, F., & Li, W. (2020). Assessing road traffic safety under COVID-19: Inequality, irregularity, and severity. arXiv preprint arXiv:2011.02289.

Lippi, M., Bertini, M., & Frasconi, P. (2013). Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning. *IEEE Transactions on Intelligent Transportation Systems, 14*(2), 871–882. https://doi.org/10.1109/TITS.2013.2247040.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297.

Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. *Transportation Research Part C: Emerging Technologies, 111*(January), 352–372. https://doi.org/10.1016/j.trc.2019.12.022.

Ma, T., Zhou, Z., & Antoniou, C. (2018). Dynamic factor model for network traffic state forecast. 118, 281–317. https://doi.org/10.1016/j.trb.2018.10.018.

Mogaji, E. (2020). Impact of COVID-19 on transportation in Lagos, Nigeria. *Transportation Research Interdisciplinary Perspectives, 6*, 100154. https://doi.org/10.1016/j.trip.2020.100154.

Molloy, J. (2020). MOBIS Covid19 Mobility Report.

Moovit. (2020). Impact of coronavirus (COVID-19) on public transit usage. Accessed 11/06/20. https://moovitapp.com/insights/en/Moovit_Insights_Public_Transit_Index-countries.

Neuburger, L., & Egger, R. (2020). Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: A case study of the DACH region. *Current Issues in Tourism*, 1–14.

Nielsen, D. (2016). Tree boosting with xgboost-why does xgboost win" every" machine learning competition? (Master's thesis, NTNU).

Papadimitriou, E., Argyropoulou, A., Tselentis, D. I., & Yannis, G. (2019). Analysis of driver behaviour through smartphone data: The case of mobile phone use while driving. *Safety Science, 119*(May), 91–97. https://doi.org/10.1016/j.ssci.2019.05.059.

Petraki, V., Ziakopoulos, A., & Yannis, G. (2020). Combined impact of road and traffic characteristic on driver behaviour using smartphone sensor data. Accident Analysis & Prevention, 144(November 2019), 105657. https://doi.org/10.1016/j.aap.2020.105657.

R Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: https://www.R-project.org/.

Roe, C. M., Rosnick, C. B., Colletta, A., & Babulal, G. M. (2020). Reaction to a pandemic: Social distancing and driving among older adults during COVID-19. *Journal of Applied Gerontology*. 0733464820966516.

Smith, T. G. et al. (2017). ARIMA estimators for Python. Http://Www.Alkaline-Ml.Com/Pmdarima.

Stavrakaki, A. M., Tselentis, D. I., Barmpounakis, E., Vlahogianni, E. I., & Yannis, G. (2020). Estimating the necessary amount of driving data for assessing driving behaviour. *Sensors (Switzerland), 20*(9). https://doi.org/10.3390/s20092600.

Stavrinos, D., McManus, B., Mrug, S., He, H., Gresham, B., Albright, M. G., et al. (2020). Adolescent driving behavior before and during restrictions related to COVID-19. *Accident Analysis & Prevention, 144*, 105686. https://doi.org/10.1016/j.aap.2020.105686.

Ting, C. Y., Tan, N. Y. Z., Hashim, H. H., Ho, C. C., & Shabadin, A. (2020). Malaysian road accident severity: Variables and predictive models. In Computational science and technology (pp. 699-708). Springer, Singapore.

TomTom. (2020). COVID-19 response | TomTom [WWW Document]. https://www.tomtom.com/covid-19/.

Tselentis, D. I., Vlahogianni, E. I., & Yannis, G. (2019). Driving safety efficiency benchmarking using smartphone data. Transportation Research Part C: Emerging Technologies, 109(June 2018), 343–357. https://doi.org/10.1016/j.trc.2019.11.006.

Vingilis, E., Beirness, D., Boase, P., Byrne, P., Johnson, J., Jonah, B., et al. (2020). Coronavirus disease 2019: What could be the effects on road safety? *Accident Analysis & Prevention, 144*, 105687.

Wang, W., Baumann, M., & Jiang, X. (Eds.). (2020). Green, smart and connected transportation systems: Proceedings of the 9th international conference on green intelligent transportation systems and safety (Vol. 617). Springer Nature.

Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data Mining and Knowledge Discovery, 13*(3), 335–364.

Wagner, E., Atkins, R. G., Berning, A., Robbins, A., Watson, C., & Anderle, J. (2020). Examination of the traffic safety environment during the second quarter of 2020: Special Report (No. DOT HS 813 011). United States. National Highway Traffic Safety Administration. Office of Behavioral Safety Research.

World Health Organization. (2020). Corona Virus 2019. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

Williams, B. M. (2003). Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results article. *Journal of Transportation Engineering, 129*(6).

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet, 395*(10225), 689–697. https://doi.org/10.1016/S0140-6736(20)30260-9.

XGBoost developer team (2019). XGBoost Documentation. Official website. Available from: https://xgboost.readthedocs.io/en/latest/index.html [Accessed 20/06/2020].

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. Finance Research Letters, April, 101528. https://doi.org/10.1016/j.frl.2020.101528.

Zhang, L., Ghader, S., Pack, M. L., Xiong, C., Darzi, A., Yang, M., & Hu, S. (2020). An interactive COVID-19 mobility impact and social distancing analysis platform. medRxiv.

Ziakopoulos, A., Tselentis, D., Kontaxi, A., & Yannis, G. (2020). A critical overview of driver recording tools. *Journal of Safety Research, 72*, 203–212.

Dr. Christos Katrakazas is a Research Associate at the Department of Transportation Planning and Engineering at the School of Civil Engineering of the National Technical University of Athens (NTUA). He holds a Civil Engineering Diploma from NTUA majoring in Transportation Engineering (2013) and a Ph.D. in safety of autonomous vehicles from Loughborough University (2017). After finishing his doctoral studies, he worked as a Research Associate at Loughborough University, and the Chair of Transportation Systems Engineering at the Technical University of Munich (TUM), Germany. He has more than 6 years of experience in several aspects of transport safety and engineering.

Ms. Eva Michelaraki is a Civil Engineer, PhD Candidate and Researcher at the Department of Transportation Planning and Engineering at the School of Civil Engineering of the National Technical University of Athens (NTUA). Since 2019, she holds a Civil Engineering Diploma from NTUA majoring in Transportation Engineering. She is pursuing her Doctoral Thesis, entitled: "Road accident risk factors and big data". She has more than 1 year of experience in several aspects of traffic engineering and road safety. The current focus of her work lies on driver behaviour, crash and statistical analysis and modeling. She has participated in 2 research projects and has published two (2) scientific papers (one (1) in scientific journal and one (1) in scientific conference proceedings after peer review).

Mr. Marios Sekadakis is a Civil – Transportation Engineer and Research Associate at the Department of Transportation Planning and Engineering at the School of Civil Engineering of the National Technical University of Athens (NTUA). Since 2020, he holds a Civil Engineering Diploma from NTUA, majoring in Transportation Engineering.

Dr. Apostolos Ziakopoulos, PhD, MSc DIC is a Research Associate at the Department of Transportation Planning and Engineering at the School of Civil Engineering of the National Technical University of Athens (NTUA). He holds a Civil Engineering Diploma from NTUA majoring in Transportation Engineering (2013). He also holds a Master of Science – DIC in Transport from Imperial College London and University College London (2014). His Ph.D. dissertation was in spatial analysis of road safety and traffic behavior from NTUA (2020). He has worked as a researcher for the University of Sheffield through the I.A.E.S.T.E. scheme. His main research interests involve road crash analyses, statistical and spatial modelling and traffic engineering. He has participated in 7 research and engineering projects and studies and has published 39 papers in scientific journals, textbooks and conferences.

Ms. Armira Kontaxi is a Ph.D. Candidate and Researcher at the Department of Transportation Planning and Engineering at the School of Civil Engineering of National Technical University of Athens. Since 2017, she holds a Civil Engineering Diploma from NTUA, majored in Transportation Engineering. She is pursuing her Doctoral Thesis, entitled: "Integrated support of driver traffic behaviour and safety by smartphone data". Her research primarily involves data analysis and statistical modelling in the field of traffic behaviour and safety.

Prof. George Yannis is a Professor in Traffic Safety and Management with particular focus on data management and analysis and currently Director of the Department of Transportation Planning and Engineering of the School of Civil Engineering at the NTUA. He has contributed extensively in more than 245 research and engineering projects and studies and in several scientific committees of the European Commission and other International Organizations (UN/ECE, OECD, WHO, World Bank, EIB, CEDR, ERF, UITP, ETSC, ECTRI, WCTR, TRB). He has published more than 560 scientific papers (169 in scientific journals) widely cited worldwide.

Journal of Safety Research 78 (2021) 9–18

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Implementation of participatory organizational change in long term care to improve safety

Dwayne Van Eerd [a,*], Teresa D'Elia [a], Era Mae Ferron [a], Lynda Robson [a], Benjamin Amick III [b]

[a] Institute for Work & Health, 481 University Ave, Toronto, ON M5G 2E9, Canada

[b] University of Arkansas Medical System, Fay W Boozman College of Public Health, 4301 West Markham, #820, Little Rock, AR 722 05, United States

ARTICLE INFO

Article history:
Received 3 March 2020
Received in revised form 2 February 2021
Accepted 5 May 2021
Available online 20 May 2021

Keywords:
Participatory ergonomics
Implementation
Long term care
Hazard reduction

ABSTRACT

Introduction: Long Term Care (LTC) facilities are fast-paced, demanding environments placing workers at significant risk for injuries. Health and safety interventions to address hazards in LTC are challenging to implement. The study assessed a participatory organizational change intervention implementation and impacts. *Methods:* This was a mixed methods implementation study with a concurrent control, conducted from 2017 to 2019 in four non-profit LTC facilities in Ontario, Canada. Study participants were managers and frontline staff. Intervention sites implemented a participatory organizational change program, control sites distributed one-page health and safety pamphlets. Program impact data were collected via Survey (self-efficacy, control over work, pain and general health) and observation (Quick Exposure Checklist). Interviews/focus groups were used to collect program implementation data. *Results:* Participants described program impacts (hazard controls through equipment purchase/modification, practice changes, and education/training) and positive changes in culture, communication and collaboration. There was a statistically significant difference in manager self-efficacy for musculoskeletal disorder (MSD) hazards between the control and intervention sites over time but no other statistical differences were found. Key program implementation challenges included LTC hazards, staff shortage/turnover, safety culture, staff time to participate, and communication. Facilitators included frontline staff involvement during implementation, management support, focusing on a single unit, training, and involving an external program facilitator. *Conclusion:* A participatory program can have positive impacts on identifying and reducing MSD hazards. Key to success is involving frontline staff in identifying hazards and creating solutions and management encouragement on a unit working together. High turnover rates, staffing shortages, and time constraints were barriers as they are for all organizational change efforts in LTC. The implementation findings are likely applicable in any jurisdiction. *Practical Application:* Implementing a participatory organizational change program to reduce MSD hazards is feasible in LTC and can improve communication and aid in identification and control of hazards.

1. Introduction

Long Term Care (LTC) facilities are fast-paced and demanding environments placing workers at significant risk for occupational-related injuries. In 2018, the Ontario healthcare sector had the third highest lost-time injury rate – overall as well as for sprains and strains (the category that captures musculoskeletal disorders (MSD)) (Workplace Safety & Insurance Board (WSIB), 2019). LTC healthcare support staff, such as personal support workers, dietary staff, nurses, and environmental workers, represent workers at high risk for occupational injuries. Personal support workers constitute a significant component of the LTC labor force and are the backbone of the LTC sector (Blair & Glaister, 2005; Canadian Research Network for Care in the Community (CRNCC), 2010). Healthcare support staff comprise approximately 70% of LTC sector employees. Personal support workers complete approximately 90% of all resident care provided (Zhang et al., 2011). The two primary hazards for LTC employees are slips, trips, and falls hazards and ergonomic hazards.

Health and safety interventions to address hazards in LTC are challenging to implement due to high staff turnover levels and the nature of the work with residents (Armstrong & Daly, 2004; Boakye-Dankwa, Teeple, Gore, & Punnett, 2017; Garg & Owen, 1992; Holmberg et al., 2013; Kurowski, Gore, Buchholz, & Punnett, 2012; Owen, Keene, & Olson, 2002). To date, hazard

* Corresponding author at: Institute for Work & Health, 481 University Ave, Suite 800, Toronto, ON M5G 2E9 Canada.

E-mail address: dvaneerd@iwh.on.ca (D. Van Eerd).

https://doi.org/10.1016/j.jsr.2021.05.002
0022-4375/

reduction interventions have yielded limited evidence of effectiveness (Kamioka et al., 2011; Kennedy et al., 2010; Wahlstrom, Ostman, & Leijon, 2012). Without well-designed and properly implemented hazard reduction programs, workplace safety hazards will continue to negatively affect workers and ultimately, resident safety.

Recent studies examining the implementation and effectiveness of participatory organizational interventions have shown some promise (Carayon et al., 2006; Framke, Sorensen, Pedersen, & Rugulies, 2018; Jakobsen, Aust, Kines, Madeleine, & Andersen, 2019; Kotejoshyer, Zhang, Flum, Fleishman, & Punnett, 2019; Rasmussen, Holtermann, Bay, Sogaard, & Birk Jorgensen, 2015; Schelvis et al., 2017; Van Eerd et al., 2018). Participatory programs can improve workplace conditions through participation, communication, and group problem solving (Carayon et al., 2006; Hignett, Wilson, & Morris, 2005; Rasmussen, Holtermann, Mortensen, Sogaard, & Jorgensen, 2013; Rasmussen, Larsen, Holtermann, Sogaard, & Jorgensen, 2014; van Eerd et al., 2010; Wilson & Haines, 1997). Kotejoshyer et al. (2019) and Jakobsen et al. (2019) have recently shown positive impacts of participatory organizational interventions in busy healthcare settings. Kotejoshyer et al. (2019) showed positive impacts on staff awareness and participation, and communication in the intervention sites even though there was little change in desired organizational outcomes. Jakobsen et al. (2019) implemented a participatory organizational intervention to increase the use of assistive devices in patient transfer and found an increase in communication and general use of assistive devices. However, the effectiveness of participatory programs has not been consistently established (Haukka et al., 2008; Pehkonen et al., 2009; Rivilis et al., 2008). Mixed effectiveness findings may be related to the implementation process (Cole et al., 2009; Driessen, Proper, Anema, Bongers, & van der Beek, 2010; van Eerd et al., 2010; Wells, Laing, & Cole, 2009). Inconsistent findings on the effectiveness of participatory programs in reducing occupational hazards requires a research focus on the implementation process.

The study objective was to follow the implementation of a participatory organizational change intervention and assess the program implementation effects on important intermediate outcomes. A mixed methods approach was chosen to minimize LTC worker burden and allow detailed data collection on participatory change program implementation. In this study, we used Meyers et al. (2012) Quality Implementation Tool (QIT) as a framework for assessing program implementation.

2. Methods

2.1. Study design

This was a mixed methods study of the implementation of a participatory organizational change intervention in non-profit LTC facilities with a control group comparison. Quantitative and qualitative data were collected regarding program impacts and qualitative data were collected about the program implementation process.

2.2. Participatory organizational change intervention

The Employees Participating in Change (EPIC) program was developed to address occupational hazards in healthcare settings with a relatively mature health and safety infrastructure, including a well-functioning internal responsibility system, to ensure that organizations are ready to support a change process. In Canadian jurisdictions, an internal responsibility system is a consistent underlying philosophy of occupational health and safety (OHS) legislation whereby both employers and employees work together to ensure workplace health and safety. EPIC program implementation involves five steps: (1) selection of a program champion to lead implementation activities and act as the main liaison between program committees and external program facilitators; (2) formation of a participatory steering committee (includes site administrators, departmental supervisors, a joint health and safety chairperson, HR, and union representatives) that supports implementation success ensuring key organizational performance metrics are achieved; (3) formation of a participatory change team (composed mainly of frontline staff) responsible for hazard identification, risk assessment, and hazard control activities; (4) provision of training, mentoring, and coaching from program facilitators; and, (5) the LTC facility further develops the OHS management system to build accountability for change and to support change processes. EPIC was implemented in two intervention sites, with implementation steps guided by a trained facilitator and safety consultant who was one of the program developers. The facilitator conducted a pre-implementation site review to ensure the sites were able to implement the program. The facilitator provided all the training and facilitated team meetings at the onset of implementation, mentored both the steering committee and the change team, gradually decreasing involvement over six months, and then became available as needed for consultation for the remainder of program implementation (approximately another six months). The overall goal was to implement a sustainable program.

2.3. Control site intervention

The control site intervention was chosen to be more broadly focused on OHS without participatory engagement. The control intervention was designed to: (1) help improve participation rates in the control groups by providing a nominal intervention and encourage leadership engagement, and (2) allow a comparison between similar facilities receiving OHS interventions over time. The concurrent comparison group allowed us to better attribute any changes observed in the intervention to EPIC and not secular trends.

The two control sites received one-page OHS information pamphlets at approximately Time 1 (denoted as C1 in Fig. 1) for distribution: (1) managers, supervisors, and joint health and safety committee (JHSC) team members received *Empowerment and Self-Protection: Occupational Health and Safety for Workers*, *Health and Safety Management Systems* and *Caught in the Middle: The Supervisor and Occupational Health and Safety* at the start of the EPIC program; (2) all staff received *Hazards in Health Care Workplaces* at approximately Time 1 (C1); (3) all staff received *Occupational Health and Safety is Everyone's Business* at approximately Time 2 (C2); and (4) all staff received *An Introduction to the Joint Health and Safety Committee* at approximately Time 3 (C3). The timing of the control intervention (informational pamphlets) delivery was chosen to coincide with key program implementation time points at the intervention sites. Once all Time 3 data collection was completed the remaining study time was for data analysis and interpretation. The study began in June 2017 and ended in June 2019. See Fig. 1 for a timeline of the study implementation and data collection.

2.4. Sample

Four non-profit LTC facilities in Ontario were recruited to participate. Intervention and control facilities were matched on municipality, facility age, and client composition. Intervention and control sites were not randomly chosen but volunteered, which was necessary to ensure willingness to implement the intervention program. EPIC program participants at each intervention

Fig. 1. Study timeline showing intervention and data collection.

facility were decided upon through a voluntary and participatory approach. Program champions and site leaders were asked to be part of the EPIC program steering committee. The steering committee with support of the facilitator then selected the target area for program implementation and sought relevant volunteer frontline staff to participate as change team members. These two groups (steering committee and change team) were each to be composed of 9–12 members who participated in program team meetings, as well as research interviews, focus groups, and self-report surveys. Members of the existing JHSC of the participating sites were approached or volunteered to be members of the steering committee or change team. A JHSC is an advisory body that helps to stimulate or raise awareness of health and safety issues in the workplace, including recognizing and identifying workplace risks and developing recommendations. In Ontario, a JHSC is to be composed of worker and employer representatives and must be in place in all workplaces with greater than 20 employees.

The intervention sites chose to focus on one unit of the facility and include all staff (from different departments) that work on the target unit. Study participants were managers and frontline staff recruited from steering committee members, change team members, and a random selection of all frontline staff from the participating unit (using site staff lists) not involved in either the committee or team. Additional frontline staff from the unit were recruited to replace those who were not available at data collection time points. Our goal was to recruit 30 participants from each intervention site (10 steering committee members, 10 change team members and 10 frontline staff) and control site.

At control sites we randomly selected a similar number of managers and frontline staff from matching units/departments so that the intervention and control participant groups would be matched by position/department. To be eligible for study participation, staff had to work at least 30 hours a week and be able to complete a self-report survey in English (Grade 6 reading level). The participants from each site were followed over time, and invited to participate in self-report surveys, observations, and at the intervention sites, interviews, and focus groups.

Project research staff met with EPIC team members at their first meeting, provided them with the study details, and asked them to read and sign an informed consent. The randomly selected staff were asked to attend a meeting where the study details were described. Consent forms were read and signed prior to data collection. If a potential study participant declined to participate, another staff member was randomly selected to ensure the project started with the same number of consenting participants at each facility.

2.5. Data collection

Data were collected using three different methods: (1) self-administered questionnaires; (2) staff observations; (3) interviews and focus groups. The questionnaires and observations concerned program impacts and the interviews and focus groups addressed the program implementation process (Pinnock et al., 2017).

2.5.1. Program impacts

Self-Administered Questionnaire: A self-report survey was administered at baseline pre-intervention (Time 1 – pre-implementation), 6 months (Time 2 – mid-point of implementation), and 12 months (Time 3 – end of implementation). There were separate questionnaires for managers and frontline staff, containing the same questions but directed to managers or workers. All questionnaires were at a grade six reading level or below. The questionnaire took less than 10 minutes to complete and was administered to the steering committee and change team members (who completed them during meetings), and directly to the frontline staff participants at each facility. All survey participants received small incentives (typically coffee and snacks provided at meetings or in the break room). All surveys were completed on paper.

Self-Efficacy: Based on an existing measure by Greene, DeJoy, and Olejnik (2005), this 7-item scale assesses self-efficacy in identifying MSD hazards and managing workplace changes in individuals regarding MSD prevention in the LTC workplace. This scale had high internal consistency (alphas > 0.8) in pilot work. Scores from 1-7 were summed for all answered items and divided by the number of items to create a scale score varying between 1 (not confident at all) and 7 (highly confident). Mean scale scores were used for all comparisons.

Pain in past week: A single visual analog scale item indicating level of pain in the past week on a 0 (no pain) to 10 (worst pain ever) scale. Higher scores indicate higher pain.

Control over work: We asked participants how much say they had in what they do at work and how much freedom they had in how they do their work. Each of these was a single item scored from 0-3. Item scores were used for all comparisons.

General health: In addition, we asked participants to rate their general health (on a 5-point scale), which we expected to remain stable (on average) in both groups over approximately a 12-month period. Mean scores were used from all comparison.

2.5.1.1. Staff observations. The Quick Exposure Check (QEC) (David, Woods, Li, & Buckle, 2008) tool was used to quantify the postural hazards at the LTC sites. The QEC covers a variety of physical risk factors including posture, load, frequency of movement, and vibration for four main body regions (back, neck, shoulder/arm, wrist/hand). The tool also captures psychosocial risk factors (e.g., work stress, pace of work) by consulting with workers during the observation. A higher QEC score indicates greater risk. The observers (TD and EMF) were trained in using the QEC by the lead author. Training consisted of using the QEC to evaluate mock-up workplace scenarios. Observers' QEC responses were compared to the lead author's and each other; differences were discussed until consensus was achieved. The QEC has good validity and reasonable reliability for an observational tool (David et al., 2008).

2.5.2. Program implementation

2.5.2.1. Interviews/focus groups. Interviews were used to evaluate EPIC program implementation. At each intervention site, members of the steering committee (program champion, JHSC committee worker representatives, participating department supervisors, facility administrators), and change team members (program co-champion, workers from the participating department) were invited to participate in interviews at three time-points (Time 1 – pre-implementation, Time 2 – mid-point of implementation, and Time 3 – end of implementation). As well, frontline staff who did not participant in the EPIC program implementation, were recruited to participate in interviews to gain a broader unit perspective. Interviews were conducted either in-person or over the phone. A semi-structured interview guide was followed. Interview questions were based on Meyers et al.'s (2012) QIT framework focusing on the structural and supportive aspects of the implementation including facilitators and barriers, and program satisfaction.

Two focus groups were conducted at one intervention site: one for managers and one for frontline staff. Separate groups were done to ensure that frontline staff would be able to speak openly without the presence of supervisors and managers. These focus groups were done in place of Time 3 interviews for this site due to a delay in program implementation (caused by factors unrelated to the intervention program or study protocol).

The interviews and focus groups explored participants' experiences regarding EPIC program activities and achievements. Participants were asked how the participatory program was implemented and what they felt they accomplished regarding hazard identification and reduction.

Ethics approval was obtained from the University of Toronto Research Ethics Board.

2.6. Data analysis

Survey data were analyzed descriptively as well as with an exploratory approach using regression models with robust covariance estimators to correct for the dependence between repeated observations at the three time points. Within and between (intervention and control) group differences were examined for self-efficacy, control over what and how work activities are performed, pain, and general health. We hypothesized: self-efficacy would improve over time in the intervention group and remain stable in the control group; pain levels would decrease over time in the intervention group and remain stable in the control group; and control over how work is done and what work activities are done would increase in the intervention group and remain stable in the control group. We did not expect to see a change in the general health measure across time in either group since the third measurement is immediately at the end of program implementation. Intervention and control group characteristics were analyzed descriptively. In the regression models, we examined the interaction between study arm (control vs. intervention) and time (Time 1(baseline), Time 2, Time 3). Separate analyses were conducted for managers and frontline staff. Statistical significance was assessed at the $p < 0.05$ level with 95% confidence limits. The observational outcomes were QEC scores (David et al., 2008). The QEC scores are based on combinations of risk factors (posture, weight, and duration) identified by the observer for each body area. These scores represent the potential relationship between the increased level of exposure and potential health outcomes. Summed scores range from 4 to 56 (with higher scores indicating higher risk) depending on the body area and risk factors observed (David et al., 2008). We examined within group differences and also between (intervention and control) group differences to see if postural hazards changed over time and between groups. All data were analyzed using Statistical Analysis Software (SAS®) V9.3.

Interviews and focus groups were transcribed and de-identified. Interview transcripts were coded using conceptual categories derived from the Meyers et al. (2012) implementation framework adapted for OHS program implementation. The analysis of interviews was iterative and reflexive. Additional codes were developed for new concepts and emergent themes. Interview coding was done independently by two coders using a coding guide until the coding guide no longer changed. Subsequent coding was completed by a first coder and checked by a second. Coding was focused on the barriers and facilitators to implementation. Data reduction and display were guided by methods described by Miles and Huberman (1994) and Saldana (2015).

3. Results

There were 132 participants with 65 participating from the control sites and 67 participating from the intervention sites. A participation rate is not calculable as the managers and frontline staff on the steering committees and change teams were recruited by the workplace. A description of the study participants is provided in Table 1. There are managers and frontline workers from multiple departments who were study participants as both intervention sites chose to implement EPIC to address MSD hazards in a single unit and include all departments that work on the unit.

Table 1
Participant characteristics.

	Control sites	Intervention sites
Facility Size	Site C1: 150 beds Site C2: 160 beds	Site I1: 300 beds Site I2: 270 beds
Number of participants	65	67
Staff type	Managers: 17 Frontline: 48	Managers: 19 Frontline: 48
Employment status (Full time or Part time/Casual)*	Full time: 53 Part time/Casual: 12	Full time: 47 Part time/Casual: 16
Shift worked** (Day/afternoons/nights)	Days: 41 Afternoons/evenings: 6 Nights: 0 Combination: 22	Days: 29 Afternoons/evenings: 6 Nights: 0 Combination: 23
Departments represented (%)	Personal support workers: 15 Nursing (RNs/RPNs): 15 Dietary, maintenance, housekeeping and recreation: 17 Management: 18	Personal support workers: 22 Nursing (RNs/RPNs): 12 Dietary, maintenance, housekeeping and recreation: 16 Management: 17

*there were missing responses in the intervention site surveys.
**multiple responses could be selected therefore will not add up to total.

Table 2
Survey results for frontline staff participants.

Variable mean (standard deviation) n	Control group T1 N = 43	Control group T2 N = 42	Control group T3 N = 24	Intervention group T1 N = 39	Intervention group T2 N = 42	Intervention group T3 N = 21	*P*-value** Group by Time
Self-efficacy	4.90 (1.1) $n = 43$	5.15 (0.9) $n = 38$	4.83 (1.3) $n = 24$	4.50 (1.2) $n = 39$	4.67 (1.2) $n = 40$	4.61 (1.3) $n = 21$	0.8461
Control over *what* activities are done	2.47 (0.7) $n = 43$	2.39 (0.7) $n = 38$	2.33 (0.7) $n = 24$	2.13 (0.7) $n = 38$	2.18 (0.8) $n = 40$	1.81 (0.8) $n = 21$	0.6973
Control over *how* activities performed	2.14 (0.7) $n = 43$	2.11 (0.7) $n = 38$	2.29 (0.6) $n = 24$	1.92 (0.6) $n = 37$	1.98 (0.7) $n = 40$	2.14 (0.7) $n = 21$	0.7179
Pain	3.86 (2.7) $n = 42$	3.95 (2.8) $n = 37$	4.32 (2.7) $n = 22$	3.29 (2.3) $n = 35$	4.08 (2.6) $n = 40$	4.40 (2.6) $n = 20$	0.6010
General health	4.08 (0.7) $n = 39$	4.00 (0.7) $n = 37$	4.08 (0.7) $n = 24$	3.95 (0.7) $n = 37$	3.92 (0.9) $n = 39$	4.05 (0.6) $n = 20$	0.5134

Table 3
Survey results table – for managers.

Variable mean (standard deviation) n	Control group T1 N = 14	Control group T2 N = 16	Control group T3 N = 12	Intervention group T1 N = 18	Intervention group T2 N = 19	Intervention group T3 N = 13	*P*-value** Group by Time
Self-efficacy	5.18 (1.0) $n = 14$	5.47 (1.0) $n = 16$	4.97 (0.9) $n = 12$	4.77 (1.1) $n = 18$	5.15 (1.0) $n = 19$	5.49 (1.0) $n = 13$	0.0488
Control over *what* activities are done	2.71 (0.5) $n = 14$	2.31 (0.7) $n = 16$	2.67 (0.5) $n = 12$	2.69 (0.5) $n = 16$	2.61 (0.5) $n = 18$	2.77 (0.4) $n = 13$	0.1158
Control over *how* activities performed	2.71 (0.6) $n = 14$	2.31 (1.0) $n = 16$	2.50 (0.7) $n = 12$	2.69 (0.5) $n = 16$	2.78 (0.4) $n = 18$	2.54 (0.5) $n = 13$	0.1528
Pain	1.93 (2.4) $n = 14$	2.50 (1.9) $n = 14$	2.42 (2.2) $n = 12$	4.25 (2.8) $n = 16$	4.00 (3.1) $n = 18$	3.23 (1.7) $n = 13$	0.3420
General health	3.93 (1.1) $n = 14$	3.94 (0.9) $n = 14$	4.08 (0.5) $n = 12$	3.81 (0.9) $n = 16$	4.06 (0.6) $n = 18$	4.31 (0.6) $n = 13$	0.7461

3.1. Program impacts

3.1.1. Self-report survey

Survey results for frontline staff revealed no significant group by time interactions (see Table 2). There is little change over time for the intervention group in any of the measures for frontline staff.

Survey results for managerial staff revealed a significant group by time interactions for self-efficacy (Chi-square 6.04, $p = 0.049$; see Table 3). Control group self-efficacy scores did not increase over time whereas for the intervention group, self-efficacy scores increase over time. Regression analysis revealed no significant group by time interactions for all other measures (p values ranged from 0.12 to 0.75).

3.1.2. Observations

There is no statistically significant interaction effect (group by time) for any QEC measure (p values ranged from 0.23 to 0.85). However, three of the six measures (back (motion), shoulder/arm, wrist/hand) favored the intervention sites (see Fig. 2).

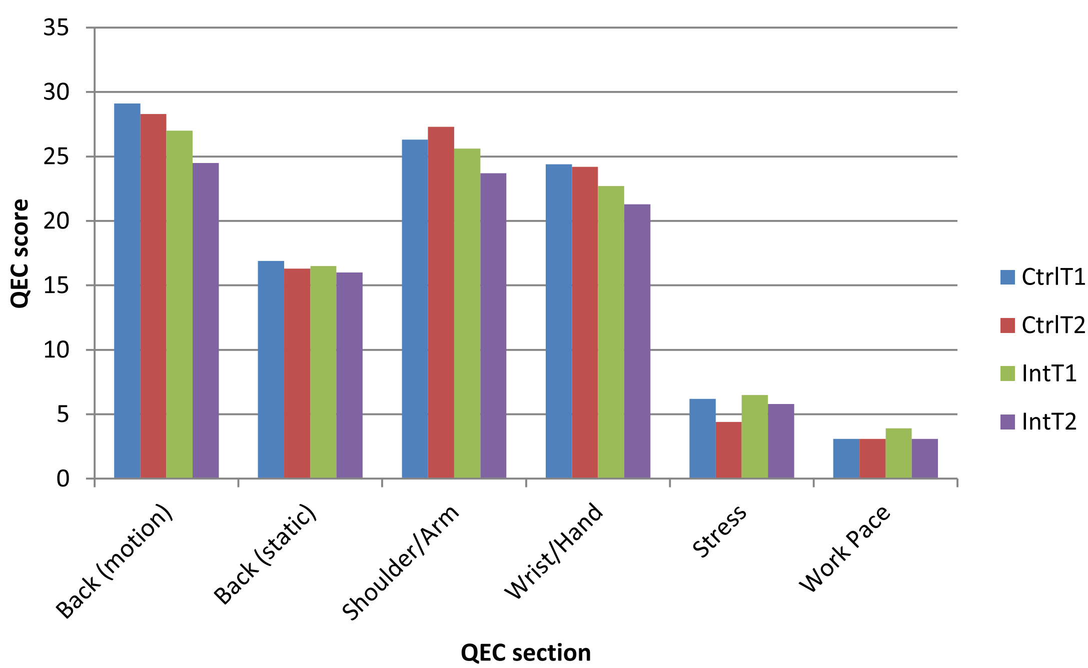

Fig. 2. QEC scores by site. (Max scores: Back (motion) = 56; Back (static) = 40; Shoulder/Arm = 56; Wrist/Hand = 46; Stress = 16; Work Pace = 9).

3.1.3. Interview/focus group: hazard changes implemented

Participants were asked to describe any changes as a result of EPIC program implementation. They reported the EPIC program implementation resulted in positive workplace changes. The implementation in both intervention sites targeted a single unit in the facility. Therefore, any solutions implemented on the unit could have originated in the housekeeping, maintenance, dietary, nursing, or recreation departments. Participants from both sites indicated that pushing, pulling, and lifting hazards were of concern. Consequently, many changes described were related to reducing hazards from pulling, pushing, or lifting. These included acquiring new (or modifying existing) equipment such as carts. There were descriptions of new carts to move laundry to and from laundry facilities. The new linen carts used on the unit were lighter and easier to pull. The new carts and bins for dirty laundry had a lower edge to reduce lifting of loads above shoulder level. Participants also mentioned new medication carts that were easier to pull with lower shelves to reduce reaching and improve visibility when moving through the hallways. There were also changes noted for garbage carts, which included acquiring an electric cart to pull the garbage bins outside of the facility. Large garbage bins were acquired that would fit under the existing garbage chutes, ensuring that the garbage was not spilled, which would require additional lifting and clean up. These large bins were dump-able to reduce forces when emptying in outdoor bins.

Additional changes were implemented to reduce push and pull forces such as adding wheels to large dining tables so they could be moved more easily for recreation activities that require space. Participants also reported reconfiguration of spaces and equipment including changing laundry and storage room layouts to reduce lifting and carrying of loads. One site described changes to decrease the distance between carts and shelves for linens to reduce lifting and carrying (or tossing) of heavy linens. These changes also reduced twisting motions with heavy loads. Shelving units were reorganized to reduce lifting of loads above the shoulder level. Higher shelves were used to store equipment that was not used frequently. At one site, the height of the dryers was adjusted to reduce bending to lift loads. In the dining area, one site added faucets to the steam tables to eliminate the need to fill them with buckets of water.

In addition to equipment purchase, modification, or reorganizing the workspace, there were education initiatives for proper lifting techniques implemented. The education was provided by onsite physiotherapy staff thereby reducing costs. The increased training opportunities for proper lifting techniques occurred on different days and different shifts to cover most staff. Participants reported that there was an increased awareness of proper lifting techniques. The education/training programs were not implemented in isolation but accompanied new/modified equipment and practice changes. Practice changes included stacking chairs to lower heights to reduce lifting above the shoulder, providing keys to the elevator for recreation staff to reserve an elevator so that there was less rush to get clients to their unit/floor. This helped resolve an issue of recreation staff pushing/pulling more than one client at a time as the rush to reach the elevator was resolved. Staff were also trained on how to tie garbage bags. Both laundry and garbage staff were asked not to fill the bags to full capacity (or overfilled!). Interviewees reported these two changes resulted in lower loads and less spills. A practice change related to volunteer scheduling was implemented to have more volunteers available when there was a need to move larger numbers of clients.

More broadly, implementing EPIC increased staff communication and collaboration. Participants consistently remarked there was increased MSD hazards communication within health and safety committee meetings and communication of solutions with staff. It was noted that MSD hazards were more consistently on the health and safety committee agenda and that there was engagement with staff to determine where changes could be made to reduce MSD hazards. Additionally, interviewees indicated there was increased awareness of near-misses and hazardous condition reporting. Often respondents noted that this led to more proactive hazard identification and solution suggestions.

3.2. Program implementation

3.2.1. Implementation facilitators and barriers themes

3.2.1.1. Theme: LTC sector-specific challenges. Respondents described some *challenges that they felt were specific to LTC*. The primary OHS challenges noted by both frontline and managerial staff concerned *pushing, pulling, and lifting activities*. These activities could involve various types of carts (laundry, food, garbage, tools) or the movement of clients in wheelchairs. Pushing and pulling activities were often linked with MSD experiences. Key aspects of the challenges of pushing and pulling were a shortage of staff (related to number of clients) and long distances between locations. The shortage of staff meant that individuals would often try to move more than they should in a given timeframe.

> Quote: "I have too much persons who are hard to move and it's affecting me as a worker, ... when you break [your body] in the next 5 years it's whatever, it's your body, we can't afford to get more staff on the floor," Frontline FG

Related to *staff shortage*, LTC staff consistently described staff turnover and absences as a common concern. Staff turnover and absence challenges were noted for day to day workload and was also a concern for implementing OHS programs such as EPIC. Absences from committee meetings and having new staff come to the team were considered somewhat disruptive and tended to slow the implementation progress. The concern about staff changes were consistent over time as one manager noted about sustaining the participatory program was at times challenged by a recent administrator change.

> Quote: "we've had so many changes I can't even keep track, we just got a new administrator who has just started and [Um hmm] we have a new assistant administrator, everybody here is brand new, so we haven't had, we've had so much on our plates, ..." Manager 10909

Respondents' concerns about safety and *worker health* most often concerned MSDs. They often referred to MSD hazards related to force and posture. In addition, some staff raised concerns about stress levels and overarching concerns about resident safety over staff safety.

> Quote: "Certainly from the perspective [of] resident care equipment and moving to a better product, a more reliable product, ... I mean if you are talking about pushing or moving shower chairs and that, a better quality that makes it easier to transport a resident." Frontline 10201

Some staff also raised issues regarding *culture* and staff dynamics. Frontline staff often felt that management did not listen to or respond to the issues frontline staff brought forward, which was considered to be a longstanding problem. Another issue all staff noted was related to a lack of engagement and uptake of practice or equipment changes from frontline staff. Here, respondents noted a culture of resistance to change, even though many felt the changes were likely to be beneficial.

> Quote: "for frontline workers, I have too much persons who are hard to move and it's affecting me as a worker, and I need more staff, what I will get back if I speak out is, we don't have it in our budget so basically you are telling me, when you break up your-

self in the next 5 years it's your body, we can't afford to get more staff on the floor, so it's going to deaf ears" Frontline FG.

3.2.1.2. Theme: implementation barriers/facilitators. Respondents were asked specifically about the barriers and facilitators related to EPIC implementation. Early in EPIC program implementation (Time 1) all staff noted that a major *barrier* was *devoting time to the program and meetings*. Getting people to meetings was a challenge. This issue seemed to be related to staff workload issues and staff turnover. In later stages of implementation (Time 2 and Time 3) the challenges were linked to issues of communication. One aspect of communication was between team members and staff to keep everyone "in the loop' about proposed changes as well as their role in the changes. Another aspect of communication mentioned at Time 3 was managerial staff felt that there could have been better communication about expectations about their role throughout the implementation process. However, staff and administrator turnover was a key element in the perceived lack of communication about roles. At Time 3 there was still some mention of time challenges, but this was related more to how to sustain the program than about working on solutions or changes on the unit.

Quote: "I think just resources, just making sure that people are available is the biggest thing" MGR 22109

Implementation *facilitators* were mentioned mostly in Time 2 and Time 3. The main facilitators mentioned were related to having *frontline staff involvement* throughout the program implementation. Managerial staff felt that the frontline staff were empowered by being involved in EPIC from the beginning.

Quote: "The frontline, having the frontline staff engaged from the beginning yeah. Yeah, that was a big help because they kind of held together even when our team was falling apart, when we had so many changes happening and leadership changes, ... it was really trying for us as the leadership team, but they were the core people so they were still there and still engaged." Manager FG

Another facilitator was the addition of the EPIC process as a standing agenda item on the monthly OHS meetings. The inclusion of the participatory program in standing procedures was consistently noted in the later stages of implementation. This was considered important for continuity and to ensure that the process of hazard identification and control was sustained in the organization.

3.2.1.3. Theme: Organizational level barriers/facilitators. Overall, respondents felt that there was good *support* from upper management for the EPIC program. Managerial staff noted that they were involved and trying to be responsive as the program was being implemented. Some managers noted they were initially concerned about the types of recommendations they might get from staff but in later stages noted that the changes suggested were not overly costly. Those that noted less support from management often pointed out that the steering committee and change team structure of EPIC allowed for changes to be proposed even when upper management support was not considered strong. There was no indication of change over time with respect to management support during EPIC implementation. It was perceived as present at the beginning of implementation and consistent throughout.

Quote: "I understand what the staff are going through, I understand what's happening in this building and I really see the need, the need for change and I think this is an excellent way for me to go forward to advocate for those things that are outside of the scope of the staff other than to recommend and we require some assistance at a divisional level ... because I think there are some things that we can do that can impact change, but I think there's some very realistic things that staff have identified that would, would require a, a higher level of, of involvement to ensure that they can have a safe effective working environment." Manager FG

Both intervention sites chose to *implement EPIC in a single unit* and include staff from all departments that work on the unit. Therefore, the participatory approach involved problem solving among different groups of workers. Both sites felt this was better than focusing on a single department. Throughout the implementation process comments about the focus unit reflected the opportunity to have frontline staff from different departments work together to solve MSD hazards. Managerial and frontline staff remarked that they saw improved communication and less 'silos.'

Quote: "so by getting everybody to give input hopefully people are also going to give support to each other throughout it and there will be less of people working in silos, like, this is my job, I do this." Manager 10,809

3.2.1.4. Theme: Benefit of a participatory process for organizational improvement. The main *benefit* of a participatory approach noted by both managerial and frontline staff was the opportunity for *frontline staff to engage* and describe hazards as well as discuss potential solution ideas. Many reported that the increased staff engagement and the opportunity for increased communication both to managers, as well as among frontline staff, was particularly beneficial. Over time, participants noted that the increased communication translated to action and changes being made in their work areas. Frontline staff felt encouraged to bring their suggestions forward. Team meetings were considered positive spaces to address concerns. Managers noted that opportunities for "continuous feedback" were very helpful when trying out potential solutions such as new equipment or procedures.

Quote: "I just like the fact that you know, we, we felt like we were being heard. That was you know, the, the biggest thing." Frontline staff 21701

The *MSD training* provided as part of EPIC implementation was considered a facilitator when it was received. Prior to program implementation, individuals on EPIC teams (both managers and frontline staff) reported getting training from the program facilitator. Frontline staff often reported not having much prior MSD training, or noted that the training they received prior to EPIC was recently moved online, which they felt was not as effective. At later stages in implementation (Time 2 and Time 3), frontline staff noted having some interaction with the EPIC facilitator related to training for hazard identification and solutions. Those that received training from the EPIC facilitator found it to be quite useful. Those that reported not receiving any training or getting only online training were somewhat frustrated.

Quote: [EPIC Facilitator] has been very good in explaining all of that kind of stuff and what we could be looking at, ..., in our meetings he has brought forward a lot of examples and um, what we should be looking for". Frontline staff 23049

It was clear that the *EPIC facilitator* was a valued part of the EPIC program implementation. All staff noted that he provided important training that was required and helpful throughout program implementation. Participants also noted that he provided important guidance about the implementation process (steps) and consistently pointed out the importance of frontline staff involvement and engagement. At later stages of implementation, he was also able to provide specific information about MSD hazard identification and control when requested.

> Quote: "I think having [the EPIC facilitator] as a resource was, was really, really helpful I, I mean I, I would call him up and I'd say do you have any information about this information or I am having a hard time with this and, and he was, he was great, he was fantastic like he, he was very key to the success of that program." Manager 22209

4. Discussion

The results from this mixed-methods implementation study suggest that an organizational change intervention such as EPIC can be successfully implemented in busy LTC facilities. We examined program impacts on important outcomes related to MSD hazard identification and control as well as program implementation. Regarding program impacts, our results suggest that the EPIC program had a positive impact on managers' self-efficacy regarding MSD hazards and a reported impact on MSD hazards. The increase in operational leadership confidence to address MSD hazards is important for program implementation but also for program impact and sustainability. However, we did not find that the program impacted other outcomes related to pain, control over work, general health, or observed postural hazards. It is not surprising that we observed few program impacts based on survey outcomes. The study was not well-powered, nor was the study long enough in duration to allow the EPIC program to develop hazard controls that could affect many outcomes (Abildgaard et al., 2019; Driessen et al., 2010; Gupta et al., 2018; Rasmussen et al., 2017).

Important barriers to consider for program implementation include LTC specific factors related to staff shortages and turnover, high levels of MSD hazards, and a non-supportive safety culture. Challenges related specifically to implementing a participatory intervention were securing sufficient staff time to be involved and ensuring there is a good level of communication about the intervention and role expectations for all staff involved. Similar challenges have been reported in the literature (Andersen & Zebis, 2014; Driessen et al., 2010; Murta, Sanderson, & Oldenburg, 2007; Rasmussen et al., 2017; van Eerd et al., 2010). An important facilitator to implementing a participatory approach is early frontline staff involvement. Early involvement helps to maintain engagement throughout the implementation process. Organizational factors important for a successful implementation process include: managerial support, incorporating good training as a starting point to implementation, and having a program facilitator and/or champion involved. These findings are consistent with other research that has evaluated participatory approaches (Abildgaard et al., 2019; Haines & Wilson, 1998; Haines, Wilson, Vink, & Koningsveld, 2002; Hignett et al., 2005; Rasmussen et al., 2015; van Beurden, Vermeulen, Anema, & van der Beek, 2012; van Eerd et al., 2010).

Staff experiences highlighted the participatory program was implemented as planned. Staff interviewees expressed satisfaction with the program and implementation process. Study participants noted that there was increased awareness and communication about MSD hazards and a change in safety culture. Previous studies have also noted worker and management satisfaction with the implementation of participatory approaches even when there have been few statistically significant changes in health or symptom outcomes (Abildgaard et al., 2019; Gupta et al., 2018). Though beyond our focus on implementation, we note that study respondents felt reducing MSD hazards and working more safely would have a positive impact on the quality of care in LTC sites. More research is required to explore the link between employee health and quality of care in LTC.

Further research is necessary to determine the effectiveness of participatory interventions. Our research shows that the choice of outcomes is paramount. The typical outcomes of lost-time claims and pain or symptoms may not be appropriate. Measures of engagement, hazard controls, communication, and culture may be more important to consider in the controlled trials since health changes may take longer to observe. In order for changes to occur in MSD symptom and lost-time claim outcomes, participatory programs must be sustained in workplaces. Longer duration studies with well implemented participatory interventions (programs) are necessary to understand overall effectiveness and contribute new knowledge to the scientific evidence. Proper implementation of programs and practices in real world settings is key for work and health research. The continued research of poorly implemented interventions does a disservice to work and health research and to improving workplace health and safety.

5. Strengths and limitations

The strengths of the study include: using a mixed methods design to collect and analyze multiple and rich sources of data on the implementation process, evaluating program implementation using an established implementation framework, and incorporating a concurrent comparison group, providing more confidence that program implementation impacts are actually due to the intervention. A process evaluation paper is being prepared describing the implementation framework and process outcomes in more detail.

Our primary limitation is a relatively small sample size decreasing the ability to detect program impacts quantitatively. It also prevents us from accounting for the nesting of individuals in workplaces, which may have increased the risk of a Type 1 error. However, we purposely restricted the sample size to ensure that we could collect good quality data but keep the burden on participants and sites low. We developed and tested data collection procedures in pilot work to ensure we collected good data with low burden. We also collected little information about study participant characteristics. This was done to reduce the burden on study participants and participating sites given busy schedules and challenging work. In busy workplaces, researchers must keep surveys short and conducting interviews and observations are key to collecting important data, even though it is more expensive and resource intensive for researchers. The relatively short follow-up time in this study also limits our ability to robustly evaluate program impacts. However, we chose outcomes that we considered could change during the study period, based on a previous study (Van Eerd et al., 2018). Overall, the proposed study resulted in rich data to increase our understanding about implementation.

Importantly the implementation of a participatory approach was possible in busy workplaces with staff shortage and turnover challenges. This suggests that a participatory program could be implemented in other jurisdictions. In addition, subsequent to implementation, the intervention sites began to incorporate key participatory elements of the EPIC program into their general health and safety procedures, suggesting that the participatory approach may be sustained. Key program impact variables such as those we included should be explored further in larger, longer duration, controlled studies specifically designed to assess intervention effectiveness and sustainability.

6. Conclusion

Our results suggest that participatory approaches to reduce MSD hazards can be implemented in LTC environments. Furthermore, the results provide some evidence that the programs can have positive impacts on staff communication as well as identifying and reducing MSD hazards. Our results also suggest that a participatory approach such as EPIC can have a positive impact on managers' self-efficacy for MSD hazards. In addition, our results regarding program implementation show that key barriers include staff shortage/turnover, staff time to participate, and communication issues. Key implementation facilitators include frontline staff involvement/engagement, support from management, MSD training, and involving an external program facilitator. The results of this research regarding facilitators and barriers to implementation will allow future research to achieve better implementation and therefore provide better evidence of effectiveness.

7. Practical applications

Implementing a participatory organizational change program to reduce MSD hazards is feasible for LTC workplaces. A participatory program can improve communication and aid in the identification and control of a variety of MSD hazards. Importantly this type of program can be implemented in busy caregiving environments regardless of common challenges such as staff shortage/turnover.

Acknowledgements

This research was funded by a grant from the Province of Ontario [Grant # 15-R-031] and a seed grant from Centre of Research Expertise for the Prevention of Musculoskeletal Disorders (CRE-MSD) [#53254-10004].

Declaration of interest

The authors report no conflicts of interest.

References

Abildgaard, J. S., Nielsen, K., Wåhlin-Jacobsen, C. D., Maltesen, T., Christensen, K. B., & Holtermann, A. (2019). 'Same, but different': A mixed-methods realist evaluation of a cluster-randomized controlled participatory organizational intervention. *Human Relations*. 0018726719866896.

Andersen, L. L., & Zebis, M. K. (2014). Process evaluation of workplace interventions with physical exercise to reduce musculoskeletal disorders. *International Journal of Rheumatology, 2014*. https://doi.org/10.1155/2014/761363 761363.

Armstrong, P., & Daly, T. (2004). *There are not enough hands: Conditions in Ontario's long-term care facilities*. Toronto: Canadian Union of Public Employees.

Blair, C. E., & Glaister, J. A. (2005). Making the most of nursing homes' number one resource. *Health Care Management Review, 30*(2), 110–115. https://doi.org/10.1097/00004010-200504000-00005.

Boakye-Dankwa, E., Teeple, E., Gore, R., & Punnett, L. (2017). Associations among health care workplace safety, resident satisfaction, and quality of care in long-term care facilities. *Journal of Occupational and Environmental Medicine, 59*(11), 1127–1134. https://doi.org/10.1097/jom.0000000000001163.

Canadian Research Network for Care in the Community (CRNCC) (2010). *Ontario personal support workers in home and community care: CRNCC/PSNO survey results*. Toronto: Canadian Research Network for Care in the Community.

Carayon, P., Schoofs Hundt, A., Karsh, B. T., Gurses, A. P., Alvarado, C. J., Smith, M., & Flatley Brennan, P. (2006). Work system design for patient safety: The SEIPS model. *Quality & Safety in Health Care, 15*(Suppl 1), i50–i58. https://doi.org/10.1136/qshc.2005.015842.

Cole, D. C., Theberge, N., Dixon, S. M., Rivilis, I., Neumann, W. P., & Wells, R. (2009). Reflecting on a program of participatory ergonomics interventions: A multiple case study. *Work, 34*(2), 161–178. https://doi.org/10.3233/wor-2009-0914.

David, G., Woods, V., Li, G., & Buckle, P. (2008). The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal disorders. *Applied Ergonomics, 39*(1), 57–69. https://doi.org/10.1016/j.apergo.2007.03.002.

Driessen, M. T., Proper, K. I., Anema, J. R., Bongers, P. M., & van der Beek, A. J. (2010). Process evaluation of a participatory ergonomics programme to prevent low back pain and neck pain among workers. *Implementation Science, 5*, 65. https://doi.org/10.1186/1748-5908-5-65.

Framke, E., Sorensen, O. H., Pedersen, J., & Rugulies, R. (2018). Can illegitimate job tasks be reduced by a participatory organizational-level workplace intervention? Results of a cluster randomized controlled trial in Danish pre-schools. *Scandinavian Journal of Work, Environment and Health, 44*(2), 219–223. https://doi.org/10.5271/sjweh.3683.

Garg, A., & Owen, B. (1992). Reducing back stress to nursing personnel: An ergonomic intervention in a nursing home. *Ergonomics, 35*(11), 1353–1375. https://doi.org/10.1080/00140139208967398.

Greene, B. L., DeJoy, D. M., & Olejnik, S. (2005). Effects of an active ergonomics training program on risk exposure, worker beliefs, and symptoms in computer users. *Work, 24*(1), 41–52.

Gupta, N., Wahlin-Jacobsen, C. D., Abildgaard, J. S., Henriksen, L. N., Nielsen, K., & Holtermann, A. (2018). Effectiveness of a participatory physical and psychosocial intervention to balance the demands and resources of industrial workers: A cluster-randomized controlled trial. *Scandinavian Journal of Work, Environment and Health, 44*(1), 58–68. https://doi.org/10.5271/sjweh.3689.

Haines, H., & Wilson, J. (1998). *Development of a framework for participatory ergonomics, report 174/1998*. London: Health & Safety Executive.

Haines, H., Wilson, J. R., Vink, P., & Koningsveld, E. (2002). Validating a framework for participatory ergonomics (the PEF). *Ergonomics, 45*(4), 309–327. https://doi.org/10.1080/00140130210123516.

Haukka, E., Leino-Arjas, P., Viikari-Juntura, E., Takala, E. P., Malmivaara, A., Hopsu, L., ... Riihimaki, H. (2008). A randomised controlled trial on whether a participatory ergonomics intervention could prevent musculoskeletal disorders. *Occupational and Environmental Medicine, 65*(12), 849–856. https://doi.org/10.1136/oem.2007.034579.

Hignett, S., Wilson, J. R., & Morris, W. (2005). Finding ergonomic solutions: Participatory approaches. *Occupational Medicine (Oxford, England), 55*(3), 200–207. https://doi.org/10.1093/occmed/kqi084.

Holmberg, M. D., Flum, M., West, C., Zhang, Y., Qamili, S., & Punnett, L. (2013). Nursing assistants' dilemma: Caregiver versus caretaker. *Hospital Topics, 91*(1), 1–8. https://doi.org/10.1080/00185868.2013.757953.

Jakobsen, M. D., Aust, B., Kines, P., Madeleine, P., & Andersen, L. L. (2019). Participatory organizational intervention for improved use of assistive devices in patient transfer: A single-blinded cluster randomized controlled trial. *Scandinavian Journal of Work, Environment and Health, 45*(2), 146–157. https://doi.org/10.5271/sjweh.3769.

Kamioka, H., Okuizumi, H., Okada, S., Takahashi, R., Handa, S., Kitayuguchi, J., & Mutoh, Y. (2011). Effectiveness of intervention for low back pain in female caregivers in nursing homes: A pilot trial based on multicenter randomization. *Environmental Health and Preventive Medicine, 16*(2), 97–105. https://doi.org/10.1007/s12199-010-0170-1.

Kennedy, C. A., Amick, B. C., 3rd, Dennerlein, J. T., Brewer, S., Catli, S., Williams, R., ... Rempel, D. (2010). Systematic review of the role of occupational health and safety interventions in the prevention of upper extremity musculoskeletal symptoms, signs, disorders, injuries, claims and lost time. *Journal of Occupational Rehabilitation, 20*(2), 127–162. https://doi.org/10.1007/s10926-009-9211-2.

Kotejoshyer, R., Zhang, Y., Flum, M., Fleishman, J., & Punnett, L. (2019). Prospective evaluation of fidelity, impact and sustainability of participatory workplace health teams in skilled nursing facilities. *International Journal of Environmental Research and Public Health, 16*(9). https://doi.org/10.3390/ijerph16091494.

Kurowski, A., Gore, R., Buchholz, B., & Punnett, L. (2012). Differences among nursing homes in outcomes of a safe resident handling program. *Journal of Healthcare Risk Management, 32*(1), 35–51. https://doi.org/10.1002/jhrm.21083.

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*(3–4), 462–480. https://doi.org/10.1007/s10464-012-9522-x.

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: A sourcebook of new methods*. Thousand Oaks, CA: Sage publications.

Murta, S. G., Sanderson, K., & Oldenburg, B. (2007). Process evaluation in occupational stress management programs: A systematic review. *American Journal of Health Promotion, 21*(4), 248–254. https://doi.org/10.4278/0890-1171-21.4.248.

Owen, B. D., Keene, K., & Olson, S. (2002). An ergonomic approach to reducing back/shoulder stress in hospital nursing personnel: A five year follow up. *International Journal of Nursing Studies, 39*(3), 295–302. https://doi.org/10.1016/s0020-7489(01)00023-2.

Pehkonen, I., Takala, E. P., Ketola, R., Viikari-Juntura, E., Leino-Arjas, P., Hopsu, L., ... Riihimaki, H. (2009). Evaluation of a participatory ergonomic intervention process in kitchen work. *Applied Ergonomics, 40*(1), 115–123. https://doi.org/10.1016/j.apergo.2008.01.006.

Pinnock, H., Barwick, M., Carpenter, C. R., Eldridge, S., Grandes, G., Griffiths, C. J., ... Taylor, S. J. (2017). Standards for Reporting Implementation Studies (StaRI) statement. *BMJ, 356*. https://doi.org/10.1136/bmj.i6795 i6795.

Rasmussen, C. D., Holtermann, A., Bay, H., Sogaard, K., & Birk Jorgensen, M. (2015). A multifaceted workplace intervention for low back pain in nurses' aides: A pragmatic stepped wedge cluster randomised controlled trial. *Pain, 156*(9), 1786–1794. https://doi.org/10.1097/j.pain.0000000000000234.

Rasmussen, C. D., Holtermann, A., Mortensen, O. S., Sogaard, K., & Jorgensen, M. B. (2013). Prevention of low back pain and its consequences among nurses' aides in elderly care: A stepped-wedge multi-faceted cluster-randomized controlled trial. *BMC Public Health, 13*, 1088. https://doi.org/10.1186/1471-2458-13-1088.

Rasmussen, C. D., Larsen, A. K., Holtermann, A., Sogaard, K., & Jorgensen, M. B. (2014). Adoption of workplaces and reach of employees for a multi-faceted intervention targeting low back pain among nurses' aides. *BMC Medical Research Methodology, 14*, 60. https://doi.org/10.1186/1471-2288-14-60.

Rasmussen, C. D., Lindberg, N. K., Ravn, M. H., Jorgensen, M. B., Sogaard, K., & Holtermann, A. (2017). Processes, barriers and facilitators to implementation of a participatory ergonomics program among eldercare workers. *Applied Ergonomics, 58*, 491–499. https://doi.org/10.1016/j.apergo.2016.08.009.

Rivilis, I., Van Eerd, D., Cullen, K., Cole, D. C., Irvin, E., Tyson, J., & Mahood, Q. (2008). Effectiveness of participatory ergonomic interventions on health outcomes: A systematic review. *Applied Ergonomics, 39*(3), 342–358. https://doi.org/10.1016/j.apergo.2007.08.006.

Saldana, J. (2015). *The coding manual for qualitative researchers* (3rd ed.). London: Sage Publications.

Schelvis, R. M., Wiezer, N. M., van der Beek, A. J., Twisk, J. W., Bohlmeijer, E. T., & Oude Hengel, K. M. (2017). The effect of an organizational level participatory intervention in secondary vocational education on work-related health outcomes: Results of a controlled trial. *BMC Public Health, 17*(1), 141. https://doi.org/10.1186/s12889-017-4057-6.

van Beurden, K. M., Vermeulen, S. J., Anema, J. R., & van der Beek, A. J. (2012). A participatory return-to-work program for temporary agency workers and unemployed workers sick-listed due to musculoskeletal disorders: A process evaluation alongside a randomized controlled trial. *Journal of Occupational Rehabilitation, 22*(1), 127–140. https://doi.org/10.1007/s10926-011-9314-4.

van Eerd, D., Cole, D., Irvin, E., Mahood, Q., Keown, K., Theberge, N., ... Cullen, K. (2010). Process and implementation of participatory ergonomic interventions: A systematic review. *Ergonomics, 53*(10), 1153–1166. https://doi.org/10.1080/00140139.2010.513452.

Van Eerd, D., Ferron, E. M., D'Elia, T., Morgan, D., Ziesmann, F., & Amick, B. C. 3rd., (2018). Process evaluation of a participatory organizational change program to reduce musculoskeletal and slip, trip and fall injuries. *Applied Ergonomics, 68*, 42–53. https://doi.org/10.1016/j.apergo.2017.10.015.

Wahlstrom, J., Ostman, C., & Leijon, O. (2012). The effect of flooring on musculoskeletal symptoms in the lower extremities and low back among female nursing assistants. *Ergonomics, 55*(2), 248–255. https://doi.org/10.1080/00140139.2011.583360.

Wells, R., Laing, A., & Cole, D. (2009). Characterizing the intensity of changes made to reduce mechanical exposure. *Work, 34*(2), 179–193. https://doi.org/10.3233/wor-2009-0915.

Wilson, J. R., & Haines, H. M. (1997). Participatory ergonomics. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 490–513). New York: John Wiley & Sons.

Workplace Safety & Insurance Board (WSIB) (2019). *By the numbers: 2019 WSIB statistical report*. Toronto, Ontario: Workplace Safety & Insurance Board.

Zhang, Y., Flum, M., Nobrega, S., Blais, L., Qamili, S., & Punnett, L. (2011). Work organization and health issues in long-term care centers. *Journal of Gerontological Nursing, 37*(5), 32–40. https://doi.org/10.3928/00989134-20110106-01.

Dr. Dwayne Van Eerd is a scientist at the Institute for Work & Health. He has a PhD in work and health from the Faculty of Applied Health Sciences at the University of Waterloo. His research focuses on MSD prevention, as well as the implementation of programs and practices to prevent work injury and disability related to physical and mental health. His research projects include evaluations of participatory ergonomics interventions, evaluations of workplace MSD and mental health practices/programs and their implementation, and systematic reviews of the prevention literature. Van Eerd also practices, and researches integrated knowledge transfer and exchange.

Teresa D'Elia has an MA in Gerontology from Simon Fraser University. Teresa is a research coordinator with an extensive background in mixed method program evaluation. Her work focuses on health promotion programs and preventative health practices in the long-term care sector. Her recent work explores occupational health and safety in long term care homes.

Dr. Era Mae Ferron graduated from Queen's University with a degree in Nursing Science and earned her PhD from the University of Toronto's Lawrence S. Bloomberg Faculty in Nursing. Her thesis focused on the impact of work environment factors on nurse faculty's intent to remain employed in academia. Dr. Ferron has since collaborated on several projects focused on workplace violence prevention in healthcare settings. She has a keen interest in putting knowledge into action and is the Manager of Knowledge Translation on the project, Violence, Aggression & Responsive Behaviour (VARB) project at the Public Services Health & Safety Association.

Dr. Lynda Robson is a scientist at the Institute for Work & Health and an adjunct professor in the School of Occupational and Public Health at Ryerson University. She obtained her PhD in biochemistry from the University of Toronto, but later changed fields to public health sciences. Robson's research publications focus on occupational health and safety (OHS) management, organizational change in OHS, and the evaluation of prevention programs.

Dr. Benjamin C. Amick III is Associate Dean for Research and Professor of Epidemiology at the Jay W Boozman College of Public Health in the University of Arkansas Medical Sciences. He conducts research on organizational change and implementation of OSH programs in workplaces. He also continues to work on life course issues in work and health.

Journal of Safety Research 78 (2021) 276–291

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Injury severity analysis of rollover crashes for passenger cars and light trucks considering temporal stability: A random parameters logit approach with heterogeneity in mean and variance

Shaojie Liu [a], Wei David Fan [b,*], Yang Li [a]

[a] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3366, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States

[b] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3261, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States

ARTICLE INFO

Article history:
Received 16 January 2021
Received in revised form 10 March 2021
Accepted 22 June 2021
Available online 6 July 2021

Keywords:
Transferability
Injury severity
Safety
Contributing factors
Vehicle type

ABSTRACT

Problem: The rollover crash is a serious crash type that often causes higher injury severities. Moreover, factors that contribute to the injury severities of rollover crashes may show instabilities in different vehicle types and time periods, which requires further investigations. This study utilizes the rollover crash data in North Carolina from Highway Safety Information System (HSIS) to study the effect instabilities of factors in vehicle type and time periods in rollover crashes. *Methods:* The injury severities of drivers are estimated using the random parameters logit (RPL) model with heterogeneity in means and variances. Available factors in HSIS have been categorized into three groups, which are drivers, road, and environment, respectively. This study also justifies the segmentations through transferability tests. The effects of identified significant factors are evaluated using marginal effects. *Results:* Factors such as FWP (farm, wood, and pasture areas), unhealthy physical condition, impaired physical condition, road adverse, and so forth have shown instabilities in marginal effects among vehicle types and time periods. *Practical Applications:* The finding of this research could provide important references for policy makers and automobile manufactures to help mitigate the injury severity of rollover crashes.

1. Introduction

Rollover crashes are dangerous incidents that often result in higher fatalities. Based on 2009–2018 public crash records (Fatality Analysis Report System and Crash Report Sampling System) from the National Highway Traffic Safety Administration (NHTSA), motor-vehicle rollover crashes are more likely to cause injuries and even fatal injuries of drivers compared to non-rollover crashes, as shown in Fig. 1. On average, the fatality rate for rollover crashes and non-rollover crashes are 0.036 and 0.004, respectively, indicating that rollover crashes are more likely to incur severe injuries. In rollover crashes, the injury severities may also vary among different vehicle types and time periods. In the traffic safety facts report (NHTSA, 2016), it was stated that, from 1982 to 2016, 22–24% passenger car rollover crashes involve fatalities, while the percentages for SUVs and pickups are 41–68%. As for the different time periods, it was found that the fatal passenger car rollover crashes had been decreasing, while the fatal light truck (SUV, vans and pickup trucks) rollover crashes were increasing from 1996 to 2000 (Deutermann, 2002).

The identification of significant factors influencing injury severities and their associated impact magnitudes can help policymakers develop effective guidelines to reduce the severity and frequency of accidents. Although much effort has been made in investigating rollover crashes in the areas of single vehicle (Anarkooli et al., 2017), gender (Wu et al., 2016), and large truck (Khattak et al., 2003; Azimi et al., 2020), to the best knowledge of the authors, research on the injury severity analysis of light truck rollover crashes are limited. Further, the temporal stability/instability in the effects of contributing factors has not been taken into account in the rollover crashes by considering passenger cars and light trucks separately. As larger vehicles become more popular in the Unites States, the safety concerns for SUVs, pickups, and minivans deserve more attention. An exploration of impacts of the environment, vehicle, and drivers' characteristics on the rollover

* Corresponding author.
E-mail addresses: sliu29@uncc.edu (S. Liu), wfan7@uncc.edu (W.D. Fan), yli107@uncc.edu (Y. Li).

https://doi.org/10.1016/j.jsr.2021.06.013
0022-4375/

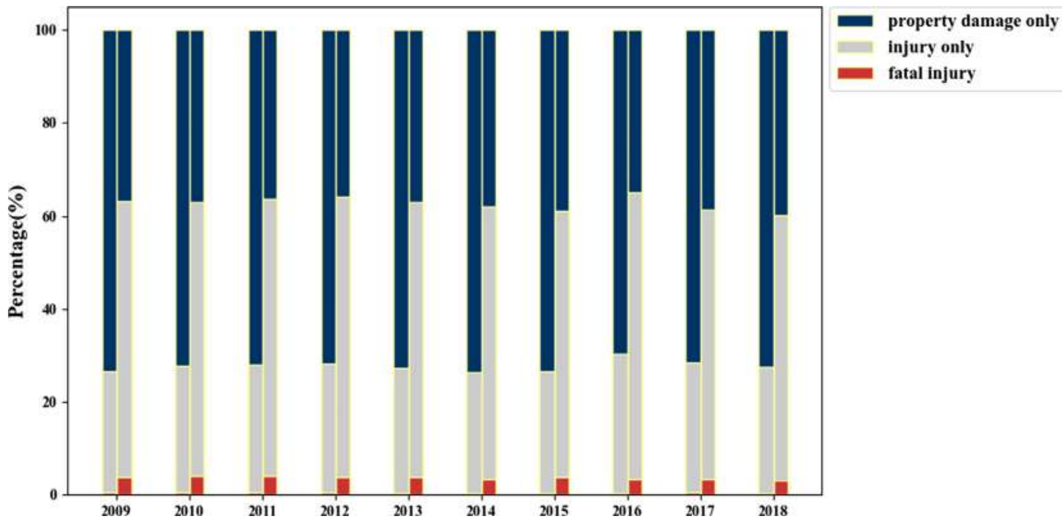

Fig. 1. Injury severity proportions for non-rollover crashes (left) and rollover crashes (right) from 2009 to 2018 based on crash records provided by NHTSA.

crashes will deliver safety insights for all stakeholders, such as governmental entities, vehicle manufacturers, and drivers.

Hence, this study aims to explore the potential influencing factors that could exert a significant impact on the drivers' injury severity in the rollover crashes considering two time periods of crash records. Also, to further explore effects of factors of different vehicle types, two types of vehicles are defined for comparison in this research, that is, passenger car and light truck (minivans, SUV and pickup).

2. Literature review

The following literature review section consists of two parts that can help illustrate the significance of this research: (a) existing studies on rollover crashes; and (b) studies on random parameter logit model with heterogeneity in means and variances.

Much effort has been put into investigating rollover crashes, with topics ranging from SUV, large trucks, severe rollover crashes, passenger vehicle, and single vehicles. This indicates the necessity of considering impacts of certain vehicle types during the investigation of rollover crashes. They are summarized and ordered by time in Table 1. Important information including authors, years of publication, models, topics, and their associated findings are presented.

Based on the literature review, although there are some rollover crash studies that considered vehicle types (such as large trucks, passenger vehicles, pickups, and SUVs), a systematic comparison has not been fully explored for different vehicle types. Recently, numerous crash studies have demonstrated temporal instabilities of contributing factors in vehicle crashes (Mannering, 2018; Behnood & Mannering, 2019; Al-Bdairi et al., 2020; Yu et al., 2020; Li & Fan, 2020). Additionally, existing studies on rollover crashes have also proved the necessity of investigating rollover crashes with the segmentation of vehicle types (Islam et al., 2016). Nevertheless, previous rollover crash studies mainly utilized the RPL model, which has proven to be less effective compared to RPL with heterogeneity in means and variances. Overestimation or underestimation for effects of certain significant parameters may occur. Hence, by utilizing the RPL model with heterogeneity in means and variances, this research attempts to study the rollover crashes with a combined consideration of temporal instabilities and different vehicle types to provide more insight on rollover crash analysis.

The RPL model is a major advancement of the discrete choice model because of its ability to account for unobserved heterogeneity by allowing parameters to vary across observations. The RPL model with heterogeneity with means and variances differs from the RPL model in that it allows for different means and variances in the parameter distribution. Such extensions for the RPL model are statistically superior to the RPL-only model, as proven in other studies (Mannering, 2018; Al-Bdairi et al., 2020; Li & Fan, 2020; Islam & Mannering, 2020). By applying the RPL model with heterogeneity in means and variances, the heterogeneity in factor effects can be captured in greater detail compared to the RPL-only model. In the case of rollover crashes, previous work with the RPL-only model may have overestimated or underestimated the effects of certain factors toward injury severities without accounting for heterogeneity in means and variances. This research aims to employ the RPL model with heterogeneity in means and variances to explore injury severity of rollover crashes by considering potential instabilities and stabilities of contributing factors in different vehicle types and time periods.

3. Data descriptions

This study uses the crash records of North Carolina from 2008 to 2017, with 10 years in total. The crash records are retrieved from the Highway Safety Information System (HSIS), which is a database that contains the crash data as well as the relevant environment characteristics across several states, including California, Illinois, Maine, Michigan, Minnesota, North Carolina, Ohio, Utah, and Washington. To guarantee the data sufficiency for measuring the model performance, two periods are defined, one being 2008–2012 and the other being 2013–2017. To explore possible heterogeneity effects between vehicle types, segmentation on light trucks and passenger vehicles are also implemented. Therefore, both segmentations result in four subgroups of data in total (passenger car in 2008–2012, light trucks in 2008–2012, passenger car in 2013–2017, light trucks in 2013–2017). The available crash attributes in the dataset are classified into three categories (driver, road, and environmental characteristics). After removing records without necessary information listed in Table 2, a total number of 18,476 observations are extracted. The property damage only (PDO) is selected as the base reference. Due to the relatively small sample size of fatal injury and incapacitating injury, they have

Table 1
Existing studies in Rollover Crashes.

Authors	Year	Model	Topics	Findings
Khattak and Rocha	2003	negative binomial model	SUV rollover crashes	curvatures and dangerous driving behavior can directly increase injury severity
Khattak et al	2003	binary logit model	Large truck rollover crashes	dangerous driving behavior, truck exposure, hazardous materials and post-crash fires can increase the risk of higher injuries
Conroy et al.	2006	binary logit model	Serious rollover crashes	intrusion at the occupant's position, the vehicle interior side and roof are identified as sources of injury
Keall and Newstead	2009	logit regression analysis	Passenger vehicle rollover crashes	teenager drivers, older vehicles are highly risky factor for rollover crashes
Hu and Donnell	2011	multinomial logit model	Cross-median and rollover crashes on rural divided highways	median cross-slopes and narrower medians were more likely to cause severer injury
Islam et al.	2016	RPL model	SUV and pickup rollover crashes	horizontal curve and intersection are random parameter in the SUV rollover crashes while horizontal curve and dry roadway surfaces are random parameter in pickup rollover crashes
Wu et al.	2016	RPL model	Single vehicle rollover crashes	female drivers are more likely to suffer severe or fatal injuries in rollover crashes than male drivers.
Anarkooli et al	2017	random effect generalized ordered probit model	Single vehicle rollover crashes	REGOP model is found to outperform the RPL model, significant factor toward injury severity were identified
Azimi et al.	2020	random parameter ordered logit model	Large truck rollover crashes	Impacts of lighting conditions and driving speed had significant variation across observations

combind as severe injury (SI). Evident injury and possible injury have also been combined as minor injury (MI).

4. Methodology

This study uses the RPL model with heterogeneity in means and variances to model the injury severities of drivers of rollover crashes for different time periods and vehicle types (i.e., passenger car in 2008–2012, light truck in 2008–2012, passenger car in 2013–2017, and light truck in 2013–2017). The injury severity function of the model is a linear function, which has two components that capture the observed variable and the unobserved factor, respectively, as shown in Equation (1).

$$U = \beta \mathbf{X} + \varepsilon \tag{1}$$

where U represents the utility, β is the vector of parameters to be estimated for the observed variable vector $\mathbf{X}$, ε collectively stands for the unobserved factors, which is represented by a certain type of distribution. For a closed-form of choice probability, the distribution of ε is assumed to follow a Gumbel distribution. The resulting closed-form expression for the choice probability for individual n given alternative j is shown in Eq. (2).

$$P_{n,j} = \frac{e^{U_{nj}}}{\sum_j e^{U_{nj}}} \tag{2}$$

where U_{nj} represents the utility for individual n given an alternative choice (or injury severity level) j. In the multinomial logit model, β is assumed to be a fixed constant for all observations. However, this assumption might not be valid since individuals have different sensitivity towards the variables. For example, in the injury severity analysis, the impact of the dark roadway with light is different for individuals with various visionary capabilities. This assumption is released in the RPL model by allowing the coefficients to follow a random distribution described by a set of parameters $\boldsymbol{\theta}$ (usually means and variances). Eq. (3) provides a form of the RPL model.

$$P_{n,j} = \int \frac{e^{U_{nj}}}{\sum_j e^{U_{nj}}} f(\boldsymbol{\beta}|\boldsymbol{\theta}) d\boldsymbol{\beta} \tag{3}$$

The distribution of $\boldsymbol{\beta}$ can be normal, lognormal, or triangular distribution. Many studies have demonstrated that valid results can be obtained if $\boldsymbol{\beta}$ is assumed to follow a normal distribution (Li & Fan, 2020; Liu & Fan, 2020). Instead of fixing the means and variances in the RPL model, the RPL model accounting for heterogeneity in means and variances allows different distribution for means and variances. The resulting form of such model is shown as below according to Greene et al. (2006) and Seraneeprakarn et al. (2017):

$$\beta_{n,j} = \beta + \delta_{\boldsymbol{n,j}} \mathbf{z}_{\boldsymbol{n,j}} + \sigma_{n,j} \exp(\omega_{\boldsymbol{n,j}} \mathbf{w}_{\boldsymbol{n,j}}) v_{n,j} \tag{4}$$

where β represents the mean parameter estimated across all observations, $\delta_{\boldsymbol{n,j}}$ denotes the vector of estimated parameters for $\mathbf{z}_{\boldsymbol{n,j}}$, which is a vector of attributes describing the heterogeneity in the mean for injury severity level j.$\sigma_{n,j}$ is standard deviation estimated across all observations, $\omega_{\boldsymbol{n,j}}$ means the vector of the estimated parameters for $\mathbf{w}_{\boldsymbol{n,j}}$, which is a vector of attributes capturing heterogeneity in standard deviation. If no heterogeneity has been found in the standard deviation, then the model would fall into an RPL model with heterogeneity in means. Further, if no heterogeneity in means is found, then the model will collapse into an RPL model. Finally, the RPL model can collapse into a multinomial logit model if no random parameter is found among identified significant parameters. The models for each segmentation have been estimated by using the maximum likelihood method. The selection of the number of Halton draws is not a definite decision and a higher number of Halton draws does not necessarily produce better fitting performances. The number of Halton draws can range from 100 to 1,000 (Train, 2009; Bhat, 2003; Gong & Fan, 2017). According to the literature, 500 Halton draws are popularly applied for their efficiencies (Moore et al., 2011; Liu & Fan, 2020). Therefore, in this research, 500 Halton draws are used to estimate the models.

5. Transferability test

In this research, separated models are developed for different vehicle types and time range. However, it needs to be answered whether these separated models are significantly different from each other or a unified model is sufficient for identified segmentations. A transferability test can be used to test the significance of separated models. First, to test whether the developed models regarding time periods and vehicle types are significantly different, Eq. (5) is applied. The test statistic $\boldsymbol{X}^2$ is assumed to follow a $\boldsymbol{\chi}^2$ distribution.

$$\boldsymbol{X}^2 = -2\left[LL(\beta_{s2,s1}) - LL(\beta_{s1})\right] \tag{5}$$

Table 2
Descriptive Statistics for the Rollover Crashes.

		Severe Injury (Incapacitating/Fatal Injury)	Minor Injury (Evident/Possible Injury)	Property Damage Only	Total
Overall Data Segmentation					
Passenger Vehicles	2008–2012	96	1913	1645	3654
	2012–2017	102	1718	1606	3426
Light Trucks	2008–2012	239	2884	2642	5765
	2012–2017	195	2645	2791	5631
Driver Attributes					
Age	young (if the age of the driver is less than 25)	209	3877	3764	7850
	middle age (if the age of driver is between 25–50) *	303	3937	3688	7928
	old (if the age of the driver is greater than 50 years)	120	1346	1232	2698
Gender	Male	439	5318	5974	11,731
	female*	193	3842	2710	6745
Physical Condition	normal*	349	6947	7508	14,804
	unhealthy physical condition (illness, fatigue, fell asleep, loss of consciousness)	24	492	272	788
	impaired physical condition (impairment due to medications, drugs or alcohol)	259	1721	904	2884
Road Attributes					
Traffic control	no control*	208	3198	2850	6256
	sign control (stop, yield, warning)	12	167	215	394
	signal control (stop and go signal, flashing signal)	1	55	77	133
	double yellow line	411	5740	5542	11,693
Speed limits	30 mph (speed limit less than or equal to 30 mph) *	2	64	88	154
	30–50 mph (speed limit between 30 mph to 50 mph)	105	1670	1809	3584
	≥50 mph (speed limit great than 50 mph)				
Road configuration	straight level*	208	3471	3290	6969
	adverse straight road (bottom, grade, crest)	61	1024	1161	2246
	curve level	237	2649	2324	5210
	curve-adverse (hillcrest, grade or bottom)	126	2016	1909	4051
Road condition	dry*	568	7116	5781	13,465
	road adverse (wet, watery, icy, snowy, sandy, mudy, dirty, or graveled)	64	2044	2903	5011
Road type	Undivided	528	7630	7421	15,579
	two-way divided*	104	1530	1263	2897
The number of lanes	1–2 lanes*	526	7556	7309	15,391
	3–4 lanes	94	1312	1066	2472
	More than 4 lanes	12	292	309	613
Road class	rural arterial	113	1426	1363	2902
	rural collector	184	2807	2555	5546
	rural local	212	2574	2440	5226
	urban arterial	68	1300	1251	2619
	urban collector	26	453	437	916
	urban local*	29	600	638	1267
Environment Attributes					
Region	Rural	615	8778	8300	17,693
	urban*	17	382	384	783
Weather	clear*	507	6464	5526	12,497
	Snow	6	256	554	816
	Cloudy	87	1549	1558	3194
	Rain	29	812	972	1813
	fog, smog, smoke	3	79	74	156
Light	daylight*	309	5224	5029	10,562
	dusk and dawn	23	350	383	756
	dark lighted road	9	126	117	252
	dark no lighted road	291	3460	3155	6906
Workzone	Workzone	5	78	59	142
	no workzone*	627	9082	8625	18,334
Terrain	flat*	184	2210	1964	4358
	Rolling	375	5692	5015	11,082
	Mountain	73	1258	1705	3036
Development	FWP (farm, woods, pasture)	508	7280	6741	14,529
	residential*	99	1477	1465	3041
	Commercial	22	363	445	830
	Institutional	3	40	33	76
Intersection	Intersection	22	332	452	806
	Non intersection*	610	8828	8232	17,670

*Indicate the base reference.

where $LL(\beta_{s2,s1})$ denotes the log-likelihood of the converged model for the data of segmentation $s1$, using the estimated parameters from data of segmentation $s2$, $LL(\beta_{s1})$ represents the log-likelihood of the converged model for the data of segmentation $s1$ using the estimated parameter from the data segmentation $s1$. The degree of freedom is the number of parameters. The reverse procedure is also conducted, in which the segmentation $s2$ and $s1$ are replaced with each other. The test statistic results are summarized in Table 3 (numbers in brackets show the confidence level to reject the null hypothesis, while numbers in parenthesis indicate the degree of freedom). The overall results demonstrate the necessity of segmentation based on time periods and vehicle types since only 1 out of 12 tests cannot be rejected at a 95% confidence level.

Additionally, the sum of the log-likelihood values from all sub-segmentation is less than the log-likelihood values of the model for the overall dataset (−14,386 vs. −14,414), indicating better fitting performance with segmentation. A further test is conducted to examine the market segmentation efficiency based on vehicle types and time periods. The test statistics also follow the $\boldsymbol{X}^2$ distribution. The degrees of freedom can be obtained as the total number of parameters from all sub segmentations, minus the number of parameters from the model for the overall dataset. The equation obtaining $\boldsymbol{X}^2$ is shown below:

$$\boldsymbol{X}^2 = -2[LL(\beta_{overall}) - LL(\beta_{s1}) - LL(\beta_{s2}) - LL(\beta_{s3}) - LL(\beta_{s4})] \qquad (6)$$

where $LL(\beta_{overall})$ denotes the log-likelihood of the best model estimated for the overall dataset, $LL(\beta_{s1})$-$LL(\beta_{s4})$ represents the four identified segmentations in this research (i.e., passenger vehicles in 2008–2012 and 2013–2017, light trucks in 2008–2012 and 2013–2017). The calculated $\boldsymbol{X}^2$ equals 56.78 with a degree of freedom being 45. This indicates an acceptable confidence level (89%) to reject the null hypotheses that the model results are transferable across all segmentations.

6. Results and discussions

In this section, the model results are discussed. The factors that are found to be significant at the 95% level in the MNL model are kept for the analysis of random parameters accounting for potential heterogeneity in means and variances. To fully capture the heterogeneity effects, this study retains the factors that are significant at the 90% level, while exploring potential heterogeneity in means and variances. Despite models for each sub dataset, an RPL model accounting for potential heterogeneity in the means has also been estimated for the overall dataset. The model results are presented in Appendix Tables A1–A5 and marginal effects are summarized in Tables 4–8. Heterogeneity in means and variances is found for passenger vehicle rollover crashes in 2012–2017; while heterogeneity in the means is found in light truck rollover crashes in 2012–2017, as well as for the whole dataset. However, the heterogeneity effect in means is not found in the other two segmentation. The following section analyzes the impacts of each significant factor using marginal effects in terms of driver, road, and environmental attributes. To provide an intuitive comparison between the RPL model and RPL model with heterogeneity in means and variances, the RPL model results are also presented, which can provide detailed information on how the effects of significant parameters vary from RPL model to RPL with heterogeneity in means and variances.

6.1. Driver attributes

6.1.1. Effect of impaired physical condition

Impaired physical condition has been identified as a significant contributing factor toward severe injury in all segmentations. For vehicle types, it can be observed that the magnitudes of the marginal effects are larger for passenger vehicles compared to light trucks in 2008–2012. Nevertheless, in 2013–2017, the effects of this factor on light trucks are greater, indicating instability in the effect of the factor between different vehicle types and over time. When examining the temporal instability, impaired physical condition is exerting more effect for light truck rollover crashes in 2013–2017 than it did in 2008–2012. In contrast, this effect is lessened for passenger car rollover crashes in 2013–2017. Light truck drivers may need immediate attention from the policy and decision-makers in order to mitigate such malignant effects.

In terms of model structure, RPL model with heterogeneity in means shows that such impact may be exaggerated by the RPL only model for the passenger car in 2013–2017 (0.0135 in RPL vs. 0.0125 in RPL with heterogeneity in means). In the RPL model with heterogeneity in means and variances for passenger car rollover crashes in 2013–2017, the mitigatory effect of this factor on the minor injury is reduced slightly from 0.0023 to 0.0018 (Fig. 2).

6.1.2. Effect of unhealthy physical condition

Overall, the unhealthy physical condition has a more significant impact on the minor injury in all sub segmentations except for the passenger car in 2013–2017, in which unhealthy physical condition contributes to severe injury instead of minor injury with less extent. To more extent, the unhealthy physical condition has a greater impact on light truck rollover crashes compared to passenger car rollover crashes in terms of minor injury. This is reasonable as light truck drivers practice more heavy driving tasks during longer driving time, and hence unhealthy physical conditions of light truck drivers might lead to minor injury with higher chances.

As for the model structure, by accounting for potential heterogeneity in means, the magnitude of impact on the minor injury is significantly higher in the model of light truck rollover crashes in 2013–2017. This implies that one might underestimate the impact of unhealthy physical conditions by referring to the RPL only model. Other than that, no other notable difference is identified between the RPL model with heterogeneity in means and RPL model with heterogeneity in means and variances (Fig. 3).

6.1.3. Effect of male gender

The factor of the male gender has a negative impact on minor injury and a positive impact on PDO. Such impacts are most significant in 2013–2017 light truck rollover crashes. Also, it can be observed that the effect of the male gender is more obvious in the light truck rollover crashes compared to passenger car rollover crashes. Understandably, light trucks attract more male drivers and therefore this impact is amplified in the light rollover crashes. For

Table 3
χ^2 value for transferability test.

	Parameter			
Data	PC 2008–2012	PC 2012–2017	LT 2008–2012	LT 2012–2017
PC 2008–2012 (18)	–	131.68 [99.99%]	79.61 [99.99%]	191.02 [99.99%]
PC 2012–2017 (17)	102.77 [99.99%]	–	180.02 [99.99%]	194.32 [99.99%]
LT 2008–2012 (18)	153.74 [99.99%]	138.42 [99.99%]	–	233.05 [99.99%]
LT 2012–2017 (20)	84.25 [99.99%]	103.29 [99.99%]	23.16 [64.2%]	–

Table 4
Marginal effects of significant factors in the model of Passenger Car Rollover Crashes in 2008–2012.

Variables	RPL Model		
	SI	MI	PDO
Driver Attributes			
young (1 if the driver is younger than 25; 0 otherwise)	−0.0057	0.0033	0.0024
impaired physical condition (1 if the driver is under the influence of medication, drugs or alcohol; 0 otherwise)	0.0198	0.0129	−0.0327
male (1 if the gender of the driver is male; 0 otherwise	0.0034	−0.0658	0.0624
unhealthy physical condition (1 if the driver is under the condition of illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	−0.0001	0.005	−0.0049
Road Attributes			
curve level road (1 if the road is curve and level; 0 otherwise)	0.0155	0.0046	−0.0201
road-adverse (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	0.0035	−0.0136	0.0102
3–4 lanes (1 if the road has 3–4 lanes; 0 otherwise)	−0.0003	0.0076	−0.0074
rural collector road (1 if the road is rural collector; 0 otherwise)	−0.0008	0.0164	−0.0156
rural local road (1 if the road is rural local; 0 otherwise)	−0.0007	0.0142	−0.0135
curve-adverse (1 if road is curve and also hillcrest, grade or bottom)	−0.0005	0.0118	−0.0113
Environment Attributes			
mountain (1 if the terrain is mountain; 0 otherwise)	0.0005	−0.0124	0.0119

Table 5
Marginal effects of significant factors in the model of Light Truck Rollover Crashes in 2008–2012.

Variables	RPL Model		
	SI	MI	PDO
Driver Attributes			
male (1 if the gender of the driver is male; 0 otherwise	−0.0043	−0.0788	0.0831
impaired physical condition (1 if the driver is under the influence of medication, drugs, or alcohol; 0 otherwise)	0.0183	−0.009	−0.0093
unhealthy physical condition (1 if the driver is under the condition of illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	−0.0005	0.0063	−0.0058
Road Attributes			
3–4 lanes (1 if the road has 3–4 lanes; 0 otherwise)	0.0034	0.0055	−0.0088
road-adverse (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−0.0044	−0.0432	0.0477
curve-level (1 if the road is curve and level; 0 otherwise)	0.0088	0.005	−0.0137
curve-adverse (1 if the road is curve with hillcrest, grade, or bottom; 0 otherwise)	0.0025	0.0081	−0.0104
30–50 mph (1 if the speed limit is between 30 mph to 50 mph; 0 otherwise)	0.0006	−0.007	0.0064
Environment Attributes			
rain (1 if the weather is raining; 0 otherwise)	−0.0001	0.0057	−0.0056
rolling (1 if the terrain is rolling; 0 otherwise)	−0.006	0.003	0.003

Table 6
Marginal effects of significant factors in the model of Passenger Car Rollover Crashes in 2013–2017.

Variable	RPL			RPL with heterogeneity in means			RPL with heterogeneity in means and variances		
	SI	MI	PDO	SI	MI	PDO	SI	MI	PDO
Driver Attributes									
impaired physical condition (1 if the driver is under the influence of medication, drugs, or alcohol; 0 otherwise)	0.0135	−0.0022	−0.0113	0.0125	−0.0023	−0.0102	0.0125	−0.0018	−0.0108
male (1 if the gender of the driver is male; 0 otherwise	0.003	−0.0457	0.0427	0.0018	−0.0264	0.0246	0.0014	−0.021	0.0196
unhealthy physical condition (1 illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	0.0018	−0.0004	−0.0015	0.0018	−0.0004	−0.0015	0.0019	−0.0003	−0.0016
Road Attributes									
undivided Road (1 if the road is undivided; 0 otherwise	0.0017	−0.0269	0.0251	0.0013	−0.0238	0.0225	0.0012	−0.0214	0.0202
road-adverse (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−0.0022	−0.0197	0.0219	−0.0021	−0.0184	0.0206	−0.0021	−0.0167	0.0188
Environment Attributes									
cloudy (1 if the weather is cloudy; 0 otherwise)	−0.0005	0.0102	−0.0097	−0.0005	0.0102	−0.0097	−0.0005	0.0094	−0.0089
rolling (1 if the terrain is rolling; 0 otherwise)	−0.0011	0.0169	−0.0158	−0.0014	0.0247	−0.0233	−0.0012	0.0219	−0.0207
FWP (1 if the development type if farm, woods, or pastures; 0 otherwise)	−0.0052	0.0876	−0.0824	−0.007	0.1133	−0.1063	−0.0038	0.0619	−0.0581
mountain (1 if the terrain is mountain; 0 otherwise)	−0.0021	0.0004	0.0016	−0.002	0.0003	0.0017	−0.002	0.0003	0.0017

Table 7
Marginal effects of significant factors in the Model of Light Truck Rollover Crashes in 2013–2017.

Variables	RPL			RPL in means		
	SI	MI	PDO	SI	MI	PDO
Driver Attributes						
old (1 if the driver is older than 50; 0 otherwise)	0.0032	−0.0017	−0.0016	0.0032	−0.0017	−0.0015
impaired physical condition (1 if the driver is under the influence of medication, drugs and alcohol; 0 otherwise)	0.0124	0.0189	−0.0314	0.0124	0.0188	−0.0311
male (1 if the gender of the driver is male; 0 otherwise)	0.0085	−0.1007	0.0922	0.0085	−0.101	0.0924
unhealthy physical condition (1 if the driver is under the condition of illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	−0.0004	0.0058	−0.0054	−0.0006	0.0085	−0.008
Road Attributes						
3–4 lanes (1 if the road has 3–4 lanes; 0 otherwise)	0.0024	−0.0013	−0.0011	0.0024	−0.0013	−0.0011
rural collector road (1 if the road is rural collector; 0 otherwise)	0.0028	0.0087	−0.0115	0.0028	0.0086	−0.0114
road-adverse (1 if the road is wet, watery, icy, snowy, sandy, mudy, dirty, or graveled; 0 otherwise)	−0.0042	−0.0385	0.0425	−0.0042	−0.0387	0.0429
undivided Road (1 if the road is undivided; 0 otherwise	0.003	−0.0398	0.0368	0.003	−0.0397	0.0366
Environment Attributes						
rolling (1 if the terrain is rolling; 0 otherwise)	−0.0003	0.0164	−0.0161	−0.0007	0.0179	−0.0172
mountain (1 if the terrain is mountain; 0 otherwise)	−0.0019	−0.0124	0.0142	−0.0019	−0.0122	0.0141
intersection (1 if the crash location is within footprint of intersection; 0 otherwise)	−0.0007	−0.0045	0.0052	−0.0007	−0.0046	0.0053

Table 8
Marginal effect of the Estimated Model for the Overall Dataset.

Variables	RPL			RPL with heterogeneity in means		
	SI	MI	PDO	SI	MI	PDO
Driver Attributes						
young (1 if the driver is younger than 25; 0 otherwise)	−0.0025	0.0014	0.0011	−0.0026	0.0014	0.0011
old (1 if the driver is older than 50; 0 otherwise)	0.0062	−0.0032	−0.003	0.0055	−0.0029	−0.0025
male (1 if the gender of the driver is male; 0 otherwise	0.0064	−0.0818	0.0754	0.0065	−0.0824	0.0759
impaired physical condition (1 if the driver is under the influence of medication, drugs, or alcohol; 0 otherwise	0.014	0.0139	−0.0278	0.014	0.0138	−0.0278
unhealthy physical condition (1 illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	−0.0003	0.0056	−0.0053	−0.0003	0.0056	−0.0053
Road Attributes						
curve level (1 if the road is curve and level; 0 otherwise)	0.0053	0.0051	−0.0104	0.0053	0.0051	−0.0105
rural arterial road (1 if the road is rural arterial; 0 otherwise)	0.0028	−0.0016	−0.0012	0.0028	−0.0016	−0.0012
rural local road (1 if the road is rural local; 0 otherwise)	0.0046	−0.0026	−0.002	−0.0044	−0.0058	0.0049
rural collector road (1 if the road is rural collector; 0 otherwise)	0.003	−0.0017	−0.0013	0.0029	−0.0016	−0.0013
road-adverse (1 if the road is wet, watery, icy, snowy, sandy, mudy, dirty, or graveled; 0 otherwise)	−0.0034	−0.0301	0.0335	−0.0036	−0.0301	0.0337
curve-adverse (1 if road is curve with hillcrest, grade or bottom)	0.0014	0.0072	−0.0086	0.0014	0.0072	−0.0086
undivided Road (1 if the road is undivided; 0 otherwise)	−0.0084	−0.035	0.0435	−0.0085	−0.035	0.0435
>=50 mph (1 if the speed limit is greater than 50 mph; 0 otherwise)	−0.0013	0.0176	−0.0163	−0.0013	0.0175	−0.0162
Environment Attributes						
cloudy (1 if the weather is cloudy; 0 otherwise)	−0.0003	0.0053	−0.005	−0.0003	0.0053	−0.005
rain (1 if the weather is rainy; 0 otherwise)	−0.0001	0.0048	−0.0047	−0.0001	0.0049	−0.0048
other adverse weather (1 if the weather is snow, sleet, hall, freezing rain/drizzle)	0.0001	−0.0027	0.0027	0	−0.0027	0.0027
commercial (1 if the development is commercial; 0 otherwise)	0.0002	−0.0027	0.0025	0.0002	−0.0027	0.0025
mountain (1 if the terrain is mountain; 0 otherwise)	−0.0014	−0.0132	0.0146	−0.0013	−0.0133	0.0146
intersection (1 if the road is within the footprint of intersection; 0 otherwise)	0.0002	−0.0032	0.003	0.0002	−0.0032	0.003

temporal instability, in more recent years (2013–2017), such impact is reduced for the passenger car rollover crashes.

In the RPL model accounting for potential heterogeneity in means for passenger rollover crashes in 2013–2017, the benign influence of male gender is decreased by 42% (from −0.0457 to −0.0264), indicating an overestimate for the influence of the male gender without accounting for the heterogeneity in means. In the RPL model with heterogeneity in means and variances, the absolute marginal effect values of male gender on all injury severity levels are reduced to some degree for passenger car rollover crashes in 2013–2017 (Fig. 4).

6.1.4. Effect of young drivers and old drivers

Young drivers (age less than 25) are a significant factor for minor injury on passenger car rollover crashes in 2008–2012. As for the model estimated from the overall dataset, the impact of young drivers is not as significant as it is in passenger car rollover crashes in 2008–2012. This reflects the necessity of segmentation in this study since the greater impact in passenger car rollover crashes in 2008–2012 will be ignored without segmentation.

Old drivers (age greater than 50) are likely to sustain severe injury in passenger car rollover crashes in 2013–2017. The effects of older drivers do not show significant differences in the RPL model with heterogeneity in means compared to the RPL-only model.

6.2. Road Attributes

6.2.1. Effect of curve-level and curve-adverse

Curve-level road and curve-adverse (curve with hillcrest, grade, and bottom) road can both contribute to the severe injury and minor injury for light truck rollover crashes. The curve-adverse road can slightly reduce the probability of the severe injury for passenger car rollover crashes in 2008–2012, demonstrating the

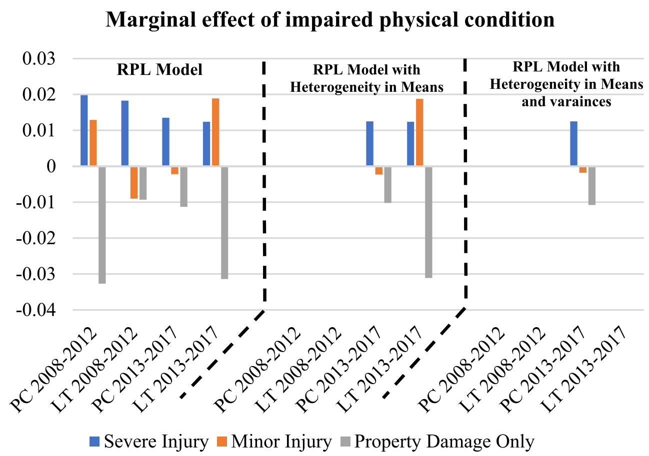

Fig. 2. Marginal effects of impaired physical condition (PC: Passenger Cars; LT: Light Trucks).

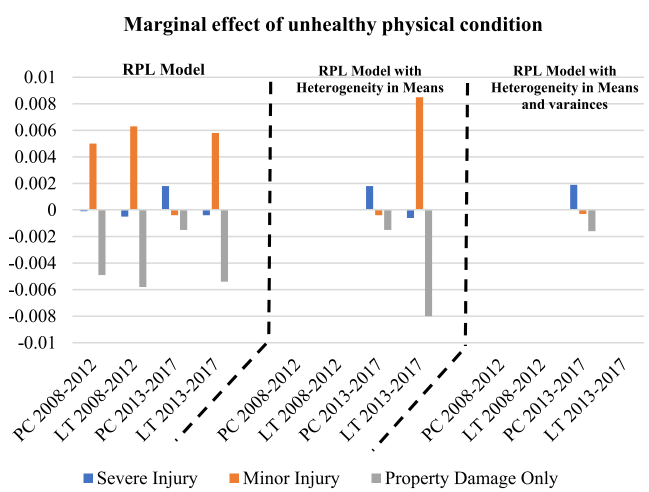

Fig. 3. The marginal effects of unhealthy physical condition (PC: Passenger Cars; LT: Light Trucks).

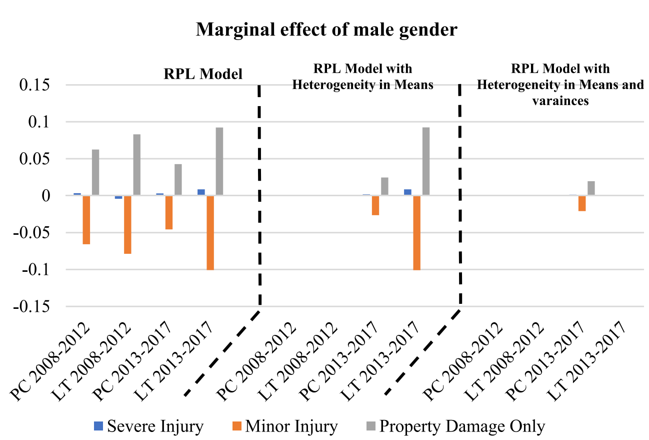

Fig. 4. Marginal effects of male gender (PC: Passenger Cars; LT: Light Trucks).

instability in effects of factors between different vehicle types. It can also be observed that road curvature contributes to the severe and minor injuries for light truck rollover crashes. This is acceptable since vehicles may easily fall off the road while they are turning around on the curve road. Nevertheless, these two factors no longer present a significant association with driver injuries in 2013–2017. In contrast to intuition, the curve-level can increase the probability of severe injury to a greater extent compared to the factor of curve-adverse. This may be because vehicles often have a relatively lower speed when they are moving on a slope, at the hillcrest, or at the bottom.

6.2.2. Effect of road-adverse (wet, watery, icy, snowy, sandy, muddy, dirty, or graveled)

Except for the model of passenger car in 2008–2012, all of the models of sub-segmentation show that the factor of road-adverse reduces the risk of the driver suffering from severe injury and minor injury, as shown in Fig. 5. Although wet, watery snowy, icy, sandy, or dirty surface can make the road slippery and the vehicle can be easily out of control, vehicles often move slower on roads under these conditions after being informed, which may explain these against-expectation results. Additionally, such impact was enhanced in 2013–2017 for passenger car rollover crashes but reduced a bit for light truck rollover crashes in 2013–2017, showing instability in the effect of the factor between different vehicle types and time periods to some extent.

As for the model structure, although not obvious, the mitigatory impact on the minor injury is lessened in the RPL model with heterogeneity in means for passenger car rollover crashes in 2013–2017. To be specific, the reduction in the probability of resulting minor injury changes from 0.0197 to 0.0184. In the RPL model with heterogeneity in means and variances, the number again decreases from 0.0184 to 0.0167. Hence the conclusion can be made that there may be an overestimation of the effect of this factor when using the RPL model without accounting for possible heterogeneity in means and variances.

6.2.3. Effect of 3–4 lanes and rural collector road

The factor of 3–4 lanes has a significant impact towards severe injury for light truck rollover crashes in 2008–2012 and 2013–3017. In 2013–2017, such impact is mitigated as the marginal effect goes down from 0.0034 to 0.0024. The reason why 3–4 lanes are riskier relative to 1–2 lanes (reference base) may be due to the higher speed limits that often present in 3–4 lanes.

Rollover crashes on the rural collector road in 2008–2012 are more likely to incur minor injury. For light trucks in 2013–2017, this factor is even more dangerous as a significant contribution toward severe injury is observed. A collector road diverts the local traffic to the arterial, hence such roads often present higher traffic volumes relative to the urban local road, which is selected as the base reference.

Referring to the model structure, the factor of 3–4 lanes and rural collector are contained in the RPL model with heterogeneity in means for light trucks rollover crashes in 2013–2017, nevertheless, no significant difference in the marginal effect is observed

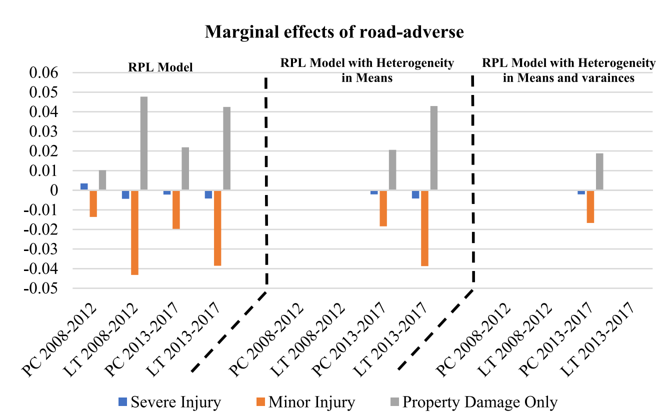

Fig. 5. Marginal effects of road adverse (PC: Passenger Cars; LT: Light Trucks).

between the RPL model and RPL model with heterogeneity in means.

6.2.4. Effect of undivided road

Undivided road (one-way undivided road and two-way undivided road) is a significant contributing factor for rollover crashes in 2013–2017. Except for a minor increase in the probability of severe injury, the undivided road is an overall benign factor as it substantially increases the probability of PDO. In terms of vehicle type, the factor of the undivided road shows a greater influence on passenger car rollover crashes than light truck rollover crashes.

As for the model structure, overestimate might occur for the mitigatory effect of the undivided road without considering potential heterogeneity in means for the passenger car rollover crashes in 2013–2017 (from 0.0269 to 0.0238) (Fig. 6).

6.2.5. Effect of other significant factors

Except for the factors mentioned above, rural local road and speed limits of 30–50 mph are also significant factors toward driver injury severities in rollover crashes. The rural local road increases the probability of minor injury by 0.0142 in the passenger car rollover crashes in 2008–2012. In the model for the overall dataset, it also increases the probability of severe injury by 0.0046. Compare to the base reference, which is the urban local road, the rural local road may present higher speed due to fewer vehicles on the road. Hence, it is more likely to result in higher injury severity levels. In the light truck rollover crashes in 2008–2012, 30–50 mph speed limit is relatively a benign factor as it can reduce the probability of minor injury (−0.007), while slightly increasing the probability of the driver suffering a severe injury (0.0006). 30–50 mph is a medium speed limit range, hence it is expected to be risker more or less. These effects are overall consistent with expectations.

The rural arterial road and the speed limits greater than 50 mph are the factors that are only found to be significant in the model estimated for the overall dataset. The rural arterial road can increase the probability of the driver suffering severe injury by 0.0028, while reducing the probability of minor injury by −0.0016. Features like higher speeds in the rural arterial road are very likely to cause severe injuries once rollover crashes occur. The speed limit of greater than 50 mph can contribute to the probability of minor injury by 0.0176, while slightly decrease the probability of severe injury by 0.0013. A speed limit that is greater than 50 mph is a risky factor since the marginal effect for the minor injury is much greater than the one for severe injury, which is understandable as a higher speed limit naturally causes higher damages in rollover crashes.

6.3. Environment attributes

6.3.1. Effect of mountain terrain

Mountain terrain is potentially a risky terrain for it is featured with the graded and hilly road, which may increase the challenges of driving task. Nevertheless, this model shows a positive association for no injury and a negative association for severe injury in 2013–2017 for both passenger car and light trucks. As for minor injury, a significant reduction in the probability of minor injury is observed for the passenger car in 2008–2012, light truck in 2013–2017, and the overall dataset as well. In terms of temporal instability, a reverse impact on the severe injury is observed for passenger car rollover crashes between 2008–2012 and 2013–2017. As for the model structure, this factor is not sensitive to the consideration of possible heterogeneity in means and variances (Fig. 7).

6.3.2. Effect of rolling terrain

The factor of rolling is found to be statistically significant in light truck rollover crashes for 2008–2012, and both two-vehicle type rollover crashes in 2013–2017. In terms of temporal instability, the effect on minor injury is intensified more than five times for light truck rollover crashes in 2013–2017 (from 0.003 to 0.0164) in the RPL-only model. However, the mitigatory impact on severe injury is also reduced. As for vehicle type, in 2013–2017, the effect of rolling has a similar influence over light truck rollover crashes and passenger car rollover crashes. In terms of model structure, compared to the RPL- only model, a higher marginal effect result on the minor injury is observed from the RPL model with heterogeneity in means. Nevertheless, the RPL model with heterogeneity in means may have overestimated the effect of rolling according to the results from the RPL model with heterogeneity in means and variances (Fig. 8).

6.3.3. Effect of cloudy weather

Similar to the effect of rain, cloudy weather can significantly contribute to the minor injury of the passenger car rollovers in 2013–2017, with a minor negative impact on the severe injury. This effect is somewhat reduced in the overall dataset with the same direction of effects on all injury severity levels. By accounting for possible heterogeneity in means and variances in passenger car rollover crashes in 2013–2017, the effect on the minor injury has decreased (0.0102 in RPL, 0.0094 in RPL with heterogeneity in means and variances), which indicates an overestimate of effects in the RPL model without accounting for possible heterogeneity in means and variances (Fig. 9).

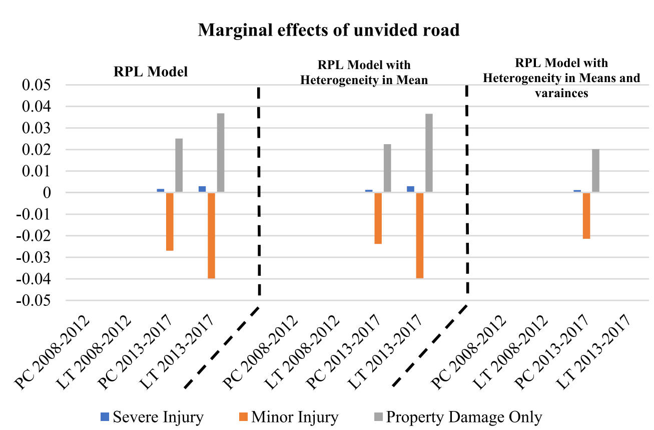

Fig. 6. Marginal effects of undivided road (PC: Passenger Cars; LT: Light Trucks).

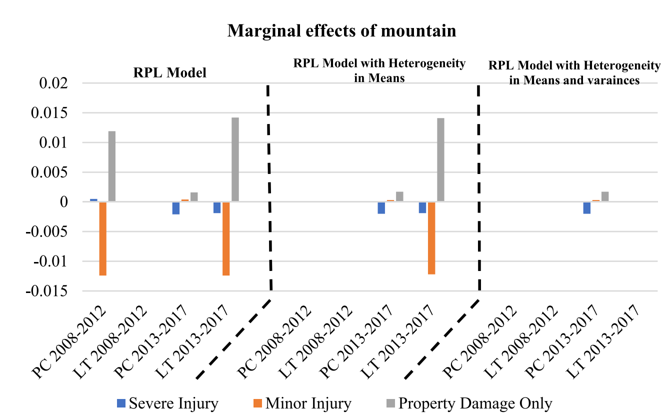

Fig. 7. Marginal effects of mountain terrain (PC: Passenger Cars; LT: Light Trucks).

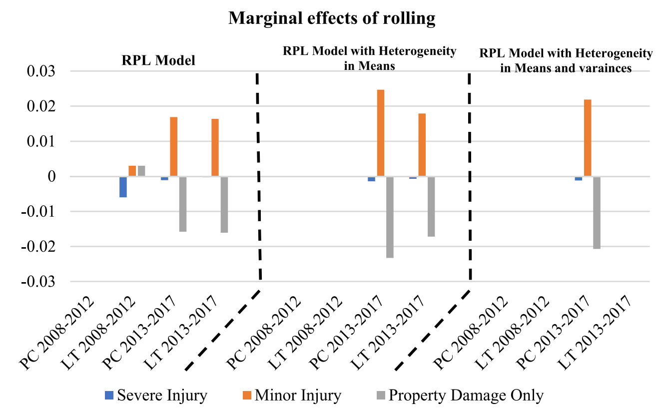

Fig. 8. Marginal effects of rolling (PC: Passenger Cars; LT: Light Trucks).

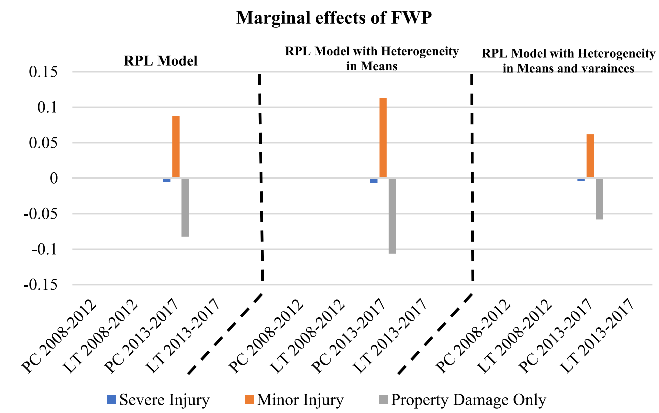

Fig. 10. Marginal effects of the FWP (PC: Passenger Cars; LT: Light Trucks).

6.3.4. Effect of FWP

FWP denotes the development of farms, woods, and pastures. These areas may present lower speed limits and few traffic volumes. Nevertheless, a considerable portion of truck volumes can be expected as many heavy commercial items may exist in the farms and woods. In the estimated model for passenger car rollover crashes in 2013–2017, FWP contributes to the minor injury to a large extent. As for the model structure, the direction of the marginal effects on each severity level is the same accounting for possible heterogeneity in means (and variances). Though the RPL model with heterogeneity in means indicates a higher effect on the minor injury, the RPL model with heterogeneity in means and variances shows a lower effect on the minor injury. With that being said, the RPL model with heterogeneity in means may have overestimated the contribution of FWP on the minor injury without further considering potential heterogeneity in variances (Fig. 10).

6.3.5. Effect of other significant factors

Except for the factors mentioned above, rain, other adverse weather (snow, sleet, hail, freezing rain/drizzle), intersection, and commercial are the factors that only the model estimated from the overall dataset contains. Commercial development is featured with high traffic volumes, mixed vehicle types, and low-speed limits. Commercial factor has a contradictory effect on severe injury and minor injury (positive marginal effect value on severe injury and negative marginal effect value on minor injury). Nevertheless, the marginal effect on the minor injury is much more significant than the one on the severe injury (0.0027 vs. 0.0002). The low speed in the commercial areas may be the reason for such a finding.

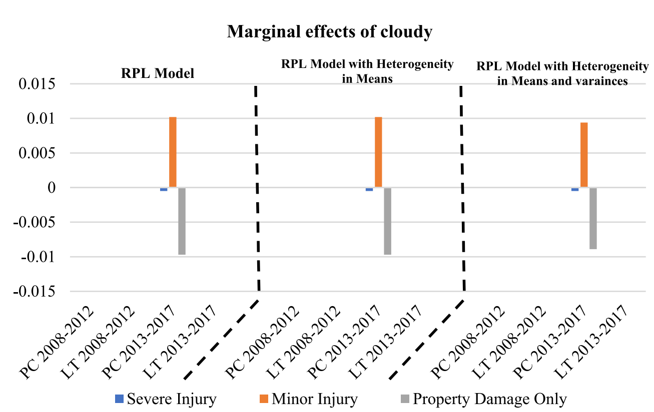

Fig. 9. Marginal effects of the cloudy (PC: Passenger Cars; LT: Light Trucks).

Under the weather of rain, the vision capability of drivers will be compromised, and rollover crashes might thus occur. Nevertheless, under the weather of rain, drivers would be advised to drive slower than normal. With a joint consideration of these two points, it would be understandable that rain increases the probability of the minor injury and decreases the probability of severe injury for light truck rollover crashes in 2008–2012. The effect of rain is identical in the RPL model and the RPL model with heterogeneity in means. In the model of the overall dataset, with the same direction (signs of marginal effect values) of effects, the effect of rain on the minor injury has decreased a bit. The intersection is featured with higher traffic volumes, low speeds, and potential turning movements. The marginal effect of intersection in light truck 2013–2017 shows that the probabilities of severe injury and minor injury both decrease if crash locations are within an intersection. As for the factor of other adverse weather, similar to rain and cloudy, a mitigatory effect is found for the minor injury.

7. Recommendations

This section provides a summary of the factors mentioned above and attempts to provide relevant recommendations to mitigate the effects of risk factors.

For driver attributes, impaired physical condition and unhealthy physical condition both have malignant effects on the injury severities of drivers. Such malignant effects of the impaired physical condition have decreased for passenger cars in more recent years, which may result from the safety policy implementations or the improvement of safety awareness of road users. Nevertheless, the malignant effect on truck rollover crashes becomes more significant in more recent years. The unhealthy physical condition seems to have a more serious impact on light truck rollover crashes compared to passenger car rollover crashes, which is understandable as light truck drivers often have more complicated driving tasks. Indeed, to mitigate the effect of impaired physical condition and unhealthy physical condition, multiple parties can be stakeholders on this issue, such as medication provider, public transportation agents, drivers, and even automobile manufacture as they can provide notable signs or signals inside the vehicles to alert drivers of the importance of maintaining proper physical condition. Even further, as technology evolves, aided driving system could potentially mitigate such injury severities.

Young drivers could contribute to the minor injury in the passenger rollover crashes in 2008–2012, while older drivers tend to suffer a severe injury in the light truck rollover crashes in 2013–2017, which are consistent with pre-expectations. According to this result, older drivers may be provided with more protective equipment when they are driving, especially light trucks.

As for road attributes, both curve level, and curve adverse deserve the attention of decision-makers. These two factors are significant in 2008–2012, but no longer significant in the models for 2013–2017. In terms of vehicle types, passenger car is more vulnerable to severe injury on the curve-level road, and minor injury on the curve-adverse road. Road signs can be installed on curve road locations to warn the drivers to slow down. Due to the relatively stable negative impacts of curvatures, designs on curved roadways might need further improvements in the near future. Moreover, road fences may also be provided if funding is available. Road-adverse seems to be mitigatory factors that do not need immediate attention. For the passenger car rollover crashes, the mitigatory effect of road adverse has enhanced in 2013–2017 compared to 2007–2012, which indicates the effectiveness of existing policies. Roads with 3–4 lanes could contribute to severe injuries in both two time periods. This contribution is reduced in 2013–2017. Compared to passenger cars, drivers operating light trucks need to practice more caution as roads with 3–4 lanes only increase the probability of minor injury for drivers operating passenger cars. Rollover crashes on rural collectors are likely to result in minor injury. The undivided road is an overall mitigatory factor and therefore does not need immediate attention either. As high-speed limits are identified as a significant contributing factor toward minor injury in this research, lower speed limits may be considered to mitigate the effects of 3–4 lanes and rural collector road.

In terms of environmental attributes, the mountain terrain and rolling terrain are both significant contributing factors to rollover crashes in 2013–2017. Compared to the mountain terrain, the rolling terrain is a riskier factor as the contribution toward the minor injury is relatively high. Possible countermeasures could be setting lower speed limits on such areas to ensure a safer driving environment, and installing road fences. Rain and cloudy weather are likely to contribute to minor injury in rollover crashes. The reason behind this may be the compromised vision capability under such weather conditions. Lower speed limits will allow drivers more reaction time to avoid the occurrence of rollover crashes. In addition, turning on the headlight of vehicles should be encouraged under such weather to improve the vision conditions. The probability of experiencing a minor injury in FWP areas (farms, woods, pastures) is higher than in areas with other land development types. When a rollover crash occurs, vehicles in these areas may hit objects such as trees and animals, which may cause minor injury to the drivers. To mitigate such effects, fences could be installed along the road if funding is available. The intersection is relatively a mitigatory factor that decreases the likelihood of severe and minor injury for drivers in light truck rollover crashes in 2013–2017.

8. Conclusions

The rollover crash is a crash type that tends to cause great economic loss to society. Investigation on the contributing factors towards injury severities of rollover crashes can help planners and decision-makers to form efficient policies mitigating the damages of rollover crashes. Nevertheless, the effects of contributing factors in rollover crashes to the injury severities may present instability with time periods and vehicle types. This research attempts to investigate these effects considering different periods and vehicle types, specifically, passenger cars and light trucks in 2008–2012 and 2013–2018. The rollover crash data were extracted from HSIS and further segmented into four groups based on the vehicle types and time periods. The RPL model with heterogeneity in means and variances was employed to analyze the factors contributing the injury severities of drivers in rollover crashes. Heterogeneity in means was found in the passenger car and light truck rollover crashes in 2013–2017, as well as the overall dataset. Besides, heterogeneity in means and variances was found in the passenger car rollover crashes in 2013–2017. The segmentations in this research were also justified through test statistics. RPL models with heterogeneity in means and variances were also able to yield better fitting model performances. This research further discussed possible solutions to mitigate the undesirable effects of factors.

It is found that factors may show a different effect on certain injury severities in the RPL with heterogeneity in means and RPL with heterogeneity in means and variances, such as FWP, unhealthy physical condition, rolling terrain, and so forth. Without accounting for such heterogeneity, one may underestimate or overestimate the effects of such factors. Additionally, some factors present vehicle type and temporal instabilities in terms of marginal effects, such as impaired physical condition, unhealthy physical condition, road adverse, and so forth.

The findings from this research demonstrate the importance of segmentation when investigating the injury severities of rollover crashes. Policy and decision-makers can utilize the findings of this research as their important references for their strategies in improving the safety condition of existing transportation infrastructures. Since this research has underscored the effects of driver attributes on the injury severity of rollover crashes, technologies such as aided driving systems could have huge potential for alleviating injuries of rollover crashes.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors want to express their deepest gratitude for the financial support by the United States Department of Transportation, University Transportation Center through the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at The University of North Carolina at Charlotte (Grant Number: 69A3551747133).

References

Anarkooli, A. J., Hosseinpour, M., & Kardar, A. (2017). Investigation of factors affecting the injury severity of single-vehicle rollover crashes: A random-effects generalized ordered probit model. *Accident Analysis & Prevention, 106*, 399–410.

Al-Bdairi, N. S. S., Behnood, A., & Hernandez, S. (2020). Temporal stability of driver injury severities in animal-vehicle collisions: A random parameters with heterogeneity in means (and variances) approach. *Analytic methods in accident research, 26* 100120.

Azimi, G., Rahimi, A., Asgari, H., & Jin, X. (2020). Severity analysis for large truck rollover crashes using a random parameter ordered logit model. *Accident Analysis & Prevention, 135* 105355.

Behnood, A., & Mannering, F. (2019). Time-of-day variations and temporal instability of factors affecting injury severities in large-truck crashes. *Analytic Methods in Accident Research, 23*, 1001102.

Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B: Methodological, 37*(9), 837–855.

Conroy, C., Hoyt, D. B., Eastman, A. B., Erwin, S., Pacyna, S., Holbrook, T. L., ... Velky, T. (2006). Rollover crashes: Predicting serious injury based on occupant, vehicle, and crash characteristics. *Accident Analysis & Prevention, 38*(5), 835–842.

Deutermann, W. 2002. Characteristics of fatal rollover crashes (No. HS-809 438).

Gong, L., & Fan, W. D. (2017). Modeling single-vehicle run-off-road crash severity in rural areas: Accounting for unobserved heterogeneity and age difference. *Accident Analysis & Prevention, 101*, 124–134.

Greene, W. H., Hensher, D. A., & Rose, J. (2006). Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research Part B: Methodological, 40*(1), 75–92.

Highway Traffic Safety Administration, National and National Center for Statistics. n.d. TRAFFIC SAFETY FACTS 2016 A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System.
Hu, W., & Donnell, E. T. (2011). Severity models of cross-median and rollover crashes on rural divided highways in Pennsylvania. *Journal of Safety Research, 42* (5), 375–382.
Islam, M., & Mannering, F. (2020). A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. *Analytic Methods in Accident Research, 27* 100128.
Islam, S., Hossain, A. B., & Barnett, T. E. (2016). Comprehensive injury severity analysis of SUV and pickup truck rollover crashes: Alabama case study. *Transportation Research Record, 2601*(1), 1–9.
Keall, M., & Newstead, S. (2009). Induced exposure estimates of rollover risk for different types of passenger vehicles. *Traffic Injury Prevention, 10*(1), 30–36.
Khattak, A. J., & Rocha, M. (2003). Are SUVs "supremely unsafe vehicles"?: Analysis of rollovers and injuries with sport utility vehicles. *Transportation Research Record, 1840*(1), 167–177.
Khattak, A. J., Schneider, R. J., & Targa, F. (2003). Risk factors in large truck rollovers and injury severity: Analysis of single-vehicle collisions. *Transportation Research Record, 40*.
Li, Y., & Fan, W. (2020). Mixed logit approach to modeling the severity of pedestrian-injury in pedestrian-vehicle crashes in North Carolina: Accounting for unobserved heterogeneity. *Journal of Transportation Safety & Security*, 1–22.
Liu, P., & Fan, W. (2020). Exploring injury severity in head-on crashes using latent class clustering analysis and mixed logit model: A case study of North Carolina. *Accident Analysis & Prevention, 135* 105388.
Mannering, F. (2018). Temporal instability and the analysis of highway accident data. *Analytic Methods in Accident Research, 17*, 1–13.
Moore, D. N., Schneider, W. H., IV, Savolainen, P. T., & Farzaneh, M. (2011). Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accident Analysis & Prevention, 43* (3), 621–630.
Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 15*, 41–55.
Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.
Wu, Q., Zhang, G., Chen, C., Tarefder, R., Wang, H., & Wei, H. (2016). Heterogeneous impacts of gender-interpreted contributing factors on driver injury severities in single-vehicle rollover crashes. *Accident Analysis & Prevention, 94*, 28–34.
Yu, M., Zheng, C., & Ma, C. (2020). Analysis of injury severity of rear-end crashes in work zones: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 100126*.

Shaojie Liu is a Ph.D. student in the Infrastructure and Environmental Systems (INES) program at the University of North Carolina at Charlotte, Charlotte, NC, U.S.A. He is with the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE). His research topics include transportation safety data analysis and connected and autonomous vehicles.

Wei David Fan currently serves as a full professor in the Department of Civil and Environmental Engineering at The University of North Carolina at Charlotte (UNC Charlotte). He is the Director of the USDOT Tier I University Transportation Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE). Dr. Fan holds a Ph.D. (May 2004) in Civil Engineering - Transportation from the University of Texas at Austin (Hook 'em Horns!). His primary research interests include operations research, transportation network modeling & optimization, multimodal transportation systems planning and operations, traffic operations and simulation, transportation safety and crash analysis, intelligent transportation systems (ITS), and computer software development.

Yang Li is a Ph.D. student in the Infrastructure and Environmental Systems (INES) program at the University of North Carolina at Charlotte, Charlotte, NC, U.S.A. He is with the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE). Her research topics include transportation safety data analysis, and non-motorized transportation.

A. Appendix

Table A1
Estimated models for the passenger car rollover crashes in 2008–2012.

Variable	RPL model	
	Parameters estimate	t-stat
Constant [SI]	−3.578	−12.7
Constant [MI]	0.116	1.32
Driver Attributes		
young [SI] (1 if the driver is younger than 25; 0 otherwise)	−1.134	−3.38
impaired physical condition [SI] (1 if the driver is under the influence of medication; 0 otherwise	2.682	7.86
impaired physical condition [MI] (1 if the driver is under the influence of medication; 0 otherwise)	0.862	8.33
unhealthy physical condition [MI](1 illness, fatigue, fell asleep, loss of consciousness; 0 otherwise)	0.478	2.84
male [MI] (1 if the gender of the driver is male; 0 otherwise	−0.512	−7.38
Road Attributes		
curve level [SI] (1 if the road is curve and level; 0 otherwise)	−2.557	−1.45
standard deviation	3.672	2.96
curve level [MI] (1 if the road is curve and level; 0 otherwise)	0.194	2.33
curve-adverse [MI] (1 if road is curve with also hillcrest, grade or bottom)	0.206	2.31
3–4 lanes [MI] (1 if the road has 3–4 lanes; 0 otherwise)	0.274	2.27
rural collector [MI] (1 if the road is rural collector; 0 otherwise)	0.213	2.42
rural local [MI] (1 if the road is rural local; 0 otherwise)	0.209	2.28
road-adverse [SI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−5.527	−1.37
road-adverse [MI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−5.527	−1.37
standard deviation	4.087	1.91
Environment Attributes		
mountain (1 if the terrain is mountain; 0 otherwise)	−0.362	−3.68
Model Statistics		
Number of observations	3654	
Log-likelihood at convergence	−2763.5	
McFadden ρ2	0.31	
Akaike information criterion (AIC)	5563	

Table A2
Estimated models for the Light Truck Rollover Crashes in 2008–2012.

Variable	RPL model	
	Parameters estimate	t-stat
Constant [SI]	−2.345	−12.6
Constant [MI]	0.521	7.57
Driver Attributes		
male [SI] (1 if the gender of the driver is male; 0 otherwise	−0.419	−2.73
male [MI] (1 if the gender of the driver is male; 0 otherwise	−0.548	−8.15
impaired physical condition [MI] (1 if the driver is under the influence of medication, drugs, or alcohol; 0 otherwise)	1.223	8.37
unhealthy physical condition [MI](1 if the driver is illness, fatigue, fell asleep, loss of consciousness; 0 otherwise)	0.749	4.91
Road Attributes		
50 mph [MI] (1 if the speed limit is between 30 mph to 50 mph; 0 otherwise)	−0.238	−2.76
3–4 lanes [SI] (1 if the road has 3–4 lanes; 0 otherwise)	0.606	3.11
3–4 lanes [MI] (1 if the road has 3–4 lanes; 0 otherwise)	0.210	2.58
curve-level [SI] (1 if the road is curve and level; 0 otherwise)	0.678	4.12
curve-level [MI] (1 if the road is curve and level; 0 otherwise)	0.222	2.65
standard deviation	1.361	2.11
curve-adverse [SI] (1 if the road is curve with hillcrest, grade, or bottom; 0 otherwise)	0.678	4.12
curve-adverse [MI] (1 if the road is curve with hillcrest, grade, or bottom; 0 otherwise)	0.200	2.72
road adverse [SI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−1.660	−6.75
road adverse [MI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−0.720	−9.61
Environment Attributes		
rain [MI] (1 if the weather is raining; 0 otherwise)	0.273	2.4
rolling [SI] (1 if the terrain is rolling; 0 otherwise)	−0.307	−2.25
Model Statistics		
Number of observations	5765	
Log-likelihood at convergence	−4620.5529	
McFadden ρ2	0.2704582	
Akaike information criterion (AIC)	9277.1	

Table A3
Estimated Model for Passenger Car Rollover Crashes in 2013–2017.

Variable	RPL Model		RPL model with heterogeneity in the mean		RPL model with heterogeneity in the means and variances	
	Parameters estimate	t-stat	Parameters estimate	t-stat	Parameters estimate	t-stat
constant [SI]	−2.826	−19.38	−2.834	−19.44	−2.833	−18.43
constant [MI]	−0.218	−0.62	−0.981	−2.02	−0.519	−1.21
Driver Attributes						
impaired physical condition [SI] (1 if the driver is under the influence of medication, drugs, or alcohol; 0 otherwise)	1.301	5.70	1.365	5.98	1.369	5.73
male [MI] (1 if the gender of the driver is male; 0 otherwise	−1.084	−4.21	−0.640	−1.99	−0.673	−2.06
unhealthy physical condition [SI](1 illness, fatigue, fell asleep, loss of consciousness; 0 otherwise)	0.915	2.15	0.904	2.13	0.922	2.11
Road Attributes						
undivided Road [MI] (1 if the road is undivided; 0 otherwise	−0.561	−2.43	−0.180	−0.52	−0.048	−0.11
standard deviation	3.643	3.21	5.049	2.78	3.990	2.33
road adverse [SI] (1 if the road is wet, watery, icy, snowy, sand, mud, dirt, and gravel; 0 otherwise)	−0.718	−2.33	−0.717	−2.33	−0.715	−2.24
road adverse [MI] (1 if the road is wet, watery, icy, snowy, sandy, mudy, dirty, or graveled; 0 otherwise)	−1.068	−3.680	−1.044	−3.25	−1.259	−3.55
Environment Attributes						
cloudy (1 if the weather is cloudy; 0 otherwise)	0.820	2.610	0.876	2.500	−1.259	−3.55
mountain (1 if the terrain is mountain; 0 otherwise)	−1.217	−2.840	−1.217	−2.840	−1.231	−2.85
rolling (1 if the terrain is rolling; 0 otherwise)	0.701	2.330	1.561	2.760	0.886	2.20
standard deviation	6.156	2.600	−1.118	−1.830	5.793	2.72
FWP (1 if the development type if farm, woods or pastures; 0 otherwise)	1.461	3.450	1.952	3.740	1.514	3.45
Heterogeneity in mean of random parameters						
undivided Road: Male (1 if the gender of the driver is male; 0 otherwise)	NA	NA	−1.122	−1.770	−1.700	−1.990
undivided Road: impaired physical condition (1 if the driver is under the influence of medication, drugs or alcohol; 0 otherwise)	NA	NA	1.708	2.670	2.330	2.680
rolling: FWP (1 if the development is farm, woods, or pasture; 0 otherwise)	NA	NA	−1.122	−1.770	NA	NA
Heterogeneity in variances of random parameters						
Undivided Road: Mountain (1 if the terrain is mountain; 0 otherwise)	NA	NA	NA	NA	0.717	2.080
Model Statistics						
Number of observations	3426.000					
Log-likelihood at convergence	−2686.829		−2674.110		−2673.243	
McFadden ρ2	0.286		0.290		0.290	
Akaike information criterion (AIC)	5401.700		5382.200		5380.500	

Table A4
Estimated Models for Light Trucks in 2013–2017.

Variable	RPL Model		RPL model with heterogeneity in the means	
	Parameters estimate	t-stat	Parameters estimate	t-stat
Constant [SI]	−2.816	−21.020	−2.813	−21.080
Constant [MI]	0.617	5.470	0.578	5.120
Driver Attributes				
Old [SI] (1 if the driver is older than 50; 0 otherwise)	0.427	2.420	0.423	2.650
impaired physical condition [SI] (1 if the driver is under the influence of medication, drugs or alcohol; 0 otherwise	1.725	10.410	1.724	10.400
impaired physical condition [MI] (1 if the driver is under the influence of medication, drugs or alcohol; 0 otherwise)	1.050	10.090	1.029	9.900
male [MI] (1 if the gender of the driver is male; 0 otherwise	−0.704	−8.580	−0.681	−8.280
unhealthy physical condition [MI] (1 illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	0.644	3.960	0.932	4.540
Road Attributes				
3–4 lanes [SI] (1 if the road has 3–4 lanes; 0 otherwise)	0.447	2.110	0.443	2.070
rural collector [SI] (1 if the road is rural collector; 0 otherwise)	0.382	2.220	0.379	2.190
rural collector [MI] (1 if the road is rural collector; 0 otherwise)	0.188	2.530	0.180	2.440
intersection [SI] (1 if the road is within footprint of intersection; 0 otherwise)	−1.487	−2.500	−1.480	−2.490
road adverse [SI] (1 if the road is wet, watery, icy, snowy, sandy, mudy, dirty, or graveled; 0 otherwise)	−1.430	−5.890	−1.428	−5.880
road adverse [MI] (1 if the road is wet, watery, icy, snowy, sandy, mudy, dirty, or graveled; 0 otherwise)	−0.684	−8.900	−0.668	−8.690
undivided Road [MI] (1 if the road is undivided; 0 otherwise	−0.231	−2.560	−0.222	−2.470
Environment Attributes				
mountain [SI] (1 if the terrain is mountain; 0 otherwise)	−0.639	−2.760	−0.646	−2.790
mountain [MI] (1 if the terrain is mountain; 0 otherwise)	−0.342	−3.760	−0.338	−3.540
rolling [MI] (1 if the terrain is rolling; 0 otherwise)	0.140	1.900	0.172	2.120
standard deviation	0.140	1.900	0.888	2.150
Heterogeneity in Mean				
rolling: unhealthy physical condition [MI] (illness, fatigue, fell asleep, loss of consciousness; 0 otherwise)	–	–	−0.695	−2.340
Model Statistics				
Number of observations	5631			
Log-likelihood at convergence	−4332.175		−4329.560	
McFadden ρ2	0.3		0.3	
Akaike information criterion (AIC)	8702.0		8699.0	

Table A5
Estimated Models for the Overall Dataset.

Variable	RPL Model		RPL model with heterogeneity in means	
	Parameters estimate	t-stat	Parameters estimate	t-stat
Constant [SI]	−2.826	−18.830	−2.807	−18.970
Constant [MI]	0.491	8.070	0.494	8.120
Driver Attributes				
young [SI] (1 if the driver is younger than 25; 0 otherwise)	−0.233	−2.470	−0.236	−2.510
impaired physical condition [SI] (1 if the driver is under the influence of medication; 0 otherwise	1.658	17.010	1.642	16.840
impaired physical condition [MI] (1 if the driver is under the influence of medication; 0 otherwise)	0.737	16.090	0.738	16.110
male [MI] (1 if the gender of the driver is male; 0 otherwise	−0.539	−16.980	−0.539	−16.980
unhealthy physical condition MI] (1 illness, fatigue, fell asleep, loss of consciousness: 0 otherwise)	0.574	7.430	0.572	7.410
old [SI] (1 if the driver is older than 50; 0 otherwise)	−1.567	−1.680	−1.567	−1.680
standard deviation	2.303	3.250	−0.530	−0.660
Road Attributes				
curve level [SI] (1 if the road is curve and level; 0 otherwise)	0.547	5.240	0.539	5.250
curve level [MI] (1 if the road is curve and level; 0 otherwise)	0.129	3.430	0.129	3.430
rural arterial [SI] (1 if the road is rural arterial; 0 otherwise)	0.529	3.650	0.514	3.620
rural local [SI] (1 if the road is rural local; 0 otherwise)	0.453	3.430	0.441	3.400
rural collector (1 if the road is rural collector; 0 otherwise)	0.332	2.480	0.320	2.440
road-adverse [SI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−1.259	−8.680	−1.301	−8.840
road-adverse [MI] (1 if the road is wet, watery, icy, snowy, sandy, muddy, dirty, or graveled; 0 otherwise)	−0.507	−9.540	−0.508	−9.570
curve-adverse [SI] (1 if road is curve and also hillcrest, grade or bottom)	0.321	2.580	0.315	2.580
curve-adverse [SI] (1 if road is curve and also hillcrest, grade or bottom)	0.160	3.860	0.160	3.860
undivided Road [SI] (1 if the road is undivided; 0 otherwise)	−0.450	−3.120	−0.442	−3.130
undivided Road [MI] (1 if the road is undivided; 0 otherwise	−0.206	−4.510	−0.207	−4.520
>=50 mph [MI] (1 if the speed limits is greater than 50 mph; 0 otherwise)	0.092	2.390	0.092	2.380
Environment Attributes				
mountain [SI] (1 if the terrain is mountain; 0 otherwise)	−0.558	−4.010	−0.555	−4.070
mountain [MI] (1 if the terrain is mountain; 0 otherwise)	−0.374	−8.700	−0.374	−8.700
cloudy [MI] (1 if the weather is cloudy; 0 otherwise)	0.129	2.720	0.128	2.720
rain [MI] (1 if the weather is rainy; 0 otherwise)	0.204	2.880	0.208	2.930
commercial [MI] (1 if the development is commercial; 0 otherwise)	−0.252	−3.310	−0.253	−3.320
intersection [MI] (1 if the road is within the footprint of intersection; 0 otherwise)	−0.320	−4.250	−0.319	−4.240
Heterogeneity in mean				
Old: Male (1 if the driver is male; 0 otherwise)	–	–	−0.465	−1.950
Old: Rain	–	–	0.879	2.000
Model Statistics				
Number of observations	18,476			
Log-likelihood at convergence	−14417.691		−14414.480	
McFadden ρ2	0.29		0.29	
Information criterion (AIC)	28891.4		28,889	

Journal of Safety Research 78 (2021) 251–261

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigating the effectiveness of safety countermeasures at highway-rail at-grade crossings using a competing risk model

Amin Keramati [a], Pan Lu [b,*], Yihao Ren [a], Denver Tolliver [a], Chengbo Ai [c]

[a] Upper Great Plains Transportation Institute/North Dakota State University, NDSU Dept. 2880, P. O. Box 6050, Fargo, ND 58108-6050, United States
[b] Department of Transportation, Logistics, and Finance, Upper Great Plains Transportation Institute/North Dakota State University, NDSU Dept. 2880, P. O. Box 6050, Fargo, ND 58108-6050, United States
[c] Department of Civil and Environmental Engineering/University of Massachusetts Amherst, 142A Marston Hall, University of Massachusetts, Amherst, MA 01003, United States

ARTICLE INFO

Article history:
Received 15 December 2020
Received in revised form 27 January 2021
Accepted 30 April 2021
Available online 14 May 2021

Keywords:
Accident prediction
Railroad grade crossing
Competing risk models
Counter-measure effectiveness

ABSTRACT

Introduction: Highway-rail at-grade crossings (HRGCs) are critical locations where a railway and a roadway intersect with one another. Crashes at those locations often result in fatalities and economic and social damages due to the impacts on both road and rail users. The main purpose of countermeasures at HRGCs is to permit safe and efficient rail and highway operations. *Method:* Countermeasures at highway-rail grade crossings (HRGCs) considered in this study include all traffic control devices and other warning and barrier devices at or on approaches to crossings. In general, active devices are commonly accepted as more effective countermeasures than passive devices. However, many of the previous effectiveness studies are either at the project level or were conducted without considering the before-improvement condition. This study focuses on the network-level marginal effectiveness of countermeasures on crash rate and severity levels during the 29-year study period from 1990 to 2018 by fully considering before-improvement control levels. A competing risk model (CRM) is able to accommodate the competing nature of crash severities as multiple outcomes from the same event of interest, which is crash occurrence in this study. Subsequently, CRM is used in this study as an integrated one-step estimation approach that investigates both crash frequency and severity likelihood over time. *Results:* The study findings indicate that adding audible devices to crossings already equipped with gates will result in a considerable annual decline in crash occurrence likelihood (0.25%). The same device installed at crossings already controlled by gates and flashing lights results in less reduction in crash occurrence likelihood of 0.14%. Moreover, adding a stop sign to the active crossing controls of gates, standard flashing lights, and audible devices will lead to a decrease in the probability of crash occurrence and severe crashes (injury and fatal). However, adding stop signs to crossings equipped only with crossbucks will increase the crash occurrence.

1. Introduction and literature review

Locations where a roadway and a railway intersect each other at the same level are recognized as highway-rail grade crossings (HRGCs). These locations are potential points of conflict between roadway traffic and trains. At HRGCs, roadway users should always yield to the train (trains have the right-of-way); and the objective of countermeasures at crossing locations is to efficiently assist roadway users in recognizing the need to yield to train traffic and safely cross over HRGCs. Between 1981 and 2018, U.S. crash frequency declined by about 76% at HRGCs (FRA, 2018). This reduction can be attributed to upgrades of passive crossing controls to active controls (Lenné et al., 2011; Meeker, Fox, & Weber, 1997; Millegan, Yan, Richards, & Han, 2009). Passive crossing controls (e.g., crossbucks signs and stop signs) are generally believed to be less effective warnings to highway users compared with active controls such as flashing lights, audible devices (bells), and gates. Although grade crossing collisions, fatalities, and injuries all have fallen nearly every year since 1981, both fatality and injury rates per crash at HRGCs have increased by about 5% and 2%, respectively, from 1981 to 2018 (FRA, 2018). Research focusing on countermeasures' effects on crash occurrence likelihood is important, however understanding countermeasures' effects on crash severity is also greatly needed. Therefore, investigating and quantifying the countermeasures' effects on HRGC's safety performance in terms of both crash occurrence and crash severity simultaneously is needed.

* Corresponding author.
E-mail addresses: Amin.Keramati@ndsu.edu (A. Keramati), pan.lu@ndsu.edu (P. Lu).

https://doi.org/10.1016/j.jsr.2021.04.008
0022-4375/

Considerable research has focused on HRGC crash frequency using the crossing inventory database (Guadamuz-Flores & Aguero-Valverde, 2017; Heydari, Fu, Thakali, & Joseph, 2018; Lu & Tolliver, 2016; Zheng, Lu, & Pan, 2019). Another group of studies focused on crash severity analysis. Categorical outcome modeling, historical police reports, and FRA HRGC crash datasets are often used in those studies (Fan, Kane, & Haile, 2015; Ghomi, Bagheri, Fu, & Miranda-Moreno, 2016; Zhao, Iranitalab, & Khattak, 2018; Zheng, Lu, & Lantz, 2018). All those analyses primarily focused on improving the models' prediction performance.

Relatively few studies have focused on quantifying safety factors' effects on crash occurrence and severity levels in the same model and the same database (Abdel-Aty & Nawathe, 2006; Keramati, Lu, Zhou, & Tolliver, 2020; Ye, Pendyala, Shankar, & Konduri, 2013; Zalinger, Rogers, & Johri, 1977). It is important for road and railway safety agencies to predict crash frequency and the likelihood of crash severity outcomes consistently and simultaneously because they need to analyze the impact of common contributors on both crash frequencies and severities based on the same set of available data. Separate forecasting models are useful to identify affecting factors for crash occurrence and severity, respectively, however, several application obstacles exist including inconsistent contributors, nontransferable probabilities among severity and occurrence, and unmeasurable variance accounted in different error terms (Keramati et al., 2020). In this study, the competing risk model (CRM) is selected because of its ability to quantify contributors' effects on crash occurrence and severity likelihood simultaneously by estimating contributors' (particularly countermeasures) marginal effects and instantaneous risk.

Some publications and reports focused on the effect of specific crossing controls (e.g., crossbucks signs, gates, and stop signs) on HRGC safety outputs (Carroll, Lee, Haines, & Hellman, 2002; Eluru, Bagheri, Miranda-Moreno, & Fu, 2012; FHWA, 2009; Haleem, 2016; Heathington, Fambro, & Richards, 1989; Kim et al., 2002; Lee, Nam, & Moon, 2004; Lerner, 2002; Liu & Khattak, 2017; Liu, Khattak, Richards, & Nambisan, 2015; Noyce & Fambro, 1998; Ogden & Cooper, 2019; Siques, 2002; Washington & Oh, 2006; Yan, Richards, & Su, 2010; Zhao et al., 2018). Table 1 lists the most common traffic controls for HRGCs that are researched as countermeasures in literature, and their design details are also summarized (FHWA, 2021).

Haleem (2016) used the mixed logit and binary logit models to investigate the significant traffic causality contributors at private HRGC. Their study results suggested that the presence of warning devices at both approaching roads of private crossings result in fewer injuries and fatalities compared to warning devices installed only on one side of the road. In addition, their results indicate that non-presence of advance warning signs increases the probability of injury and fatal crashes. Liu and Khattak (2017) conducted a spatial analysis of gated crossing collisions by using a methodology that combined path analysis modeling and geo-spatial analysis. Their findings indicate that the probability of a gate violation resulting in a crossing crash is associated with factors such as the presence of two or three quadrant gates, higher train speeds, and male drivers. Liu et al. (2015) also used path analysis to investigate indirect effects of crossing controls on crash severity changes. Their findings indicate that there is not a significant direct association between crossing controls and crash severity changes. However, their results showed a significant correlation between crossing controls and pre-crash behaviors, and also between pre-crash behavior and crash severity. Eluru et al. (2012) proposed a latent segmentation-based ordered logit model to assess the contributors' effects on crash severity at HRGCs. Their findings indicate that low-risk crossing segments are defined by higher train traffic, lower class roadways, pavement marking instead of stop signs, and the absence of permanent structures such as gates and stop signs. Washington and Oh (2006) described and applied the formalization and application of a methodology to rank 18 countermeasures (including gates, stop signs, and others) from "best" to "worst." Their results indicated that the top three safest countermeasures are in-vehicle warning systems, obstacle detection, and constant warning time. Yan et al. (2010) applied the hierarchical tree-based regression model as a nonparametric approach to predict annual crash frequency at passive crossings with crossbucks-sign-only or stop-sign-only.

Table 1
Summary of highway-rail grade crossing countermeasures.

Traffic Control Device		Design details
Passive	Crossbucks sign	Installed on right side of the highway approach to the crossing. The sign shall be retroreflectorized white. It requires highway users to yield the right-of-way to the train.
Passive	Stop sign	Installed on right side of the highway. The sign means the same as it does at a highway intersection. Highway users should stop, look, and listen for the train. Proceed when it is safe to do so.
Passive	Pavement marking	Highway-rail grade crossing pavement markings shall be 100 mm solid and retroreflectorized white. The markings should be placed on the highway 1.8 m from the nearest rail.
Active	Flashing lights	Flashing lights can be post-mounted or cantilever-mounted. Two right lights will alternatively flash to indicate highway users that train is presence. It requires do no cross when it is flashing even though the highway users could not see trains within sight distance.
Active	Bells	Bells are audible control devices. The bells will provide sound warning to highway users that train is presence. Bells are considered primarily a pedestrian warning device. It requires no crossing when it is ringing.
Active	Gate	Gate consists of a drive mechanism and a fully retroreflectorized gate arm that can extend across the approaching lanes of highway traffic. When gate arm is down, it indicates that a train is present and the road is closed. It is illegal and unsafe to cross.

All of these findings are useful in understanding the impact of specific crossing controls on either crash frequency or severity changes. However, these studies do not account for (a) pre-improvement condition differences (in other words, the effects of modifying the crossing controls on crash frequency and severity changes should be different for crossings with different pre-improvement control status), and (b) the long-term time effects of crossing-control improvements (modifications) on these changes. Moreover, HRGC characteristics, including crossing controls, may change over time, such as before and after a collision occurrence (Liu & Khattak, 2017). Consequently, quantifying crash frequency and its severity level changes the need to consider the long-term time impact and record information changes for all contributors annually. This research will focus on identifying countermeasures' effects on crash frequency and severities in one model, taking into consideration both pre-improvement control conditions and contributors' value change over a long-term analysis period (29 years in this study).

Estimating the long-term effect of countermeasures on crash occurrence and severity likelihoods can increase modeling complexity. To quantify such time effects, the primary focus shifts to the time until the crash occurred. Such analysis may be complicated because: (a) the algorithm must handle a set of contributors' associations with both the time of a crash and the crash severity level in a prediction model, and (b) crossings' crash records are collected for the limited time of the study period (e.g., 29 years in this study), so only the crash occurrence time during the study period is accessible. This situation causes right-censored data because of event-free crossings during the analysis period. Consequently, specific algorithms are needed to take these characteristics into account.

Considering these needs, in this study, the CRM is selected to investigate the crash occurrence and severity likelihood, and cause-specific Cox regression (Cox, 1972) is selected as an approach to competing risk. More detailed information regarding this method is introduced in the following section. In transportation safety analysis, the objective of CRM is quantifying crash occurrence likelihood considering the crash severity levels (PDO, injury, fatal) as competing risk events. In other words, the CRM function in HRGC safety output analysis can be defined by estimating HRGC crash likelihood during a 29-year span considering the probability of crash occurrence with one of three severity levels. In addition, censoring techniques in CRM also lead to consolidation and utilization of all the available data (including crossings with no crash records), while previous studies only used crossings with crash records as their analysis input dataset (Eluru et al., 2012; Liu & Khattak, 2017; Liu et al., 2015). In this study, by utilizing CRM, the authors investigated: (a) countermeasures' significance on both crash severity and crash occurrence of HRGCs in the same model, and (b) HRGC countermeasures' instantaneous and long-term impacts on crossings' safety performance considering different pre-improvement conditions.

2. Methodology

Survival analysis is commonly used to estimate the specific failure likelihood at a specific time, such as likelihood of death in medical science or a mechanical failure in engineering fields. Consequently, in survival analysis, modeling the time-to-event data plays a key role, and such models are designed to censor data.

The CRM is one survival analysis modeling approach and its main objective is to quantify events of interest considering more than one incident event. Correspondingly, the CRM's objective in crash analysis is estimating the likelihood of crashes as an event of interest and likelihood of experiencing each severity level as failure causes. Crash severity levels in this study are defined as property damage only, injury, and fatal. Therefore, crash severity levels, as multiple outcomes, are competing with each other, and as each crossing's crash record experienced one of these events, it is not going to experience the others. Moreover, crossings with crash records (with any severity levels) are counted as observations with specific event experience and, correspondingly, crossings with no crash records during the entire study time span are counted as censored observations. As a result of this unique survival analysis feature, all the available data, even characteristics of the crossings with no crashes, can be utilized in the study analysis (De Wreede et al., 2010; Keramati et al., 2020).

The instantaneous crash risk experiencing a specific severity level is defined as a hazard ratio and can be estimated by the cause-specific hazard function in Eq. (1) (Ishwaran et al., 2014; Putter, Fiocco, & Geskus, 2007).

$$\alpha_j(t|x) = \lim_{\Delta t \to 0} \frac{P\left\{t \le T^0 \le t + \Delta t, \delta^0 = j | T^0 \ge t, x\right\}}{\Delta t} \tag{1}$$

where

- T^0 is time of crash occurrence.
- δ^0 is crash severity level, $\delta^0 \in \{1, \cdots, J\}, J = 3$.
- x is the vector of the contributing variables.

$\alpha_j(t|x)$ denotes a cause-specific hazard function demonstrating the instantaneous crash occurrence rate at time t with j level of severity in the situation of not failing form severity level j by time t given covariate x. Note, the model is able to estimate the instantaneous crash occurrence rate by substituting $\delta^0 \in \{1, \cdots, J\}$ with $\delta^0 \in \{0, 1\}$, while $\delta^0 = 1$ indicates the occurrence of a crash regardless of the severity level and 0 indicates no crash. Let T_i^0 indicates the crash occurrence time for the i^{th} crossing, $i = 1, \ldots, n$, δ_i^0 denotes its severity level, and C_i^0indicates censoring time of crossing i. It can be indicated that T_i^0 as actual crash occurrence at the time when crossing i is unobserved and one is able to observe $T_i = min\left(T_i^0, C_i^0\right)$ and the event $\delta_i = \delta_i^0 I\left(T_i^0 \le C_i^0\right)$. Clearly, if $\delta_i = 0$, crossing i is defined as censored at time T_i (Ishwaran et al., 2014).

The Cox proportional hazard regression (Cox, 1972) is selected as an approach to quantify the cause-specific hazard regression model. Cox regression can estimate the effects of both quantitative and categorical contributors on both crash frequency and severity changes. Eq. (2) indicates the Cox proportional hazard regression formulation:

$$\alpha_j(t|x) = \alpha_{0,j}(t) \times exp\left(\beta_j^\top X\right) \tag{2}$$

where

- $\beta_j^\top$ is coefficient vector estimation of contributors' (X) effect on severity level j.
- $\alpha_{0,j}(t)$ is the baseline hazard indicating the hazard value with the condition of all X are equal to zero for severity level of j.

$exp\left(\beta_j^\top X\right)$ in Eq. (2) are equations to quantify the hazard ratio (HR). HR estimates conditional instantaneous probability of crash severity j at time t for the crossing with contributors' vector X given that the crossing is crash-free just before time t. When the HR numerical value is positive, it indicates an interesting measurement of contributors' effects on crash likelihood. Contributors with an HR value equal to 1 have no effect on crash/severity likelihood, and those with HR greater than 1 have a positive effect on hazard. Finally, contributors with an HR of less than 1 have a negative impact on hazard; in other words, the competing events are independent of each other.

Although estimates resulting from cause-specific hazard functions can shed light on the instantaneous likelihood of crash occurrence and crash severity, the model assumptions of independent censoring lead to separate estimations of the crash occurrence rate with each crash severity level. However, the probability of event occurrence in a specific range of time (e.g., [0, t]) is dependent on the cause-specific hazards of the other events (Gray, 1988). Consequently, to consolidate the estimates of competing event rates (likelihood of crash severity levels), and to calculate their marginal probability, the cumulative incident function (CIF) as another applicable output of CRM is adopted to the crash analysis problem. The event probability estimated by CIF is interpreted as the likelihood of crash occurrence with j severity by time t. CIF estimation and its association with cause-specific hazard function are defined in Eq. (3):

$$CIF_j(t|X) = p(T^0 \le t, \delta^0 = j|x) = \int_0^t \alpha_j(t|X) S(t|X) dt \tag{3}$$

In Eq. (3), $CIF_j(t|X)$ denotes the crash severity j likelihood by time t and before the occurrence of the crashes with other severity levels given contributor vector X. In addition, $S(t|X)$ indicates the event-free survival likelihood function given contributor vector X.

Simply, Eq. (4) shows that the cumulative incidence function (CIF) of crash occurrence is equal to the sum of CIF related to each crash severity level (K = 3 in this study).

$$CIF_c(t|X) = \sum_{j=1}^{3} CIF_j(t|X) \tag{4}$$

The above equation indicates that the cumulative probability of crash occurrence at time t is equal to the sum of the CIF of PDF, injury, and fatal crash.

Although the mathematical part of CRM and its outputs are complicated, the model's unique capabilities to account for time effects, to predict crash severity and occurrence in the same model, and to calculate interpretable marginal effects justify the complications. For more details regarding competing risk analysis, please refer to Kalbfleisch and Prentice (2002) and Putter et al. (2007).

3. Data

This research dataset includes both crossings with no crashes and crossings with crash records with three potential severity levels of PDO, injury, and fatal. All variables' values for 29 years are summarized in Table 2. Because nearly all variables' values changed every year, Table 2 summarized their minimum and maximum annual values for the 29-year analysis period.

Table 2
Summary statistics of considered variables in the study.

Variable Names	Categorical Variable Values	Min Frequency	Max Frequency
Crash Severity			
	No Crash	3163	3192
	PDO	2	18
	Injury	0	11
	Fatal Crash	0	6
Type of Train Service			
	Freight	2718	2807
	Intercity Passenger	387	476
Pavement Markings			
	No Marking	3111	3124
	Stop Lines	49	67
	RR Xing Symbols	16	25
Train Detection System			
	None	2398	2402
	Constant Warning Time (CWT)	375	378
	Motion Detection (MD)	42	43
	PTC	1	1
	DC	374	376
Commercial Power			
	Available	2107	2107
	Not Available	1087	1087
Roadway Paved Condition			
	Paved	563	563
	Not Paved	2631	2631
Crossing Control Types			
	Gate	4	22
	Gate + Audible	6	92
	Crossbucks + Stop Sign	44	78
	Gates + StandardFLS + Audible + Stop Signs	2	14
	Gates + StandardFLS + Audible	27	184
	Crossbucks Only	2451	2676
	Gates + CantileverFLS + Audible	2	28
	CantileverFLS + StandardFLS + Audible	2	6
	Gates + CantileverFLS + StandardFLS	1	9
	Gates + CantileverFLS + StandardFLS + Audible	2	21
Total Day Time Through Trains		0	35
Total Night Time Through Trains		0	33
Annual Average Daily Traffic		5	25,600
Highway Speed Limit (MPH)		1	70
Percent of Trucks (Percentage)		1	22.67
Distance to the Nearest Intersections (Meter)		0.78	2502
Crossing Angles (Degree)		7.9	90
Number of Traffic Lanes		1	4

The data are merged from the following three sources: (a) North Dakota roadway network, railway network, roadway intersections, and HRGCs from the ND Department of Transportation (ND GIS Hub Data Portal), which is utilized to extract crossings' geometric features; (b) highway-rail grade crossing accidents/incidents provided by the Federal Railway Administration (FRA, 2011); and (c) highway-rail grade crossing inventory from the FRA data sources, which provide current and historical crossing inventory information. The final cleaned and integrated dataset has 3,194 unique public highway-rail grade crossings in North Dakota from 1990 to 2018. Since operating and physical characteristics of a grade crossing (e.g., control devices and highway/railway traffic volume) may change over time, each crossing's contributors' records are recorded for 29 years in the analysis.

4. Results

4.1. Estimated coefficient and hazard ratio interpretation

All significant contributors' estimated coefficient (Coe) for each severity level (PDO, injury, and fatal), and crash occurrence (crash) are summarized in Table 3. The regression coefficient is estimated based on Eq. (2) (cause-specific hazard model), which shows the corresponding magnitude change in the cause-specific hazard function for each one-unit change in the contributor. As indicated in Table 3, a positive coefficient of 0.55 indicates a significant increase in hazard ratio associated with crossings with passenger train service as compared with freight train service. And a negative coefficient of −2.24 indicates a significant decrease in fatal crash hazard ratio associated with crossings with stop lines as compared with a crossing with no pavement markings.

For crash occurrence likelihood, one can see from Table 3 that pavement markings, freight train service (compared with passenger train service), DC train detection system, unavailable commercial power (compared with available commercial power), unpaved roadway (compared with paved roadway), all types of crossing control devices (except crossbucks + stop sign compared with crossbucks-only), nighttime train traffic, and truck percentage reduce the crash hazard ratio. CWT, MD train detection system, daytime train traffic, roadway travel speed, AADT, crossing angle, and number of lanes have positive impacts on crash hazard. Most of the control device impact findings met the expectations with current understanding from literature (Lenné et al., 2011; Meeker et al., 1997; Millegan et al., 2009; Raub, 2009). The reason for these impacts is primarily because it is possible that the active controls are able to better attract a driver's attention and result in greater compliance compared with the passive controls. However, several findings are worthy of additional attention and detailed research. Crossbucks and stop sign (compared with crossbucks-only) have a positive impact on crashes, indicating that adding a stop sign to a crossing already equipped with a crossbucks-only control might increase crash likelihood. The detailed marginal effect will be conducted later. Such a finding could be caused by indiscriminate use of stop signs at passive grade crossings. The FHWA established a 10-year crossbucks assemblies requirement (stopsign + crossbucks sign) for all passive crossings in 2009 (FHWA, 2009). One of the potential rationales for indiscriminate use of the stop sign could be public acceptance and understanding of how to use stop signs correctly before 2009 and after. Traffic exposures are normally believed to have an increased impact on crashes. In other words, higher traffic exposure results in higher crash likelihood. However, this study found that nighttime train traffic and truck

Table 3
Coefficient estimation.

Variable	PDO		Injury		Fatal		Crash	
	coef	Pr(>|z|)	coef	Pr(>|z|)	coef	Pr(>|z|)	coef	Pr(>|z|)
Type of Train Service (Reference: Freight Train)								
Intercity Passenger Train	0.55	0.02**	−0.08	0.82	0.62	0.14	0.34	0.06*
Pavement Markings (Reference: None)								
Stop Lines	−0.40	0.41	−0.46	0.48	−2.24	0.10*	−0.55	0.16
RR Xing Symbols	−0.68	0.28	−0.90	0.27	0.50	0.69	−0.64	0.14
Train Detection (Reference: None)								
CWT	1.533	0.00***	1.63	0.00***	2.30	0.00***	1.73	<2.2e−16***
MD	1.58	0.00***	−0.65	0.39	1.38	0.28	1.29	0.0021659***
DC	−0.65	0.02**	−0.25	0.59	−1.18	0.26	−0.44	0.06*
Is Commercial Power Available? (Reference: Yes)								
No	−0.15	0.46	−0.53	0.03**	0.26	0.42	−0.19	0.18
Is Roadway/Pathway Paved? (Reference: Yes)								
No	−0.79	0.00***	−0.25	0.38	−0.74	0.09*	−0.65	0.00***
Crossing Control (Reference: Crossbucks-only)								
Gates + CantileverFLS + StandardFLS	−1.56	0.108	−1.41	0.132	−12.71	<2.e−16***	−1.616	0.0005****
Gates + CantileverFLS + StandardFLS + Audible	−13.95	<2.e−16***	−1.76	0.020**	−1.07	0.249	−2.749	0.000***
Gates + CantileverFLS + Audible	0.51	0.412	−13.73	<2.e−16***	−18.33	<2.e−16***	−0.594	0.410
Gates + StandardFLS + Audible	−2.10	0.000***	−2.04	0.000***	−2.40	0.000***	−2.215	<2.e−16***
Gates + Audible	−12.31	<2.e−16***	0.11	0.908	−17.85	<2.e−16***	−1.202	0.262
Gates + StandardFLS + Audible*StopSigns	−1.01	0.176	−12.59	<2.e−16***	−16.27	<2.e−16***	−1.793	0.009***
Crossbucks + StopSigns	0.85	0.007***	1.36	0.000***	1.30	0.016**	1.143	0.000***
Gates	0.03	0.948	−12.65	<2.e−16***	−17.98	<2.e−16***	−0.748	0.070*
CantileverFLS + StandardFLS + Audible	−1.02	0.222	−13.27	<2.e−16***	−18.79	<2.e−16***	−1.424	0.049**
Total Daylight Thru Trains	0.23	0.00***	0.15	0.09*	−0.12	0.69	0.16	0.00***
Total Night time Thru Trains	−0.18	0.00***	−0.17	0.05*	0.17	0.57	−0.14	0.01***
Highway Speed Limit	0.01	0.53	0.02	0.28	0.04	0.10*	0.01	0.15
Annual Average Daily Traffic	0.00	0.07*	0.00	0.99	0.00	0.25	0.00	0.08*
Estimated Percent Trucks	−0.16	0.01***	−0.24	0.00***	−0.10	0.26	−0.18	0.00***
Crossing Angle	0.00	0.38	−0.01	0.08*	0.00	0.55	0.00	0.12
Number of Traffic Line	0.38	0.04**	0.20	0.39	0.05	0.88	0.31	0.03**

* Significant at a 90% confidence level.
** Significant at a 95% confidence level.
*** Significant at a 99% confidence level.

percentage have decreased impacts on crashes in general. The potential rationales behind these effects need further research. They may be rooted in operating changes and their interdependent effects. For example, a higher train volume at night might indicate reduced train volume at night, which will reduce conflict since road traffic often drops significantly at night. Few researchers investigate traffic contributions at different times. However, a few recent studies also observed the similar nighttime train traffic effects (Zheng et al., 2019; Zhou, Lu, Zheng, Tolliver, & Keramati, 2020). To truly understand independent marginal effects, one needs to conduct more detailed comparable studies.

Regarding contributors' impacts on severity levels, some factors show significant effects on specific crash severity(s) likelihoods but not on all three of them, except two crossing control combinations, gates and standard flashing lights and audible devices, and crossbucks and stop signs. The independent censoring assumption can be the main reason for such under-estimated results. Crossings with gates, standard flashing lights, and audible devices show significant negative impact on all crash severity likelihoods compared with crossings with crossbucks-only. These results match those from previous studies. For instance, Liu et al. (2015) indicated that highway users are more likely to stop at crossings with gates that also have flashing lights and audible warnings. They also showed that highway users stopping at gates is associated with lower crash severity. However, in comparison with crossbucks-only, crossbucks and stop signs indicate a significant positive impact on all PDO, injury, and fatal crashes. The reason might be rooted in the fact that stop signs are among the traffic controls often used at regular highway intersections, possibly resulting in confusion among drivers at HRGCs (Jeng, 2005).

As indicated earlier in Eq. (2), regression coefficient (β_k) represents the corresponding magnitude change in the cause-specific hazard function associated with a one-unit change in the contributor. However, hazard ratio (HR) represents the magnitude of the corresponding change in the crash occurrence and severity likelihood. In other words, HR ($exp(\beta_k^{\top}X)$) shows contributors' instantaneous crash occurrence or severity likelihoods, while the regression coefficient measures the contributors' significance in effects on HR. Table 4 represents HR for all severity levels and crash occurrence for each crossing control based on Eq. (2). In terms of a categorical variable, HR estimates the crossing's relative risk with a specific contributor's value-level compared to the reference level. For a continuous contributor, HR indicates the relative independent risk related to a one-unit change in variable (Logan, Zhang, & Klein, 2006). As indicated earlier, an HR greater than 1 indicates an increase in hazard risk, and an HR below 1 represents a decline in hazard risk. The percentage change in risk probability for each crossing control change compared with the crossbucks-only (reference level) is calculated as $|\text{HR-1}| \times 100$ and can be seen as "%impact" in Table 4.

In general, Table 4 indicates that all control devices reduce crash occurrence and fatal crash likelihood compared with crossbucks-only except crossbucks + stopsigns because all of their corresponding HR values are less than 1. Regarding PDO crash likelihood, gates + cantilevelfls + audible and gates seem to have higher PDO crash likelihoods compared with crossbucks only crossings. Regarding injury crash likelihood, gates + audible crossings will have a higher injury likelihood compared with crossbucks-only crossings. These three findings seem counterintuitive. The potential rationale could be related to roadway users'

Table 4
Crossing control hazard ratio estimation.

Variable	PDO		Injury		Fatal		Crash	
	Impct	HR	Impct	HR	Impct	HR	Impct	HR
Crossing Control (Reference: Crossbucks-only)								
Gates + CantileverFLS + Audible	67%	1.67	100%	0.000001	100%	0.00000001	45%	0.55
Gates	3%	1.03	100%	0.000003	100%	0.00000002	53%	0.47
Gates + Audible	100%	0.000005	12%	1.12	100%	0.00000002	70%	0.30
CantileverFLS + StandardFLS + Audible	64%	0.36	100%	0.000002	100%	0.00000001	76%	0.24
Gates + CantileverFLS + StandardFLS	79%	0.21	76%	0.24	100%	0.000003	80%	0.20
Gates + StandardFLS + Audible + StopSigns	64%	0.36	100%	0.0000034	100%	0.0000001	83%	0.17
Gates + StandardFLS + Audible	88%	0.12	87%	0.13	91%	0.09	89%	0.11
Gates + CantileverFLS + StandardFLS + Audible	100%	0.000001	83%	0.17	66%	0.34	94%	0.06
Crossbucks + StopSigns	134%	2.34	288%	3.88	267%	3.67	213%	3.13

psychological aggressive behavior around crossing gates (Ma, Hao, Xiang, & Yan, 2018). Note that even though they all have a greater-than-1 HR value, all of them are near 1 except for the crossbucks + stopsign, which means they are indicating a slightly positive impact. Gates for PDO shows a 3% positive impact, which is very close to no difference. Gates + audible for injury shows a 12% positive impact. For crossbucks + stopsign control device, this study consistently found that adding a stop sign to a crossing that currently has only crossbucks may increase crash occurrence, PDO, injury, and fatal crash likelihoods. As indicated earlier, this could be caused by the fact that the crossbucks assemblies requirement is relatively new and stop signs were used as a traffic control device for highway intersections rather than grade crossings in the community. For highway intersections, highway users encountering stop signs normally only need to stop and check for approaching traffic in a limited distance range. However, for grade crossings, the distance to be checked should be much longer to ensure safe operation. Moreover, because stop signs are usually found at regular highway intersections, their presence at HRGCs may cause confusion for vehicle users (Jeng, 2005). Burnham (1995) research showed that only 18% of motorists might be alerted to the stop signs and 82% were confused or semi-confused about the stop signs' presence at grade crossings. To truly and fully understand the marginal effects of such control devices, a carefully designed before-and-after comparative study is needed.

Although an analysis of both coefficient and HR provides useful information, such analyses do not yield direct estimations related to the marginal magnitude of contributors' long-term impact on risk. For example, estimated HR represents only the directional relative crash/severity likelihood changes (percent impact in Table 4). Crash/severity likelihood changes corresponding to a one-unit increase in the specific contributor while all other contributors' impacts are also included. Therefore, the CIF analysis is performed to calculate contributors' marginal effects while considering crossing controls' cumulative long-term time impacts when taking into consideration the dependency of competing events in this study.

4.2. CIF analysis

Estimation of contributors' long-term robust effects is one of the competing risk model advantages. This effect is provided by estimation of cumulative probability of the crash severity levels (PDO, injury, and fatal) and crash occurrence as the CIF with Eqs. (3) and (4), respectively.

In this section, the cumulative probability marginal effect of the following combinations of active and passive controls are assessed: (1) gates, (2) gates and audible, (3) gates and standard flashing lights and audible, (4) gates and standard flashing lights and audible and stop signs, (5) gates and cantilever flashing lights and audible, (6) cantilever flashing lights and standard flashing lights and audible, (7) gates and cantilever flashing lights and standard flashing lights and audible, (8) gates and cantilever flashing lights and standard flashing lights, (9) crossbucks (only), and (10) crossbucks and stop signs.

To estimate the cumulative probability marginal effect of each crossing control, first, the predicted CIF_j, and CIF_c for all crossings during the 29-year period are calculated by using Eqs. (3) and (4), respectively. Second, using Eq. (5), the average annual CIF of each severity level (j) is estimated for all the CIFs with the same type of crossing control. Then, the marginal countermeasure difference can be defined by Eq. (6).

$$\overline{CIF_j}(t|x_p) = \frac{\sum_{i=1}^{n} CIF_j(t|x_{pi})}{n} \tag{5}$$

where x_p is variable of specific crossing control p and $\overline{CIF_j}(t|x_p)$is the average CIF for every crossing control p and severity level j.

$$D_{j,p-q}(t) = \overline{CIF_j}(t|x_p) - \overline{CIF_j}(t|x_q) \tag{6}$$

where $D_{j,p-q}(t)$ is the marginal effect of changing crossing control from q to p for severity level j at year t.

Figs. 1 and 2 present the 29-year prediction of cumulative crash severity and occurrence likelihoods by comparing eight pairs of crossing controls. The authors compared alternative options by adding a specific device into a base option except for Fig. 1(a). The base case for Fig. 1(a) is crossbucks-only and the alternative case is changing the control device to gate-only.

According to Fig. 1(a), switching from passive control to active control will likely reduce injury and fatal crash likelihood. However, this change will increase crash occurrence and PDO likelihoods. These results seem counterintuitive because one normally expects improved safety performance in crash occurrence and in all severities if a change is made from passive control to active control. The results indicate switching the control device to gate from crossbucks-only has effects on more severe crashes, but this does not reduce PDO crashes and crash occurrence in general.

Fig. 1, part b indicates that adding gate to crossings equipped with cantilevered flashing lights, standard flashing lights, and audible warnings will reduce crash occurrence and PDO likelihoods, but will increase injury and fatal crash likelihood. In other words, upgrading crossings already equipped with flashing lights and audible devices will only reduce crash occurrences and PDO crashes but will not reduce more severe crashes. Similar findings are found in the literature. Gates were found to reduce the crash frequencies as they provide physical barriers leading to a decline in the chance of vehicle-train collisions (Austin & Carson, 2002; Elvik, Høye, Vaa, & Sørensen, 2009; Ogden & Cooper, 2019; Ogden, 2007; Park & Saccomanno, 2005; Raub, 2009). Alternatively, as a result of some drivers' pre-crash aggressive behavior, such as going around gates, gate-violations can also result in more severe crash occurrences. Therefore, some research findings indicated that the gated crossing crashes are associated with a higher

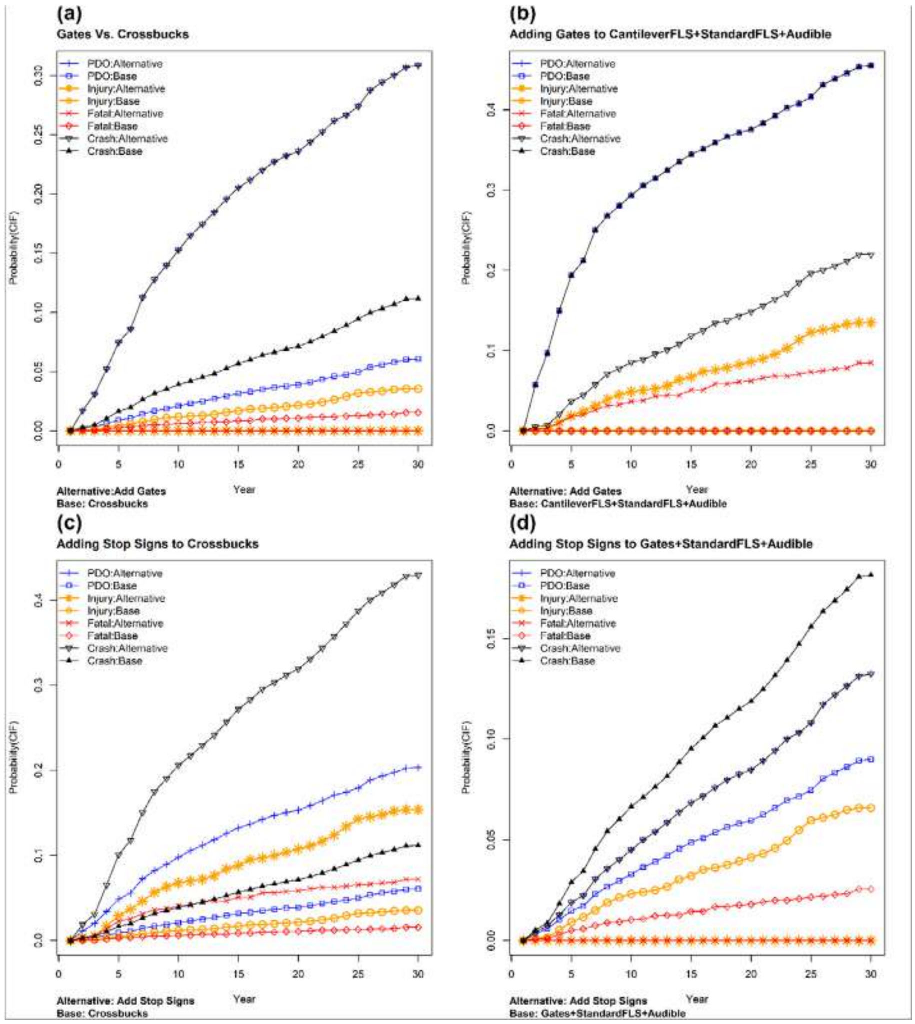

Fig. 1. Cumulative probabilities of crash severity and frequency for first crossing control pairs.

likelihood of more severe crashes (Cooper & Ragland, 2012; Raub, 2009).

Fig. 1, part c indicates that adding stop signs to crossings equipped with crossbucks will consistently increase the crash occurrence and severity likelihoods. This finding is consistent with previous coefficient and hazard ratio analyses. These probabilities are increased significantly by 284%, 235%, 333%, and 364%, respectively (annually). The potential rationales are discussed in an earlier section. However, according to Fig. 1, part d, one can see that adding stop signs to actively controlled crossings will reduce crash occurrence, injury, and fatal crash likelihood. However, PDO, a less severe crash likelihood, can be increased cumulatively by 47% in the 29-year period. These results are consistent with previous studies (Bezkorovainy & Holsinger, 1966; Burnham, 1995; Russell & Burnham, 1999; Sanders, McGee, & Yoo, 1978). For instance, according to Lerner (2002), widespread use of stop signs might have a negative effect on other passive crossing controls function (i.e., crossbucks or yield signs) as their use can decrease passive crossing controls' credibility. But in this research, the authors found adding stop signs to crossings with crossbucks-only will have a negative effect on crash occurrence and all severity levels; but adding a stop sign to an already actively controlled crossing will have additional positive effects on reducing crash occurrence and more severe crashes, but it has the negative effect of increasing the likelihood of less severe crashes such as PDO.

As can be seen from Fig. 2, part a, adding audible devices to crossings with gates, cantilevered flashing lights, and standard flashing lights will reduce crash occurrences and PDO crashes by 16% and 100%, respectively, each year. Doing so will also result in slightly reduced injury crash likelihood between years 4 and 25, and shows no effect on injury crash probability for the rest of the study period. These results are expected as the presence of audible devices warn drivers approaching the crossing (Haleem & Gan, 2015). However, this type of crossing control upgrade could result in a considerable increase in fatal crash likelihood, which is counterintuitive. To better understand and verify why adding audible devices to crossings equipped with gate and flashing lights would increase fatal crash rate, further data and research are needed.

Fig. 2, part b indicates adding an audible device to crossings with gates will reduce PDO and fatal crashes to nearly zero. In addition, this type of improvement will reduce crash occurrence by 24% cumulatively during the 29-year period. However, in this study, adding bells at crossings with gates and flashing lights will increase injury crash likelihood. In other literature, totally different conclusions were drawn. The Federal Railroad Administration (2011) research indicated that driving around or through the gates is more likely to happen at crossings with gates and flashing lights without bells, which suggests intentional trespassing behavior. However, Liu et al. (2015) suggest there is a higher possibility of driving around or through the gate at crossings with gates and

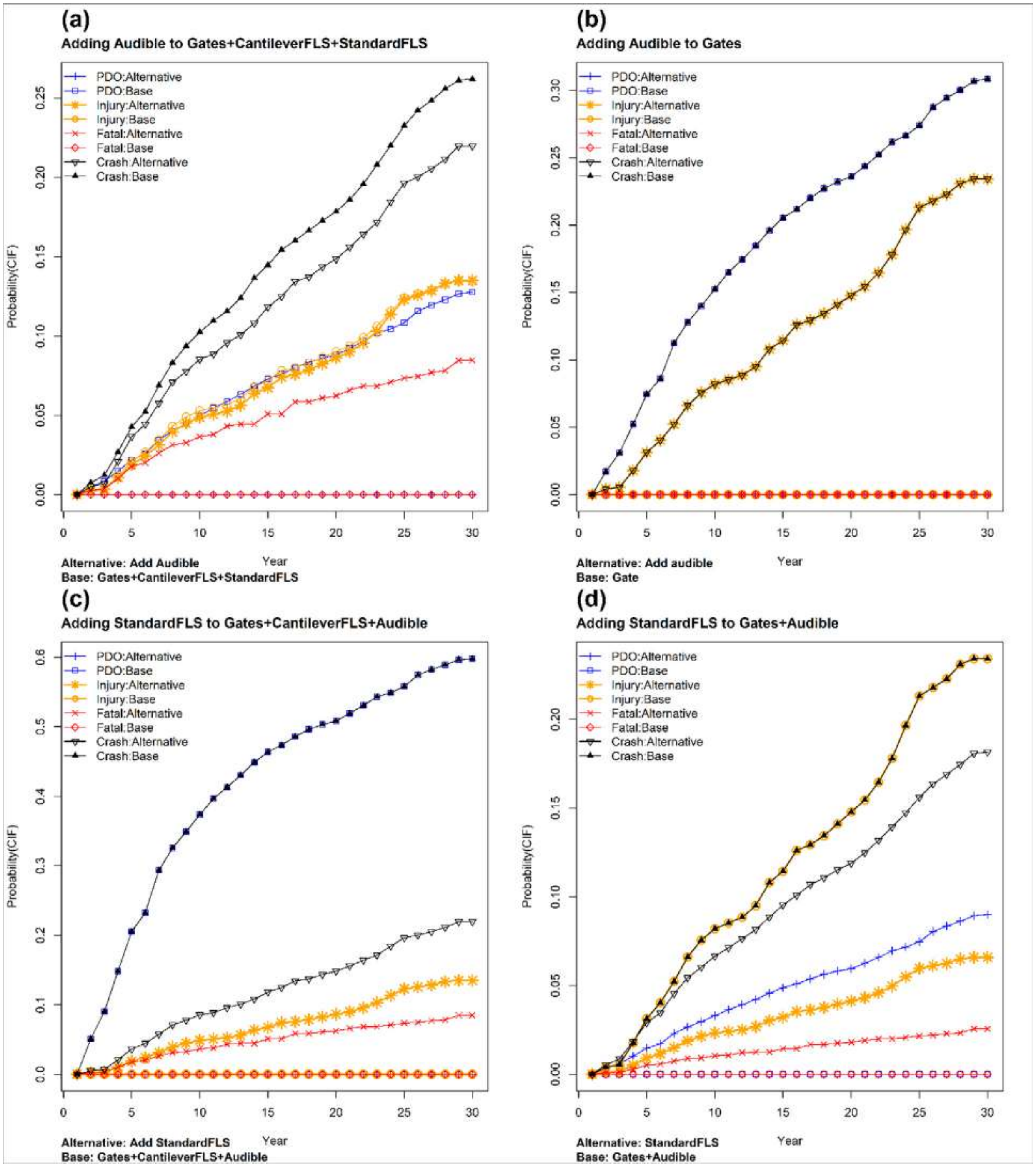

Fig. 2. Cumulative probabilities of crash severity and frequency for second crossing control pairs.

audible warnings compared with those with gates only. Increased trespassing behavior with bells may be the type of behavior that tends to result in injury crashes.

Fig. 2, part c, indicates that adding standard flashing lights to crossings with gates, cantilevered flashing lights, and audible devices, will result in decreases in PDO crash and crash occurrence but will result in increases in injury and fatal crashes. Installing standard flashing lights as supplemental flashing-light signals or side lights at the crossings with cantilevered flashing lights will increase the visibility of a crossing, thus making a higher number of vehicles aware that they are approaching a crossing or that a train is approaching. Consequently, it can be expected that crash frequency will decrease (Ogden & Cooper, 2019). With regard to its negative impact on more severe crashes, one potential reason could be associated with acute angles. Guided by the traffic control device installation manual, crossings equipped with additional pairs of light units need to be directed toward vehicular traffic approaching the HRGC from highway closely adjacent to and in parallel with the railroads. Such installations can create an acute angle between the railway track and the highway at crossings (Ogden & Cooper, 2019). Previous research (Austin, 2000; A. Keramati et al., 2020; Liu & Khattak, 2017; Oh, Washington, & Nam, 2006; Wigglesworth, 2001; Yan et al., 2010; Zhao et al., 2018) all agreed that acute crossing angles are often associated with higher levels of crash severity. Consequently, it is expected that crossings with standard flashing lights used as additional warning lights are more likely to have crashes with higher levels of severity (injury and fatal).

Fig. 2, part d, indicates that adding standard flashing lights to crossings with gates and audible devices will reduce crash occurrence and injury likelihoods, but will increase PDO and fatal crash rates. According to Lenné et al. (2011), mean vehicle speed on approaches to HRGCs declined more promptly in response to flashing lights in comparison with traffic signals. Therefore, crash occurrence is expected to be reduced. Although adding flashing lights to crossings with gates and audible devices increases the crossings' fatal and PDO crash likelihoods, the difference is small, both less than 0.1%. However, the authors believe additional analysis is needed to better understand the increase

To facilitate the quantifying of such marginal effects and to better understand the control devices' upgrading effects, Table 5 summarizes the estimated marginal effect of each crossing control change in average annual probability change during the 29-year period. Table 5 also provides summary information of Figs. 1 and 2 by providing average annual absolute crash likelihood changes.

One can see from Table 5 that upgrading crossing control to gates from crossbucks-only will likely reduce injury likelihood 0.12% while holding other variables unchanged and reduce fatal crash likelihood by 0.05%. However, the likelihood of a PDO crash will increase by 0.83% annually. Adding a gate to a crossing already controlled by flashing lights and bells will significantly reduce PDO likelihood by 1.52%, but will increase injury and fatal crash likelihood by 0.45% and 0.28%, respectively. Adding stop signs to crossings with crossbucks signs will increase crash likelihood for all three levels; and adding stop signs to a crossing already actively controlled will reduce the overall crash occurrence, but will increase PDO likelihood by 0.14%. Adding audible devices to a crossing already actively controlled will reduce overall crash occurrence but will slightly increase more severe crash likelihood. Note that adding audible devices to a crossing actively controlled by gate-only will again, in general, reduce crash occurrence likelihood by 0.25% but increase injury likelihood by 0.78%. Adding standard flashing lights to crossings that are actively controlled will reduce crash occurrence likelihood.

5. Concluding remarks

5.1. Summary

The study findings are based on 29-year empirical HRGC safety performance data in North Dakota. The study extends knowledge of countermeasures' effects on HRGC crash frequency and severity likelihood changes considering various pre-improvement condition situations. In addition, this study provided understandings on the long-term time-effects of HRGCs' countermeasures.

Estimation of countermeasures' long-term impact on both crash occurrence and severity likelihoods can increase modeling complexity while considering effects on crash occurrence and severity simultaneously. Moreover, it is always a challenge to select an unbiased predicting model to predict minority class with imbalanced data set. The competing risk model was selected as a novel method to address those issues and predict likelihoods of crash occurrence and severity simultaneously, as this model is able to: (a) estimate time-to-event outputs, (b) predict crash severity and occurrence simultaneously, and (c) quantify the effect of crossings with no crash records during the entire study time span by its ability in handling right-censored observations.

5.2. Findings and practical implications

The research findings improve the knowledge of control devices' long-term marginal effects on HRGC crash occurrence and severity likelihoods. The main takeaways are summarized below:

1) In general, adding a control device to a crossing will reduce crash occurrence likelihood except when adding stop sign to a crossing already controlled by crossbucks only.
2) Adding a control device to a crossing will reduce crash occurrence likelihood, but the effects on the three severity levels can be very different. For example, adding stop signs to a crossing passively controlled by gate, flash lights, and audible devices will reduce injury and fatal crash likelihood. However, doing so will increase PDO crash likelihood even though the overall crash occurrence likelihood is reduced. Because of this finding, the authors suggest that weighted improvement benefits among the three severity levels should be considered when agencies making safety improvement decisions
3) The same control device updating will have different marginal effect if the pre-improvement control level is different. For example, adding audible device to a crossing already controlled by gate-only compared to a crossing controlled by gate and flashing lights are different. The marginal effects on the reduced crash occurrence likelihood is higher for gate only (0.25% reduction) than for gate and flashing lights (0.14% reduction) from Table 5. Marginal effects of a control device upgrading can only be fully understood if one considers the different pre-upgrading control levels.

5.3. Study limitations

To understand the countermeasure's marginal effectiveness, before-and –after practical implementation is needed, then the precise practical effectiveness can be quantified and analyzed. However, those measurements and data have not been available and this study used ND empirical data to apply countermeasures effectiveness analysis. Consequently, the uncertainty of this empirical pilot study needs further investigation before using this information as HRGC safety decision-making frame work.

Table 5
Average annual crash likelihood change for crossing control change.

Crossing Control Change			PDO	Injury	Fatal	Crash
Gate	Replace	Crossbucks	0.83	−0.12	−0.05	0.66
	Add to	CantileverFLS + StandardFLS + Audible	−1.52	0.45	0.28	−0.79
Stop Signs	Add to	Crossbucks	0.48	0.39	0.19	1.06
		Gate + StandardFLS + Audible	0.14	−0.22	−0.09	−0.16
Audible	Add to	Gate + CantileverFLS + StandardFLS	−0.43	0.002	0.28	−0.14
		Gates	−1.03	0.78	0.00000001	−0.25
StandardFLS	Add to	Gate + CantileverFLS + Audible	−1.99	0.45	0.28	−1.26
		Gate + Audible	0.30	−0.56	0.09	−0.18

Interaction effects of contributors are not considered and tested in this study, including interaction factors that might alter the estimated coefficient of crossing controls.

5.4. Recommended directions for future research

The effect of geometric parameters of HRGCs and countermeasures (traffic control devices) interaction on their safety performance is still under researched. Therefore, future studies on interaction effect of HRGCs geometric and countermeasures are recommended to evaluate both crash likelihood and crash severity changes. Moreover, countermeasures' effects with more pre-improvement conditions should be further researched when supporting data become available. In this study, the authors find that adding stop signs to a crossing equipped with crossbucks will only increase crash occurrence and all three severity crash likelihoods. Further, better controlled experiments should be conducted to better understand the effects of cross-buck assembly with stop signs. Moreover, safety improvement decision making cannot be solely determined by the marginal countermeasures' effects, life time total cost analysis including initial cost of construction, operational cost, and maintenance cost should be conducted in future research to fully understand each countermeasure's cost-effectiveness.

Acknowledgements

The authors express their gratitude to the following funding agencies for their support: North Dakota State University and the Mountain-Plains Consortium (MPC), a university transportation center funded by the U.S. Department of Transportation.

References

Abdel-Aty, M., & Nawathe, P. (2006). A novel approach for signalized intersection crash classification and prediction. *Advance Transport Studies*, 9.

Accident Data as reported by Railroads [WWW Document], n.d. URL https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx (accessed 9.24.19).

Austin, R. D. (2000). *An accident prediction model for highway-rail interfaces*. Montana State University.

Austin, R. D., & Carson, J. L. (2002). An alternative accident prediction model for highway-rail interfaces. *Accident Analysis and Prevention, 34*(1), 31–42.

Bezkorovainy, G., & Holsinger, R. G. (1966). The use of stop signs at railroad crossings. *Traffic Engineering, 37*(2), 54–59.

Burnham, A. (1995). Stop sign effectiveness at railroad grade crossings (abuse without excuse). *Third international symposium on railroad-highway grade crossing research and safety*. Southeastern Transportation Center.

Carroll, A. A., Lee, M., Haines, M., & Hellman, A. D. (2002). Preliminary evaluation of the school street four-quadrant gate highway-railroad grade crossing. *9th world congress on intelligent transport systems ITS America, ITS Japan, ERTICO (Intelligent Transport Systems and Services-Europe)*.

Cooper, D.L., Ragland, D.R., 2012. Applying safety treatments to rail-highway at-grade crossings.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B, 34*(2), 187–202.

Eluru, N., Bagheri, M., Miranda-Moreno, L. F., & Fu, L. (2012). A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings. *Accident Analysis and Prevention, 47*, 119–127.

Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *Driver training and regulation of professional drivers', The handbook of road safety measures*. Emerald Group Publishing Limited.

Fan, W., Kane, M. R., & Haile, E. (2015). Analyzing severity of vehicle crashes at highway-rail grade crossings: Multinomial logit modeling. *Journal of the Transportation Research Forum*, 39–56.

Federal Railroad Administration (2011). FRA guide for preparing accident/incident reports. *Publication*. DOT/FRA/RRS-22.

FHWA. (2009). Manual on uniform traffic control devices (MUTCD).

FHWA – MUTCD. (2003). Edition Chapter 8 table of contents [WWW Document], n.d. URL https://mutcd.fhwa.dot.gov/htm/2003/part8/part8-toc.htm (accessed 1.23.21).

FRA (2018). All highway-rail incidents at public and private crossings, 1981–2018 [WWW Document] accessed 9.27.19 https://oli.org/about-us/news/collisions-casulties.

Ghomi, H., Bagheri, M., Fu, L., & Miranda-Moreno, L. F. (2016). Analyzing injury severity factors at highway railway grade crossing accidents involving vulnerable road users: A comparative study. *Traffic Injury Prevention, 17*(8), 833–841.

Gray, R. J. (1988). A class of $ K $-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics, 16*(3), 1141–1154.

Guadamuz-Flores, R., & Aguero-Valverde, J. (2017). Bayesian spatial models of crash frequency at highway–railway crossings. *Transportation Research Record, 2608*(1), 27–35.

Haleem, K. (2016). Investigating risk factors of traffic casualties at private highway-railroad grade crossings in the United States. *Accident Analysis and Prevention, 95*, 274–283. https://doi.org/10.1016/j.aap.2016.07.024.

Haleem, K., & Gan, A. (2015). Contributing factors of crash injury severity at public highway-railroad grade crossings in the US. *Journal of Safety Research, 53*, 23–29.

Hao, W., & Daniel, J. (2016). Driver injury severity related to inclement weather at highway–rail grade crossings in the United States. *Traffic Injury Prevention, 17*(1), 31–38.

Hao, W., & Daniel, J. (2014). Motor vehicle driver injury severity study under various traffic control at highway-rail grade crossings in the United States. *Journal of Safety Research, 51*, 41–48.

Hao, W., & Daniel, J. R. (2013). Severity of injuries to motor vehicle drivers at highway–rail grade crossings in the United States. *Transportation Research Record, 2384*(1), 102–108.

Heathington, K. W., Fambro, D. B., & Richards, S. H. (1989). Field evaluation of a four-quadrant gate system for use at railroad-highway grade crossings. *Transportation Research Record*, 1244.

Heydari, S., & Fu, L. (2015). Developing safety performance functions for railway grade crossings: a case study of Canada. *2015 joint rail conference*. American Society of Mechanical Engineers, p. V001T06A017-V001T06A017.

Heydari, S., Fu, L., Thakali, L., & Joseph, L. (2018). Benchmarking regions using a heteroskedastic grouped random parameters model with heterogeneity in mean and variance: Applications to grade crossing safety analysis. *Analytic Methods in Accident Research, 19*, 33–48. https://doi.org/10.1016/j.amar.2018.06.003.

Hu, S.-R., Li, C.-S., & Lee, C.-K. (2012). Model crash frequency at highway–railroad grade crossings using negative binomial regression. *Journal of the Chinese Institute of Chemical Engineers, 35*(7), 841–852.

Hu, S.-R., Li, C.-S., & Lee, C.-K. (2010). Investigation of key factors for accident severity at railroad grade crossings by using a logit model. *Safety Science, 48*(2), 186–194.

Hu, S., & Lin, J. (2012). Effect of time to train arrive on crash frequency at highway-railroad grade crossings: A general classification regression model. *Proceedings of the 91st annual meeting of the transportation research board, Washington, DC*.

Ishwaran, H., Gerds, T. A., Kogalur, U. B., Moore, R. D., Gange, S. J., & Lau, B. M. (2014). Random survival forests for competing risks. *Biostatistics, 15*(4), 757–773. https://doi.org/10.1093/biostatistics/kxu010.

Jeng, O.-J. (2005). Survey of driver perceptions of railroad and light rail warning devices/grade crossings.

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data 2002 New York*. NY Wiley Crossref.

Kang, Y., & Khattak, A. (2017). Cluster-based approach to analyzing crash injury severity at highway–rail grade crossings. *Transportation Research Record, 2608*(1), 58–69.

Keramati, A., Lu, P., Tolliver, D., & Wang, X. (2020). Geometric effect analysis of highway-rail grade crossing safety performance. *Accident Analysis and Prevention, 138*, 105470. https://doi.org/10.1016/j.aap.2020.105470.

Keramati, A., Lu, P., Zhou, X., & Tolliver, D. (2020). A simultaneous safety analysis of crash frequency and severity for highway-rail grade crossings: The competing risks method. *Journal of Advanced Transportation*. https://doi.org/10.1155/2020/8878911 8878911.

Khattak, A., & Luo, Z. (2011). Pedestrian and bicyclist violations at highway–rail grade crossings. *Transportation Research Record, 2250*(1), 76–82.

Khattak, A., Sharma, A., Luo, Z. (2012). Implications of using annual average daily traffic in highway-rail grade crossing safety models.

Kim, G., Kang, K., Ji, S., Chung, H., Shin, K., & Lee, C. K. (2002). Railway Gate Control System at Railroad-Highway Grade Crossing in Korea. *9th world congress on intelligent transport systems, ITS America, ITS Japan*. ERTICO (Intelligent Transport Systems and Services-Europe).

Lee, J., Nam, D., & Moon, D. (2004). A zero-inflated accident frequency model of highway–rail grade crossing. *Proceedings of the Transportation Research Board Annual Meeting. Washington, DC*.

Lenné, M. G., Rudin-Brown, C. M., Navarro, J., Edquist, J., Trotter, M., & Tomasevic, N. (2011). Driver behaviour at rail level crossings: Responses to flashing lights, traffic signals and stop signs in simulated rural driving. *Applied Ergonomics, 42*(4), 548–554.

Lerner, N. D. (2002). Traffic-control devices for passive railroad-highway grade crossings. *Transportation Research Board*.

Liu, J., & Khattak, A. J. (2017). Gate-violation behavior at highway-rail grade crossings and the consequences: Using geo-Spatial modeling integrated with path analysis. *Accident Analysis and Prevention, 109*, 99–112. https://doi.org/10.1016/j.aap.2017.10.010.

Liu, J., Khattak, A. J., Richards, S. H., & Nambisan, S. (2015). What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of pre-crash behaviors. *Accident Analysis and Prevention, 85*, 157–169. https://doi.org/10.1016/j.aap.2015.09.004.

Logan, B. R., Zhang, M.-J., & Klein, J. P. (2006). Regression models for hazard rates versus cumulative incidence probabilities in hematopoietic cell transplantation data. *Biol. Blood Marrow Transplant, 12*(1), 107–112.

Lu, P., & Tolliver, D. (2016). Accident prediction model for public highway-rail grade crossings. *Accident Analysis and Prevention, 90*, 73–81. https://doi.org/10.1016/j.aap.2016.02.012.
Ma, C., Hao, W., Xiang, W., & Yan, W. (2018). The impact of aggressive driving behavior on driver-injury severity at highway-rail grade crossings accidents. *Journal of Advanced Transportation, 2018*.
Medina, J. C., & Benekohal, R. F. (2015). Macroscopic models for accident prediction at railroad grade crossings: Comparisons with US Department of Transportation accident prediction formula. *Transportation Research Record, 2476*(1), 85–93.
Meeker, F., Fox, D., & Weber, C. (1997). A comparison of driver behavior at railroad grade crossings with two different protection systems. *Accident Analysis and Prevention, 29*(1), 11–16.
Millegan, H., Yan, X., Richards, S., & Han, L. (2009). Evaluation of effectiveness of stop-sign treatment at highway–railroad grade crossings. *Journal of Transportation Safety and Security, 1*(1), 46–60.
North Dakota GIS Hub Data Portal [WWW Document], n.d. URL https://www.nd.gov/gis/apps/Download/ (accessed 6.12.19).
Noyce, D. A., & Fambro, D. B. (1998). Enhanced traffic control devices at passive highway-railroad grade crossings. *Transportation Research Record, 1648*(1), 19–27.
Ogden, B. D. (2007). *Railroad-highway grade crossing handbook*. United States: Federal Highway Administration.
Ogden, B.D., Cooper, C. (2019). Highway-rail crossing handbook.
Oh, J., Washington, S. P., & Nam, D. (2006). Accident prediction model for railway-highway interfaces. *Accident Analysis and Prevention, 38*(2), 346–356. https://doi.org/10.1016/j.aap.2005.10.004.
Park, Y.-J., & Saccomanno, F. F. (2005). Collision frequency analysis using tree-based stratification. *Transportation Research Record, 1908*(1), 121–129.
Putter, H., Fiocco, M., & Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine, 26*(11), 2389–2430.
Raub, R. A. (2009). Examination of highway–rail grade crossing collisions nationally from 1998 to 2007. *Transportation Research Record, 2122*(1), 63–71.
Russell, E. R., & Burnham, A. (1999). A review of past and present research, guidelines and practice for the proper use of stop signs at rail-highway grade crossings. *Enhancing transportation safety in the 21st century ITE international conference*. Institute of Transportation Engineers (ITE).
Saccomanno, F. F., & Lai, X. (2005). A model for evaluating countermeasures at highway–railway grade crossings. *Transportation Research Record, 1918*(1), 18–25.
Saccomanno, F. F., Park, P.-J., & Fu, L. (2007). Estimating countermeasure effects for reducing collisions at highway–railway grade crossings. *Accident Analysis and Prevention, 39*(2), 406–416.
Sanders, J. H., McGee, H. W., Yoo, C. S. (1978). Safety features of stop signs at rail-highway grade crossings.
Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention, 43*(5), 1666–1676.
Siques, J. T. (2002). Effects of pedestrian treatments on risky pedestrian behavior. *Transportation Research Record, 1793*(1), 62–70.
Washington, S., & Oh, J. (2006). Bayesian methodology incorporating expert judgment for ranking countermeasure effectiveness under uncertainty: Example applied to at grade railroad crossings in Korea. *Accident Analysis and Prevention, 38*(2), 234–247. https://doi.org/10.1016/j.aap.2005.08.005.
Wigglesworth, E. C. (2001). A human factors commentary on innovations at railroad–highway grade crossings in Australia. *Journal of Safety Research, 32*(3), 309–321.
Yan, X., Richards, S., & Su, X. (2010). Using hierarchical tree-based regression model to predict train-vehicle crashes at passive highway-rail grade crossings. *Accident Analysis and Prevention, 42*(1), 64–74. https://doi.org/10.1016/j.aap.2009.07.003.
Ye, X., Pendyala, R. M., Shankar, V., & Konduri, K. C. (2013). A simultaneous equations model of crash frequency by severity level for freeway sections. *Accident Analysis and Prevention, 57*, 140–149.
Zalinger, D. A., Rogers, B. A., & Johri, H. P. (1977). Calculation of hazard indices for highway-railway crossings in Canada. *Accident Analysis and Prevention, 9*(4), 257–273.
Zhao, S., Iranitalab, A., & Khattak, A. J. (2018). A clustering approach to injury severity in pedestrian-train crashes at highway-rail grade crossings. *Journal of Transportation Safety & Security, 11*(3), 305–322. https://doi.org/10.1080/19439962.2018.1428257.
Zhao, S., & Khattak, A. (2015). Motor vehicle drivers' injuries in train–motor vehicle crashes. *Accident Analysis and Prevention, 74*, 162–168.
Zheng, Z., Lu, P., & Lantz, B. (2018). Commercial truck crash injury severity analysis using gradient boosting data mining model. *Journal of Safety Research, 65*, 115–124.
Zheng, Z., Lu, P., & Pan, D. (2019). Predicting highway-rail grade crossing collision risk by neural network systems. *Journal of Transportation Engineering, Part A: Systems, 145*(8), 04019033. https://doi.org/10.1061/JTEPBS.0000257.
Zheng, Z., Lu, P., & Tolliver, D. (2016). Decision tree approach to accident prediction for highway-rail grade crossings: Empirical analysis. *Transportation Research Record, 2545*(1), 115–122.
Zhou, X., Lu, P., Zheng, Z., Tolliver, D., & Keramati, A. (2020). Accident prediction accuracy assessment for highway-rail grade crossings using random forest algorithm compared with decision tree. *Reliability Engineering and System Safety*.

Amin Keramati (First author) PhD student, Department of Transportation, Logistics, and Finance. Research assistant, Upper Great Plains Transportation Institute North Dakota State University (NDSU). Mr. Keramati is current a PhD student with NDSU. His research and dissertation focus is in the area to evaluate and understand highway rail grade crossing safety performance and to design a safety improvement system to assist HRGC safety improvement decision makings. He holds 5 journal publications in the related area.

Pan Lu (Second author and corresponding author), PhD, Associate Professor, Department of Transportation, Logistics, and Finance, Associate Research Fellow, Upper Great Plains Transportation Institute North Dakota State University (NDSU). One of Dr. Lu's research area is in transportation safety.

Yihao Ren (The third author), PhD student, Department of Transportation, Logistics, and Finance. Research assistant, Upper Great Plains Transportation Institute, North Dakota State University (NDSU). Mr. Ren is a current PhD student with the department of transportation, logistics, and finance at NDSU. His research focus is in the area of data mining, machine learning, and deep learning algorithms application in infrastructure condition monitoring.

Denver Tolliver (The fourth author), Director, Upper Great Plains Transportation Institute (UGPTI), North Dakota State University (NDSU). In addition to being the director of the UGPTI, Dr. Tolliver is the director of the Mountain-Pains Consortium (MPC) – a multi-university consortium that serves as the university transportation center in federal region 8. He has been awarded more than 20 million in grant funding and authored and co-authored more than 150 transportation research papers.

Chengbo Ai (The fifth author), Assistant Professor, Department of Civil and Environmental Engineering, University of Massachusetts, Amherst. Dr. Ai's interdisciplinary research focuses on developing computational models, automated algorithms and hardware systems as they are applied in the fields of transportation asset management, geometry design and transportation safety.

Journal of Safety Research 78 (2021) 59–68

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigation of crashes at pedestrian hybrid beacons: Results of a large-scale study in Arizona

Kay Fitzpatrick [a,*], Eun Sug Park [a], Michael J. Cynecki [b], Michael P. Pratt [a], Michelle Beckley [c]

[a] Texas A&M Transportation Institute, 3135 TAMU, College Station, TX 77843-3135, USA
[b] Lee Engineering, LLC, 3610 N 44th Street, Phoenix AZ, 85018, USA
[c] Y2K Engineering, LLC, 1921 S Alma School Road Suite 204, Mesa, AZ, 85210-3035, USA

ARTICLE INFO

Article history:
Received 24 June 2020
Received in revised form 19 January 2021
Accepted 19 April 2021
Available online 4 May 2021

Keywords:
Pedestrian hybrid beacons
Crashes
Crash modification factors
Pedestrian crossings
Street crossings
HAWK

ABSTRACT

Introduction: The pedestrian hybrid beacon (PHB) is a traffic control device used at pedestrian crossings. A recent Arizona Department of Transportation research effort investigated changes in crashes for different severity levels and crash types (e.g., rear-end crashes) due to the PHB presence, as well as for crashes involving pedestrians and bicycles. *Method:* Two types of methodologies were used to evaluate the safety of PHBs: (a) an Empirical Bayes (EB) before-after study, and (b) a long-term cross-sectional observational study. For the EB before-after evaluation, the research team considered three reference groups: unsignalized intersections, signalized intersections, and both unsignalized and signalized intersections combined. *Results:* For the signalized and combined unsignalized and signalized intersection groups, all crash types considered showed statistically significant reductions in crashes (e.g., total crashes, fatal and injury crashes, rear-end crashes, fatal and injury rear-end crashes, angle crashes, fatal and injury angle crashes, pedestrian-related crashes, and fatal and injury pedestrian-related crashes). A cross-sectional study was conducted with a larger number of PHBs (186) to identify relationships between roadway characteristics and crashes at PHBs, especially with respect to the distance to an adjacent traffic control signal. The distance to an adjacent traffic signal was found to be significant only at the $\alpha = 0.1$ level, and only for rear-end and fatal and injury rear-end crashes. *Conclusions:* This analysis represents the largest known study to date on the safety impacts of PHBs, along with a focus on how crossing and geometric characteristics affect crash patterns. The study showed the safety benefits of PHBs for both pedestrians and vehicles. *Practical Applications*: The findings from this study clearly support the installation of PHBs at midblock or intersection crossings, as well as at crossings on higher-speed roads.

1. Introduction

The pedestrian hybrid beacon (PHB, also known as a High-intensity Activated crossWalK [HAWK] beacon) is a traffic control device used at pedestrian crossings. It was added to the *Manual on Uniform Traffic Control Devices* (MUTCD) in 2009 (Federal Highway Administration, 2009). The PHB's vehicular display faces are typically located on mast arms over the major approaches to an intersection and in some locations on the roadside. See Fig. 1 for an example installation. The face of the PHB consists of two red indications above a single yellow indication that rests in a dark mode. When activated by a pedestrian, it first displays a few seconds of flashing yellow followed by a steady yellow change interval, and then displays a WALK indication to pedestrians and a steady red indication to drivers, which creates a gap for pedestrians to cross the major roadway. During the flashing pedestrian clearance interval, the PHB displays an alternating flashing red indication to allow drivers to proceed after stopping if the pedestrians have cleared the drivers' half of the roadway, thereby reducing vehicle delays. An official interpretation by the Federal Highway Administration (FHWA) (Federal Highway Administration, 2011) allows agencies to provide a brief all-red interval before the onset of the WALK interval and a brief interval between the onset of the steady DON'T WALK and the termination of the flashing red signal for drivers.

The focus of this Arizona Department of Transportation (ADOT) research effort (Fitzpatrick, Cynecki, Pratt, Park, & Beckley, 2019) was to investigate changes in crashes for different severity levels and crash types (e.g., angle or rear-end crashes) due to the PHB presence and in crashes involving pedestrians and bicycles.

* Corresponding author.
E-mail addresses: K-Fitzpatrick@tamu.edu (K. Fitzpatrick), e-park@tamu.edu (E. S. Park), mcynecki@lee-eng.com (M.J. Cynecki), m-pratt@tti.tamu.edu (M.P. Pratt), mbeckley@y2keng.com (M. Beckley).

https://doi.org/10.1016/j.jsr.2021.04.001
0022-4375/

Fig. 1. Example PHB Installation.

The PHB has shown considerable potential for improving driver yielding (Fitzpatrick, Turner, Brewer, Carlson, Lalani, Ullman, & Whitacre, 2006; Turner, Fitzpatrick, Brewer, & Park, 2006; Fitzpatrick, Avelar, Pratt, Brewer, Robertson, Lindheimer, & Miles, 2016; Fitzpatrick & Pratt, 2016) and pedestrian safety (Fitzpatrick & Park, 2010; Zegeer, C., Srinivasan, R., Lan, B., Carter, D., Smith, S., Sundstrom, C., Thirsk, N.J., Zegeer, J., Lyon, C., Ferguson, E., and Van Houten, R., 2017a). Previous studies have proven PHBs' effectiveness in reducing pedestrian crashes; however, questions on the effect of PHBs on rear-end crashes or severe crashes could not be fully addressed because of statistically insignificant results due to limited sample sizes. One of the main limitations was the relatively small sample sizes for crash data associated with PHBs. In addition to the focus on crash severity and type, a request was made to investigate the relationship between the PHB location and spacing to nearby signalized intersections.

The objective of this study is to investigate the safety performance of the PHB using different reference groups, with a focus on crash severity and type, especially with respect to pedestrians and rear-end crashes.

2. Previous studies

In a 2010 FHWA study, researchers conducted a before-and-after evaluation of the safety performance of the HAWK, now known as the PHB (Fitzpatrick & Park, 2010). Using an Empirical Bayes (EB) method, their evaluations compared the observed crash frequency after installation of the treatment (PHB) to the EB estimate of the expected crash frequency for the same after period without the treatment. To develop the datasets used in the evaluation, researchers counted the crashes occurring 3 years before and up to 3 years after installation of the PHB. The crash categories examined in the study included total, severe injury (fatal, incapacitating injury, non-incapacitating injury, and possible injury), and pedestrian crashes. After considering data for 21 treatment sites and 102 unsignalized intersections (reference group), the researchers found in their evaluation the following changes in crash statistics after installation of the PHBs:

- A 29% reduction in total crashes (statistically significant).
- A 15% reduction in severe crashes (not statistically significant).
- A 69% reduction in pedestrian crashes (statistically significant).

A 2017 National Cooperative Highway Research Program study (Fitzpatrick & Park, 2010) and paper (Zegeer, C., Srinivasan, R., Lan, B., Carter, D., Smith, S., Sundstrom, C., Thirsk, N.J., Zegeer, J., Lyon, C., Ferguson, E., and Van Houten, R., 2017a) investigated the safety effectiveness of the PHB and developed several crash modification factors (CMFs). For a PHB with advance yield or stop pavement markings and signs, the CMF was 0.432 for pedestrian crashes, 0.820 for total crashes, and 0.876 for rear-end or sideswipe crashes.

3. Methodology

For the ADOT safety evaluations, researchers used two methods:

- EB before-after study
- Cross-sectional observational study.

The EB before-after evaluation method estimated the changes in crashes after installation of PHBs by comparing the observed crash frequency after installation of the PHB to the EB estimate of the expected crash frequency for the same after period without the PHB (the counterfactual crash frequency). The counterfactual crash frequency was obtained by combining the observed before-period crash frequency, prediction from the safety performance function (SPF) based on reference sites (similar in site characteristics but without a PHB), and an adjustment factor that accounts for time trends and traffic volume changes between before and after periods. Details on the EB method are available elsewhere (Fitzpatrick, Cynecki, Pratt, Park, & Beckley, 2019; Hauer, 1997).

In the cross-sectional observational study, the crashes for a group of PHB sites were examined to investigate the effects of other site characteristic variables on crashes at the PHB sites. The cross-sectional observational analysis included a larger number of PHB sites because it included PHB sites that were installed for longer periods of time.

One potential bias that needs to be carefully considered in a before-after safety evaluation is the fact that PHB sites may be overrepresented with pedestrian crashes in the before period. One of the major considerations in the PHB selection and ranking criteria used by Tucson, Phoenix, and ADOT is the presence of pedestrian crashes during the prior 3 to 5 years. For example, ADOT Traffic Engineering Guidelines and Process 640 (entitled "Pedestrian Hybrid Beacons") (ADOT Traffic Engineering Guidelines, 2015) provides a point system for evaluating candidate locations, and states that "a minimum score of 35 points merits Pedestrian Hybrid Beacon consideration." Five points are added for each pedestrian or bicyclist crash occurring within the most recent 5 years of crash data.

Therefore, a potential regression to the mean (RTM) bias may exist at PHB locations for before-after evaluations based on the PHB selection criteria used by Arizona agencies. Because crashes during the before period are unnaturally high, crashes tend to regress toward the true long-term averages during the after period, and as a result, those sites might experience a reduction in crashes even without the PHB. Not accounting for this bias will result in overestimation of the safety effectiveness of PHBs. The EB before-after evaluation method properly accounts for the RTM bias that may exist by combining information from two sources: the observed crash frequency in the before period at PHB locations, and the predicted crash frequency (that is expected to be close to the true long-term mean crash frequency) based on reference sites with similar traffic and site characteristics as PHB locations.

The research team learned of another potential condition that can notably affect a before-after crash study. Tucson police stopped responding to non-injury/property damage only (PDO) crashes in December 2010. Tucson motorists can still submit PDO reports, but a majority of motorists do not do so. Two approaches are available to help account for this change and both were used in this study. First, the reference sites associated with Tucson PHBs were located in Tucson so that the change in reporting practices would affect both the treatment and reference sites. The second approach is to focus on the non-PDO crash (i.e., the fatal and injury crash) results over the all-crash results when sample size permits.

3.1. Site identification and geometric data

The research team obtained PHB locations and installation dates from the state and from several cities and identified 209 known PHBs in Arizona. Roadway characteristics data were obtained using aerial photographs for each of these sites, and Google Street View was used to determine the posted speed limit at the crossing. Table 1 lists the roadway variables that were considered in the safety analysis. In some cases, a variable had to be regrouped during the analysis; for example, the posted speed limit along the major street (i.e., the street with the PHB) was regrouped into 35 mph (56.3 kph) and below and 40 mph (64.4 kph) or higher.

3.2. Sites used in analyses

PHBs installed from 2011 to 2015 were considered for the EB before-after evaluation. Sites were removed from the before-after study if major roadway improvements occurred during that period (e.g., if a crossing site had been widened from two lanes to four lanes, or if a driveway was added at the site). For the before-after study, 52 PHB sites were available. At three of the 52 before-after sites, the agency added a short median island within the space previously provided for a two-way left-turn lane as part of the PHB improvement. The research team checked to see how the results would change if those three sites were omitted and found no material difference; therefore, the three sites were retained in the analysis.

The cross-sectional study included more PHB sites because crash data before the installation of the PHB was not needed; therefore, more of the older installations (prior to 2011) were considered. The cross-sectional observational study included 186 PHB sites and up to 10.75 years of crash data per intersection. Once again, locations with geometric changes were eliminated from the analysis.

3.3. Reference groups

Crash evaluations are beneficial when a reference group of similar sites without treatment is identified. Three potential reference groups were identified for before-after evaluation. The research team selected intersections near the PHB on the major roadway with the goal of finding intersections with a similar roadway cross section (e.g., number of lanes or median type), speed limit, ADT, and number of legs where pedestrians can be expected to cross the street. Number of legs could not be matched for those PHBs installed midblock. In general, one signalized and two unsignalized intersections were identified for use as comparisons for each PHB site included in the before-after evaluation.

3.4. Vehicle counts

Several sources were used to obtain vehicle counts, including traffic counts (or historical traffic flow maps) available on the Web and historical counts from ADOT, the Pima Association of Governments (PAG), the Maricopa Association of Governments, and various cities and county agencies. Vehicle counts from existing sources were identified for most of the major streets at PHB crossings. For most sites, traffic counts were available for about every third or fourth year. When a count was not available for a given year, the count was estimated to be equal to the count from the most recent year with a known count. This method was used to estimate traffic volumes for the major streets at the study sites. In almost all cases, vehicle count data were not available for the side streets because the side streets were low-volume local residential streets that are typically not counted.

3.5. Pedestrian counts

The research team contacted the appropriate roadway owners for any available historical pedestrian count data, which included City of Phoenix data, PAG pedestrian and bicyclist data, historical count data from a prior FHWA study (Fitzpatrick & Park, 2010), and miscellaneous pedestrian and bicyclist data sources. The available historical data were used to establish typical pedestrian volumes by general level of pedestrian activity, as shown in Table 2. The pedestrian volume values are primarily based on the data from the 2010 FHWA study (Fitzpatrick & Park, 2010), with the other counts being used to judge the reasonableness of the values. *Highway Safety Manual* (HSM) (American Association of State Highway and Transportation Officials (AASHTO), 2010) data are included in the table as a comparison.

For sites that did not have any historical pedestrian or bicyclist data, the Arizona-based research team members provided their judgment on the general level of pedestrian activity at each site using their local knowledge and a review of the land development near the site. The general level of pedestrian activity was then translated to pedestrian volume based upon the traffic control pre-

Table 1
Roadway Variables Considered in Safety Analyses.

Variable	Description
C_Lanes	Cross: total number of lanes on the cross street (side street) for intersection PHBs
Legs	Number of legs at the intersection (2, 3, or 4), 2-legs are midblock crossings
M_Bike_01	Major: is a bike lane present? (1 = bike lane on one or both sides, 0 = none)
M_Lanes	Major: number of through lanes
M_LTL	Major: is a left-turn lane present on the major street? (0 = neither approach has a left-turn lane, 1 = at least one of the approaches has a left-turn lane)
M_LTL_A	Major: number of approaches with an exclusive left-turn lane (0, 1, or 2)
M_MT	Major: median type (raised, two-way left-turn lane [TWLTL], none, flush)
M_MT_R	Major: median type (raised = raised [0], all others, e.g., flush, TWLTL, none = not raised [1])
M_PK_01	Major: is a parking lane present? (1 = parking lane on one or both sides, 0 = none)
Ped or PB_Vol_MC	Daily number of pedestrians at the intersection, sum of the pedestrian volume on the major and on the cross street
PSL	Major: posted speed limit (mph)
PSL_group	Major: posted speed limit for the main street grouped into either 35 mph (56.3 kph) and below or 40 mph (64.4 kph) and higher
Sig_Dist	Major: distance between the PHB and the nearest traffic signal in feet
Veh	Major: daily number of vehicles on the major street, also called average daily traffic (ADT)

Table 2
Pedestrian volume by general level of pedestrian activity and traffic control.

General Level of Pedestrian Activity[a]	PHB[b] Ped Major 24 hr	PHB Ped Cross24 hr	PHB Ped All 24 hr	Unsig[c] Ped Major 24 hr	Unsig Ped Cross24 hr	Unsig Ped All 24 hr	Sig[d] Ped Major 24 hr	Sig Ped Cross 24 hr	Sig Ped All 24 hr	HSM[e] Sig 3-Leg	HSM Sig 4-Leg
High	950	1,180	2,130	320	290	610	820	700	1,520	1,700	3,200
Medium High	490	480	970	190	180	370	410	530	940	750	1,500
Medium	170	220	390	90	90	180	210	290	500	400	700
Medium Low	90	40	130	40	40	80	110	170	280	120	240
Low	40	20	60	10	20	30	60	60	120	20	50

Notes:
[a] The team assumed the general level of high pedestrian activity to be the 90th percentile value (rounded to the nearest 10) for the group of sites. The medium high was the 75th percentile, the medium was the 50th percentile, the medium low was the 25th percentile, and the low was the 10th percentile value (rounded to the nearest 10). Other assumptions include that the PHB is controlling the vehicles on the major street and that the pedestrian count for all is the sum of the pedestrians crossing the major legs and the pedestrians crossing [the] cross-street legs (if any).
[b] PHB values based on 52 PHB (HAWK) intersections.
[c] Unsig values based on 98 unsignalized intersections.
[d] Sig values based on 33 signalized intersections.
[e] HSM values are from the *Highway Safety Manual* Table 12-15, pp. 12–37.

sent using the values shown in Table 2. For example, if the general level of pedestrian activity at a PHB site was judged to be medium, then 170 pedestrians (daily) were assumed to cross the major street. Although this approach has limitations, the resources available, along with the large number of sites, required a different approach than collecting actual pedestrian crossing volumes at the sites.

3.6. Crash data

Crash data were supplied by ADOT for the 10.75-year period of January 1, 2007, to September 30, 2017. The incident records included latitude and longitude coordinate variables for the crash, which were used to identify crashes relevant to the study sites. About 2.8% of the crashes had to be discarded because their coordinate variables were not populated.

A database was developed with the coordinates for every PHB study and reference site, and crashes were extracted if they occurred within 250 ft (76.2 m) of the center of the intersection (or midblock crossing site). A total of 17,400 crashes were identified at the study sites, 5,383 of which were at PHB sites (the amount includes periods both before and after PHB installation). The following crash types were considered in the safety analysis:

- Total crashes
- Fatal and injury (FI) crashes, which consist of the following severity levels: fatal, incapacitating injury, non-incapacitating injury, and possible injury
- Rear-end crashes
- Angle crashes
- Pedestrian-related crashes
- FI rear-end crashes
- FI angle crashes
- FI pedestrian-related crashes

The collision manner variable was used to identify rear-end crashes and angle crashes (defined as front to side crashes, excluding left turns). Pedestrian-related crashes were identified by merging the incident records with their corresponding unit records using the incident ID variable that was common to both files. The units were identified as vehicles, bicycles, or pedestrians. Any crash involving one or more pedestrian units was coded as a pedestrian-related crash. Note that the crash types identified in the preceding list are not mutually exclusive; for example, a rear-end crash involving two vehicles and one pedestrian would be classified as both a rear-end crash and a pedestrian-related crash.

4. Results

4.1. EB before-after evaluation

The before-after evaluation included 52 intersections as treatment sites for which the PHB were installed during the study period. Reference Group 1 consisted of 101 unsignalized intersections, Reference Group 2 consisted of 56 signalized intersections, and Reference Group 3 consisted of 157 unsignalized or signalized intersections. The reference groups represent sites similar to the treatment sites but without the PHB.

The before period at each site was defined as January 1, 2007, to 2 months prior to the PHB installation date. Crashes occurring in the 2 months prior to the installation date were removed because they were assumed to have occurred during construction. Crashes occurring in the 2 months following the PHB installation were also removed because they were assumed to occur during the acclimation period while drivers were becoming familiar with the treatment. The after period consisted of 2 months following PHB installation until September 30, 2017.

The number of months in the after period for the 52 PHBs varied depending on when installation occurred. The average number of months in the before period was 79 months, with a range of 50–107 months. For the after period, the average number of months was 52 months, with a range of 23–81 months. Reference group sites were assigned the same time period in the before and after periods as their corresponding PHB site.

Table 3 contains the total number of crashes, the annual crashes adjusted by period duration, and the percentage for each type of crash by site type for the before and after study periods. Table 4 contains the summary of site characteristic variables for PHB sites, unsignalized intersections, and signalized intersections used in before-after evaluations.

Because the success of an EB evaluation largely depends on reliable estimation of SPFs, it is important to identify a reference group that is similar enough to the treatment group with respect to roadway characteristics, weather, and traffic volumes. The following three reference groups were employed to assess the robustness of results and conclusions from the EB before-after analysis:

- Reference Group 1: Unsignalized Intersections
- Reference Group 2: Signalized Intersections

Table 3
Crashes during each study period (before vs. after).

Crash Type	PHB Bef Num[a]	PHB Bef ACC[b]	PHB Bef %[c]	PHB Aft Num[a]	PHB Aft ACC[b]	PHB Aft %[c]	Unsig Bef Num[a]	Unsig Bef ACC[b]	Unsig Bef %[c]	Unsig Aft Num[a]	Unsig Aft ACC[b]	Unsig Aft %[c]	Sig Bef Num[a]	Sig Bef ACC[b]	Sig Bef %[c]	Sig Aft Num[a]	Sig Aft ACC[b]	Sig Aft %[c]
Total crashes	1,064	3.14	100%	600	2.72	100%	1,446	2.20	100%	940	2.24	100%	5,594	15.56	100%	3,627	14.94	100%
FI crashes	408	1.20	38.4%	230	1.04	38.3%	529	0.81	36.6%	337	0.80	35.9%	2,063	5.74	36.9%	1,421	5.86	39.2%
Rear-End crashes	468	1.38	44.0%	206	0.94	34.3%	561	0.89	38.8%	303	0.72	32.2%	2,230	6.20	39.9%	1,182	4.87	32.6%
Angle crashes	199	0.59	18.7%	112	0.51	18.7%	285	0.43	19.7%	184	0.44	19.6%	1,114	3.10	19.9%	718	2.96	19.8%
Ped-related crashes	70	0.21	6.6%	19	0.09	3.2%	48	0.07	3.3%	19	0.05	2.0%	114	0.32	2.0%	101	0.42	2.8%
FI_Rear-End crashes	175	0.52	16.5%	78	0.35	13.0%	179	0.27	12.4%	118	0.28	12.6%	701	1.95	12.5%	408	1.68	11.3%
FI_Angle crashes	66	0.19	6.2%	42	0.19	7.0%	111	0.17	7.7%	66	0.16	7.0%	464	1.29	8.3%	314	1.29	8.7%
FI_Ped related crashes	62	0.18	5.8%	19	0.09	3.2%	44	0.07	3.0%	18	0.04	1.9%	99	0.28	1.8%	96	0.40	2.7%

Notes:
Bef = before and Aft = after
[a] Number of crashes.
[b] Crashes adjusted by period duration, i.e., adjusted crash count (ACC) = crash count/number of days in each period*365, crashes/year.
[c] Percent crashes = number of crashes of each type/number of total crashes.
# of sites: 52 for PHB, 101 for unsignalized, and 56 for signalized.
# of days in each period (summed over sites): 123,677 for PHB before; 80,423 for PHB after; 239,543 for unsignalized before; 152,957 for unsignalized after; 131,210 for signalized before; and 88,590 for signalized after.

- Reference Group 3: Unsignalized Intersections + Signalized Intersections

The first step in the EB before-after method was to develop and calibrate SPFs using data from a reference group. Development of the SPFs involved determining which predictor variables should be used in the model, how the variables should be grouped, and what model should be used. The vehicle volume values (i.e., ADT) are often the key variables in developing SPFs for intersections. In addition, pedestrian volumes are likely to play an important role in pedestrian crashes. To account for additional intersection-to-intersection variability (other than that caused by the differences in traffic volumes and pedestrian volumes), number of legs (Legs), number of through lanes (M_Lanes), existence of raised median (M_MT_R), existence of left-turn lanes (M_LTL), existence of on-street parking (M_PK_01 with 0 = no on-street parking, 1 = on-street parking exists), existence of a bike lane (M_Bike_01, with 0 = no bike lanes, 1 = bike lanes exist), and total number of entering lanes (C_lanes) were also considered in the SPF predictions.

The negative binomial regression models with indicator variables for time period to control for general trends from before to after periods along with the aforementioned variables as independent variables were employed to develop SPFs based on the reference group. The estimated coefficients for SPFs along with the dispersion parameter and the goodness of fit measure (Pearson chi-square statistic divided by degrees of freedom) for total, FI, rear-end, angle, pedestrian-related, FI rear-end, FI angle, and FI pedestrian-related crashes based on the three reference groups are presented in Table 5. Table 6 shows the results of an EB before-after evaluation for multiple crash types considered along with three different reference groups employed.

In general, the results support positive safety effects of PHBs for crash types considered regardless of reference group. Based on Reference Group 1 (unsignalized intersections), the effects of PHBs are statistically significant for total, fatal and injury, rear-end, and pedestrian-related crashes but not significant for angle, FI angle, and FI pedestrian-related crashes, although the effects are still positive. Because the estimated SPF for FI pedestrian-related crashes is subject to larger uncertainty due to a small sample size, researchers also performed a sensitivity analysis by estimating the expected number of FI pedestrian-related crashes during the before period in EB implementation using predictions from pedestrian-related crash SPF (developed based on a larger sample size and consequently subject to smaller uncertainty) after multiplying the ratio of FI pedestrian-related crashes and pedestrian-related crashes at unsignalized intersections. The estimated crash reduction for FI pedestrian-related crashes based on pedestrian-related crash SPF was statistically significant, which is deemed to be a consequence of more precise SPF estimation based on a larger sample size.

For Reference Group 2 and Reference Group 3, all crash types evaluated were statistically significant.

4.2. Cross-sectional observational safety evaluation for PHB installations

ADOT was also interested in assessing the effects of site characteristic variables on crashes at PHB sites, especially the distance between an adjacent traffic control signal and a PHB. There were 186 PHB sites (with PHBs installed between 2000 and 2016) available for this analysis. Table 7 contains the summary of site characteristics variables for PHB sites used in this cross-sectional observational evaluation. The geometric site characteristics for these PHB crossings were reviewed to verify that their geometry had not changed during the study period.

Table 4
Descriptive statistics for PHB sites used in EB before-after evaluations.

Variable[a]	PHB (52 sites)			Unsignalized intersections (101 sites)			Signalized intersections (56 sites)		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
Legs	2	4	3.3	2	4	3.6	3	4	3.9
M_Lanes	2	7	5.0	2	9	5.3	2	10	5.5
M_LTL	0	2	0.9	0	2	1.5	0	2	1.9
M_PK_01	0	1	0.1	0	1	0.1	0	1	0.1
M_Bike_01	0	1	0.6	0	1	0.6	0	1	0.6
C_Lanes	0	3	1.3	0	6	1.6	2	12	6.0
Veh (ADT)	5,400	47,627	23,959	4,937	47,627	24,377	5,400	48,512	24,421
PB_Vol_MC (Ped)	10	1,670	297	10	480	99	30	1,520	308
M_MT_R, Value (# of sites)	Not Raised (33), Raised (19)			Not Raised (69), Raised (32)			Not Raised (35), Raised (21)		

Note:
[a] See Table 1 for description of roadway variables.

Table 5
Estimates of coefficients[a] for SPFs developed.

RG[b]	Parameter	Total	FI	Rear – End	Angle	FI Rear – End	FI Angle	Ped_rel	FI Ped_rel
1	β_{0B}	−11.9403	−11.5226	−15.7498	−12.378	−15.7896	−12.1244	−15.4578	−15.9687
1	β_{0A}	−12.0754	−11.6312	−16.0486	−12.478	−15.8129	−12.2466	−16.0408	−16.5126
1	β_{Legs}	0.1009	0.0596	0[c]	0.2010	0[c]	0.2454	0.1656	0.2603
1	β_{M_Lanes}	0.1478	0.1821	0.0561	0.2362	0.1299	0.2097	0.0988	0.0873
1	$\beta_{M\_MT\_R}$	0.3926	0.4216	0.3285	0.4808	0.3397	0.6767	0[c]	0[c]
1	β_{M_LTL}	−0.1509	−0.0977	−0.2209	−0.1071	−0.1902	−0.0817	0[c]	0[c]
1	$\beta_{M\_PK\_01}$	0.3809	0.2857	0.6189	0.3198	0.4417	0.2968	0[c]	0[c]
1	$\beta_{M\_Bike\_01}$	−0.7960	−0.5834	−0.8038	−0.7240	−0.6199	−0.6689	−1.2943	−1.3360
1	β_{C_Lanes}	0.2059	0.1528	0.2600	0.2874	0.1915	0.1656	0.0787	0.0877
1	$\beta^{d}_{n(Veh)}$	0.5514	0.3470	0.9454	0.2988	0.7101	0.1673	0.3713	0.4483
1	$\beta^{d}_{Ln(Ped)}$	0.0324	0.1433	−0.0211	0.0778	0.1703	0.0960	0.5696	0.4292
1	Dispersion	0.4573	0.4431	0.5481	0.7814	0.4770	0.8854	0.7133	0.7122
1	Person Chi-Square/DF [e]	1.2012	1.2260	1.0526	1.1970	1.0774	1.1832	1.0007	0.9980
2	β_{0B}	−11.421	−13.334	−13.225	−12.8435	−15.1351	−15.522	−21.434	−22.1790
2	β_{0A}	−11.658	−13.508	−13.578	−13.1207	−15.4234	−15.761	−21.509	−22.1791
2	β_{Legs}	0.9119	0.8809	0.8945	1.4091	0.6067	1.8794	1.9768	1.7445
2	$\beta_{\_Lanes}$	0.1157	0.1356	0.0534	0.1975	0.0295	0.1825	0.0926	0.0939
2	$\beta_{M\_MT\_R}$	0.4685	0.4399	0.3674	0.5880	0.4159	0.4601	0.3341	0.4363
2	β_{M_LTL}	−0.4095	−0.3507	−0.3751	−0.4328	−0.2963	−0.3904	0[c]	0[c]
2	$\beta_{M\_PK\_01}$	0.1590	0.0769	0.1518	0.1149	0.0955	0.0472	0[c]	0[c]
2	$\beta_{M\_Bike\_01}$	−0.2717	−0.1457	−0.2635	−0.1358	−0.1241	−0.1009	−0.3783	−0.3167
2	β_{C_Lanes}	0.1980	0.1618	0.2428	0.1145	0.2178	0.0629	0.1761	0.1693
2	$\beta^{d}_{Ln(Veh)}$	0.2531	0.3307	0.3835	0[c]	0.5560	0.0334	0.2716	0.3955
2	$\beta^{d}_{Ln(Ped)}$	0.1480	0.1902	0.0862	0.2170	0.1271	0.2227	0.4077	0.4455
2	Dispersion	0.3306	0.3116	0.3758	0.3531	0.3188	0.3388	0.1936	0.3014
2	Person Chi-Square/DF [e]	1.1711	1.0862	1.2010	1.0820	1.2562	1.0722	1.0562	1.0050
3	β_{0B}	−12.208	−12.450	−14.6901	−13.091	−15.0960	−13.732	−15.558	−16.7892
3	β_{0A}	−12.357	−12.580	−14.9795	−13.232	−15.2513	−13.872	−15.861	−17.0110
3	β_{Legs}	0.2848	0.2699	0.1861	0.5763	0.1246	0.7313	0.3507	0.4063
3	β_{M_Lanes}	0.0823	0.1008	0.0078	0.1455	0.0240	0.0862	0.0658	0.0662
3	$\beta_{M\_MT\_R}$	0.4186	0.3653	0.3841	0.5271	0.3529	0.4530	0.1466	0.1841
3	β_{M_LTL}	−0.2440	−0.2138	−0.2805	−0.2236	−0.2574	−0.2163	0[c]	0[c]
3	$\beta_{M\_PK\_01}$	0.3835	0.2202	0.5423	0.2731	0.3434	0.1351	0[c]	0[c]
3	$\beta_{M\_Bike\_01}$	−0.5514	−0.3950	−0.5235	−0.3937	−0.3570	−0.3425	−0.7398	−0.7446
3	β_{C_Lanes}	0.3435	0.3194	0.3776	0.2942	0.3379	0.2836	0.2145	0.2093
3	$\beta^{d}_{Ln(Veh)}$	0.5035	0.3901	0.7516	0.2563	0.6646	0.1919	0.2897	0.3813
3	$\beta^{d}_{Ln(Ped)}$	0.1079	0.1826	0.0496	0.1739	0.1225	0.2165	0.5350	0.5280
3	Dispersion	0.5008	0.4867	0.5614	0.6844	0.4758	0.7036	0.3381	0.4583
3	Person Chi-Square/DF [e]	1.2621	1.2147	1.0683	1.2873	1.0616	1.2942	0.9709	0.9378

Notes:
[a] negative coefficient indicates that the number of crashes decreases with an increase in the value of the variable, while a positive coefficient indicates that the number of crashes increases with an increase in the value of the variable. For example, the coefficient for number of legs (unsig reference group) is positive for total crashes (i.e., 0.1009), which indicates that more crashes are associated with 4-leg intersections than 3-leg intersections.
[b] RG = reference group, where 1 = unsignalized intersections, 2 = signalized intersections, and 3 = both unsignalized and signalized intersections.
[c] The coefficient "0" denotes that the corresponding variable was excluded from the model.
[d] Ln = natural log.
[e] DF = degree of freedom. A value of Pearson chi-square/DF close to 1 indicates a good model fit.

Crash prediction models based on crash data from PHB sites after installation of PHBs were developed using generalized linear models. The goal of this analysis was to identify relationships between roadway characteristics and crashes by crash type. Variables were removed from the model if the variable was not significant and the sign of the estimated coefficient was counterintuitive. In some cases, variables that were not statistically significant were retained in the models (as long as the signs of

Table 6
Results of EB Before-After Safety Evaluations.

RG[a]	Crash Type	Observed	EB ($\hat{\pi}$)	$\hat{\theta}$ (SE)	95% CI for θ	90% CI for θ	%CR[b]
1	Total	600	679.1	0.883 (0.046)	(0.792, 0.973)	(0.807, 0.958)	11.7**
1	Fatal & Injury	230	283.8	0.808 (0.067)	(0.678, 0.939)	(0.699, 0.918)	19.2**
1	Angle	112	128.1	0.870 (0.103)	(0.668, 1.071)	(0.701, 1.039)	13.0
1	FI Angle	42	41.9	0.991 (0.185)	(0.628, 1.353)	(0.687, 1.294)	0.9
1	Rear-End	206	234.2	0.878 (0.074)	(0.733, 1.022)	(0.756, 0.999)	12.2*
1	FI Rear-End	78	121.5	0.639 (0.084)	(0.474, 0.805)	(0.501, 0.778)	36.1**
1	Ped-related	19	33.1	0.567 (0.143)	(0.288, 0.847)	(0.333, 0.801)	43.3**
1	FI Ped-related	19	24.9	0.755 (0.191)	(0.381, 1.128)	(0.442, 1.067)	24.5
1	FI Ped-related[c]	19	29.0	0.648 (0.164)	(0.326, 0.969)	(0.379, 0.916)	35.2**
2	Total	600	724.0	0.828 (0.043)	(0.743, 0.913)	(0.757, 0.899)	17.2**
2	Fatal & Injury	230	334.7	0.685 (0.056)	(0.575, 0.796)	(0.593, 0.777)	31.5**
2	Angle	112	157.8	0.706 (0.081)	(0.547, 0.866)	(0.573, 0.840)	29.4**
2	FI Angle	42	75.5	0.552 (0.099)	(0.359, 0.745)	(0.390, 0.713)	44.8**
2	Rear-End	206	264.1	0.778 (0.065)	(0.651, 0.906)	(0.671, 0.885)	22.2**
2	FI Rear-End	78	115.6	0.672 (0.087)	(0.501, 0.843)	(0.529, 0.815)	32.8**
2	Ped-related	19	29.9	0.630 (0.153)	(0.329, 0.930)	(0.378, 0.881)	37.0**
2	FI Ped-related	19	31.8	0.591 (0.147)	(0.303, 0.879)	(0.350, 0.832)	40.9**
3	Total	600	732.2	0.818 (0.043)	(0.734, 0.903)	(0.748, 0.889)	18.2**
3	Fatal & Injury	230	306.6	0.748 (0.062)	(0.626, 0.870)	(0.646, 0.850)	25.2**
3	Angle	112	143.9	0.774 (0.092)	(0.595, 0.954)	(0.624, 0.925)	22.6**
3	FI Angle	42	55.0	0.755 (0.141)	(0.479, 1.031)	(0.524, 0.986)	24.5*
3	Rear-End	206	258.5	0.795 (0.067)	(0.664, 0.927)	(0.685, 0.905)	20.5**
3	FI Rear-End	78	108.7	0.714 (0.094)	(0.529, 0.899)	(0.559, 0.869)	28.6**
3	Ped-related	19	34.7	0.543 (0.133)	(0.282, 0.804)	(0.324, 0.761)	45.7**
3	FI Ped-related	19	34.2	0.550 (0.137)	(0.281, 0.819)	(0.325, 0.775)	45.0**

Notes:
[a] Abbreviations used in column headings:

- RG = reference group, where 1 = unsignalized intersections, 2 = signalized intersections, and 3 = both unsignalized and signalized intersections.
- Observed = observed crashes in the after period.
- EB ($\hat{\pi}$) = EB estimate representing the predicted number of crashes in the after period had PHBs not been installed.
- $\hat{\theta}$ = estimated index of effectiveness. Note that the index of effectiveness is equivalent to the CMF.
- SE = standard error.
- CI = confidence interval.
- %CR = percent crash reduction = $100(1-\hat{\theta})$.

[b] Statistical level indications:
*Statistically significant results with 90% confidence level (also known as 10% significant level).
**Statistically significant results with 95% confidence level (also known as 5% significant level).
[c] Indicates the results from the sensitivity analysis using the prediction based on Ped-related crash SPF for prediction after adjusted by the ratio of FI Ped-related crashes and Ped-related crashes at unsignalized intersections.

coefficients were not counterintuitive) to examine trends. Table 8 contains the estimated regression coefficients, dispersion parameter, and the Pearson chi-square statistic divided by degrees of freedom for each crash type, along with the *p*-value for the variable.

The negative binomial regression models, with variables in Table 8 as independent variables, were employed to develop prediction equations for crashes at PHB sites. The basic prediction equation being considered for the different crash type was:

$$\begin{aligned}\mu = \exp\big(&\beta_0 + \beta_{\mathrm{Legs}} \times \mathrm{Legs} + \beta_{\mathrm{M_Lanes}} \times \mathrm{M_Lanes} + \beta_{\mathrm{M_MT_R}} \\ &\times \boldsymbol{I}[\mathrm{M_MT_R} = \mathit{NotRaised}] \\ &+ \beta_{\mathrm{PSL_group}} \times \boldsymbol{I}\left[\mathrm{PSL_group} = 35 \text{ or less}\right] + \beta_{\mathrm{M_LTL}} \times \mathrm{M_LTL} + \beta_{\mathrm{M_PK_01}} \\ &\times \mathrm{M_PK_01} + \beta_{\mathrm{M_Bike_01}} \times \mathrm{M_Bike_01} + \beta_{\mathrm{C_Lanes}} \times \mathrm{C_Lanes} + \beta_{\mathrm{LnVeh}} \\ &\times \mathit{Ln}(\mathrm{Veh}) + \beta_{\mathrm{LnPed}} \times \mathit{Ln}(\mathrm{Ped}) + \beta_{\mathrm{Sig_Dist}} \times \mathrm{Sig_Dist}\big)\end{aligned}$$

where:

μ = predicted daily crashes
I = indicator function taking a value 1 if the condition in [ ] is satisfied, and 0 otherwise
Ln = natural log
C_Lanes = number of through lanes on the cross street
Legs = number of legs at the intersection
M_Bike_01 = 1 if a bicycle lane is present on either side of major street and 0 otherwise
M_Lanes = number of through lanes on the major street
M_LTL = number of approaches on the major street with a left-turn lane
M_MT_R = median type, raised (0) or not raised (1)
M_PK_01 = 1 if a parking lane is present on either side of major street and 0 otherwise
Ped = major and cross-street daily pedestrian volume
PSL_group = posted speed limit group, 35 mph (56.3 kph) and less or 40 mph (64.4 kph) and more
Sig_Dist = distance between the PHB and the nearest signal in feet; if Sig_Dist > 1500, then Sig_Dist = 1500
Veh = major road ADT

For total crashes, the roadway geometry variables that have significant effects on crashes for PHBs include the number of lanes on the major roadway, median treatment, bike lane presence, and number of lanes on the cross street. These relationships are as expected, with more lanes on either the major or cross street being associated with more crashes and the presence of a raised median or pedestrian refuge island being associated with fewer crashes. The presence of a bike lane at the PHB being associated with fewer total crashes is a positive finding.

Table 7
Descriptive statistics for PHB sites used in observational analysis.

Variable	PHB (186 sites)		
	Minimum	Maximum	Average
Legs	2	4	3.4
M_Lanes	2	9	4.5
M_LTL	0	2	0.8
M_PK_01	0	1	0.1
M_Bike_01	0	1	0.6
C_Lanes	0	6	1.4
Veh (ADT)	1,385	50,510	23,500
PB_Vol_MC (Ped)	40	2,130	475
Sig_Dist	277	13,249[a]	1,548
M_MT_R	Value (# of sites)	Not Raised (119), Raised (67)	
PSL_group	Value (# of sites)	35 or less (97), 40 or more (89)	

Note:
[a]If Sig_Dist was greater than 1,500 ft (457.2 m), the value was set to 1,500 ft (457.2 m). At a certain distance, a traffic control signal will probably not affect the operations or safety of a neighboring intersection. This distance was assumed to be 1,500 ft (457.2 m) based on engineering judgment.

Table 8
Estimated regression coefficients of SPFs developed for crashes at PHB sites.

Parameter	Total	FI	Rear-End	FI Rear-End	Angle	FI Angle	Ped_rel	FI Ped_rel
β_0	−13.5812 (<0.0001)	−15.948 (<0.0001)	−18.225 (<0.0001)	−23.4423 (<0.0001)	−12.940 (<0.0001)	−15.395 (<0.0001)	−21.029 (<0.0001)	−20.9389 (<0.0001)
β_{Legs}	0.0849 (0.4947)	0.0801 (0.4931)	0.1570 (0.1331)	0.0443 (0.6682)	0.2366 (0.1910)	0.1491 (0.4617)	0.2282 (0.3358)	0.3231 (0.2142)
β_{M_Lanes}	0.1234 (0.0557)	0.0496 (0.4600)	0.1365 (0.0657)	0	0.1541 (0.0904)	0	0.3856 (0.0073)	0.3787 (0.0159)
$\beta_{M\_MT\_R}$	0.2730 (0.0621)	0.2221 (0.1055)	0.3316 (0.0502)	0.3713 (0.0112)	0.2664 (0.2479)	0.1610 (0.4254)	0.9286 (0.0028)	0.8014 (0.0419)
β_{PSL_group}	−0.1407 (0.2427)	−0.1512 (0.2070)	−0.2826 (0.0328)	−0.2668 (0.0530)	0	0	0	0
β_{M_LTL}	−0.0376 (0.6418)	0	−0.1061 (0.2522)	0	−0.1076 (0.3659)	0	0	−0.1091 (0.5913)
$\beta_{M\_PK\_01}$	0	0	0.1246 (0.6083)	0.2633 (0.3430)	0.1065 (0.7460)	0	0	0
$\beta_{M\_Bike\_01}$	−0.2107 (0.0701)	−0.1073 (0.3747)	−0.1163 (0.3547)	−0.0764 (0.5782)	−0.3113 (0.0648)	−0.2440 (0.2165)	0.3677 (0.1642)	0.2737 (0.3140)
β_{C_Lanes}	0.1802 (0.0466)	0.2052 (0.0166)	0	0	0.3706 (0.0026)	0.4765 (0.0004)	0.1342 (0.4439)	0.0966 (0.6033)
β_{LnVeh}	0.6733 (<0.0001)	0.8434 (<0.0001)	1.0465 (<0.0001)	1.5901 (<0.0001)	0.3715 (0.0369)	0.5557 (0.0009)	0.5968 (0.0315)	0.5939 (0.0464)
β_{LnPed}	0.1131 (0.0642)	0.1140 (0.0750)	0.1799 (0.0074)	0.1172 (0.1121)	0.0714 (0.4152)	0.1678 (0.1146)	0.4706 (0.0008)	0.4478 (0.0020)
β_{Sig_Dist}	0	0	−0.0004 (0.0595)	−0.0004 (0.0608)	0	0	0	0
Dispersion	0.4284	0.3446	0.3867	0.2551	0.6336	0.6552	0.2817	0.3211
Person Chi-Square/DF	1.1578	1.0822	1.4349	1.0739	1.2587	1.2004	0.9252	1.9216

Notes:
1. The coefficient "0" denotes that the corresponding variable was excluded from the model.
2. P-values are provided in parentheses.
3. Cells are highlight in light gray when the p-value is between 0.05 and 0.1.
4. Cells are highlighted in dark gray with white text when the p-value is less than 0.05.
5. DF stands for degrees of freedom.

Both pedestrian-related crashes and FI pedestrian-related crashes had stronger findings with respect to the number of lanes on the major street and the presence of a raised median. Both of these variables were significant at the 0.05 level for pedestrian-related crashes compared to them only being significant at the 0.1 level for total crashes. Several studies have documented the benefit of a raised median/refuge island for pedestrians (Zegeer, 2017b), and the findings from this ADOT study also identified benefits to pedestrians for a raised median/refuge island.

The variable that was always significant for each crash type was vehicle volume, which was expected. Pedestrian volume was significant for most of the crash types. Angle crashes are the only crash type where having the pedestrian volume in the model was of questionable value. Posted speed limits were grouped into 35 mph (56.3 kph) and below or 40 mph (64.4 kph) and higher. That variable was only statistically significant for rear-end crashes. More rear-end crashes are predicted for roads posted at 40 mph (64.4 kph) and higher than for roads posted at 35 mph (56.3 kph) and below.

The distance to the nearest traffic signal variable only remained significant in the rear-end and FI rear-end crash type models, where it was significant at the 0.1 level. More of these crash types are associated with shorter distances between a traffic control signal and a PHB; however, the impact of the distance to traffic signal variable on predicting rear-end or FI rear-end crashes is less influential than or similar to the impact of higher (compared to lower) speeds or the impact of not having a raised median (compared to having a raised median). No significant difference existed in any crash type for midblock PHBs compared to PHBs installed at intersections.

5. Summary

A total of 343 sites were included in the safety studies, consisting of 186 PHBs, 56 signalized intersections, and 101 unsignalized intersections. PHB installation dates were obtained from the various government agencies, and 52 PHBs installed between 2011 and 2015 were identified for use in the EB before-after analysis. Reference groups consisting of signalized and unsignalized intersections were chosen from intersections in close proximity to the 52 before-after PHB sites and were used in the EB before-after analysis.

Along with this ADOT study, previous PHB research studies have found safety benefits in installing a PHB. When considering the reference group consisting of unsignalized intersections, crash reductions were found for the following crash types: total crashes, FI crashes, FI rear-end crashes, and pedestrian-related crashes. Crash reductions were also found for all other crash types studied when using the unsignalized intersection reference group; however, the reductions were not statistically significant.

The safety performance of PHBs can be compared to unsignalized intersections, signalized intersections, or both unsignalized and signalized intersections. In most cases, a PHB is installed at a pedestrian crossing that previously was unsignalized; however, in a few cases, the PHB replaced a traffic control signal. The level of pedestrian activity for a PHB intersection is more similar to signalized rather than unsignalized intersections; therefore, comparing PHBs to signalized intersections may be more valid.

Each reference group has potential limitations; therefore, the research team considered three different reference groups: unsignalized intersections, signalized intersections, and both unsignalized and signalized intersections combined. For the signalized and combined unsignalized and signalized intersection groups, all crash types considered showed statistically significant reductions in crashes (e.g., total crashes, FI crashes, rear-end crashes, FI rear-end crashes, angle crashes, FI angle crashes, pedestrian-related crashes, and FI pedestrian-related crashes).

For the 52 PHB sites included in the before-after study regardless of the reference group being considered, a reduction in pedestrian-related crashes was observed, as expected. Reductions were also observed for FI crashes and for rear-end crashes, two crash types where concern existed that the installation of the PHB may increase those types of crashes.

A cross-sectional study was conducted with a larger number (186) of PHBs to identify relationships between roadway characteristics and crashes at PHB sites, especially with respect to the distance between a traffic control signal and a PHB. The cross-sectional study was able to include more PHB sites because crash data before the installation of the PHB were not needed; therefore, more of the older installations (prior to 2011) could be considered.

For total crashes, the roadway variables with relationship to crashes at PHBs include the number of lanes on the major roadway, median treatment, bike lane presence, and number of lanes on the cross street. These relationships are as expected, with more lanes on either the major or cross street being associated with more crashes and with the presence of a raised median or pedestrian refuge island being associated with fewer crashes. The presence of a bike lane at the PHB being associated with fewer total crashes is a positive finding. Several studies, including this ADOT study, have documented the benefit of a raised median/refuge island for pedestrians. The distance to the signal variable only remained significant in the rear-end and FI rear-end crash type models, where it was significant at the 0.1 level (90% confidence). More rear-end crashes are associated with the shorter distances between a traffic control signal and a PHB. No significant difference existed in any crash type for midblock PHBs compared to PHBs installed at intersections.

This ADOT study permitted the inclusion of a larger number of sites and a larger number of months of before and after data than other recent studies, which aided in the ability to find statistically significant results. Crash reductions were found to be significant at the 0.05 significance level for total crashes, FI crashes, FI rear-end crashes, and pedestrian-related crashes. Other crash types are also associated with significant reductions depending upon the reference group being used and statistical significance level being accepted.

6. Conclusions

Key findings from this project include the following

- Previous studies and this ADOT study have found a safety benefit in the installation of a PHB. The EB before-after evaluation found statistically significant reductions at the 5% significance level for several crash types, including:
  - A 25% reduction in severe (FI) total crashes (CMF of 0.75).
  - A 45% reduction in FI pedestrian-related crashes (CMF of 0.54).
  - A 29% reduction in FI rear-end crashes (CMF of 0.71).
- Midblock (two legs) versus intersection (three or four legs) does not make a difference with respect to safety at PHBs since no statistical difference in crashes between midblock locations and those PHBs at three- or four-leg intersections was found in the cross-sectional evaluation.
- The cross-sectional evaluation showed no statistically significant difference between the lower-speed and higher-speed PHB sites (posted speeds at 35 mph [56.3 kph] or below versus 40 mph [64.4 kph] or higher) for all crash types except rear-end crashes. For rear-end crashes, fewer rear-end crashes are present when the posted speed limit is 35 mph (56.3 kph) or below.

7. Discussion/practical applications

The findings from this study should encourage greater consideration for the PHB to accommodate pedestrian crossings. Inclusion of the PHB in the 2009 MUTCD has allowed more communities to use the treatment and realize the safety benefits of the device. Several previous research studies, along with this project, have clearly demonstrated that overall, the PHB is associated with fewer pedestrian and non-pedestrian vehicle crashes.

For some communities, the PHB is not considered when the roadway has a greater than 40 mph (64.4 kph) speed limit or at an intersection or driveway with stop controls on the side road.

The reasons given for limiting the PHB on higher-speed roads are the concerns that drivers are not expecting such a device on the high-speed road or that insufficient data or research findings exist to support the use of the treatment on this type of roadway. The cross-sectional evaluation finding of no statistically significant difference between the lower-speed and higher-speed PHB sites directly addresses the concern of limited data and provides the desired support for installing PHBs on higher-speed roads.

The development of the proposed language for the PHB section of the MUTCD involved a number of volunteers. When the 2009 MUTCD was published, it contained the following guidance statement that was added just prior to publication as a result of a comment made during the review process: "The pedestrian hybrid beacon should be installed at least 100 feet from side streets or driveways that are controlled by STOP or YIELD signs;" however, if followed, the PHB could not be used at intersections or many driveways. This statement thus limited the use of a very effective safety device at locations most in need of a pedestrian crossing treatment. Of particular concern was that the guidance statement was not based on research. Actually, the safety research from the pre-2009 MUTCD period (Fitzpatrick & Park, 2010) only included PHB installations at intersections and driveways. In other words, the available safety research did not support the inclusion of the aforementioned guidance statement in the 2009 MUTCD. Consequently, the National Committee on Uniform Traffic Control Devices, the volunteer organization that provides recommendations on changes to the MUTCD, approved a recommendation in 2011 to remove the guidance statement (https://ncutcd.org/wp-content/uploads/meetings/2011B/Attach-No.-4-Signals-Design-of-Pedestrian-Hybrid-Beacons-Section-4F.02.pdf, 2020). FHWA will consider that recommendation when developing the next edition of the MUTCD; however, it is not yet known if the recommendation to remove the guidance statement will occur. This ADOT study found that the number of legs—that is, two legs (midblock) versus three or four legs (intersection)—does not make a difference with respect to safety at PHBs. This finding provides additional support for the installation of PHBs at intersections or driveway approaches where pedestrian crossings are most likely to occur and for removing the guidance statement from the MUTCD.

Acknowledgement and funding

This work was supported by ADOT [contract number SPR-756].

References

ADOT Traffic Engineering Guidelines and Processes. (June 2015). Section 600—Traffic Signals, 640 Pedestrian Hybrid Beacon. Available at https://azdot.gov/docs/default-source/traffic-library/tgp0640-2015-06.pdf?sfvrsn=4. Accessed date: 5-9-19.

American Association of State Highway and Transportation Officials (AASHTO) (2010). *Highway Safety Manual*. Washington, DC: AASHTO.

Federal Highway Administration. (2009). Manual on Uniform Traffic Control Devices and Arizona Supplement, adopted January 13, 2012.

Federal Highway Administration. (2011). 4(09)-14 (I)—Red clearance interval in pedestrian hybrid beacon sequence. Available at https://mutcd.fhwa.dot.gov/resources/interpretations/pdf/4_09_14.pdf.

Fitzpatrick, K., & Park, E. S. (2010). Safety effectiveness of the HAWK pedestrian crossing treatment. Report No. FHWA-HRT-10-042, Federal Highway Administration, Washington, DC.

Fitzpatrick, K., & Pratt, M. (2016). Road user behaviors at pedestrian hybrid beacons, FHWA-HRT-16-039. Washington, DC.

Fitzpatrick, K., Turner, S., Brewer, M., Carlson, P., Lalani, N., Ullman, B., ... Whitacre, J. (2006). Improving pedestrian safety at unsignalized crossings. TCRP Report 112/NCHRP Report 562. Transportation Research Board.

Fitzpatrick, K., Avelar, R., Pratt, M., Brewer, M., Robertson, J., Lindheimer, T., & Miles, J. (2016). Evaluation of pedestrian hybrid beacon and rapid flashing beacons. FHWA Technical Report HRT-16-040. Washington, DC.

Fitzpatrick, K., Cynecki, M., Pratt, M., Park, E. S., & Beckley, M. (2019). Evaluation of pedestrian hybrid beacons on Arizona highways, ADOT-SPR-756. Available at https://apps.azdot.gov/files/ADOTLibrary/publications/project_reports/pdf/spr756.pdf.

Hauer, E. (1997). *Observational before-after studies in road safety: Estimating the effect of highway and traffic engineering measures on road safety*. Oxford, United Kingdom: Pergamon press, Elsevier Science Ltd.

NCUTCD (2020) Signal Design of Pedestrian Hybrid Beacon Section 4F.02 Available at https://ncutcd.org/wp-content/uploads/meetings/2011B/Attach-No.-4-Signals-Design-of-Pedestrian-Hybrid-Beacons-Section-4F.02.pdf. accessed on June 18, 2020.

Turner, S., Fitzpatrick, K., Brewer, M., & Park, E. S. (2006). Motorist yielding to pedestrians at unsignalized intersections findings from a national study on improving pedestrian safety. *TRR, 1982*, 1–12.

Zegeer, C., Srinivasan, R., Lan, B., Carter, D., Smith, S., Sundstrom, C., Thirsk, N.J., Zegeer, J., Lyon, C., Ferguson, E., and Van Houten, R. (2017a). Development of Crash Modification Factors for Uncontrolled Pedestrian Crossing Treatments. NCHRP Report 841.

Zegeer, C., Lyon, C., Srinivasan, R., Persaud, B., Lan, B., Smith, S., Carter, D., Thirsk, N.J., Zegeer, J., Ferguson, E., Van Houten, R., and Sundstrom, C. (2017b). "Development of Crash Modification Factors for Uncontrolled Pedestrian Crossing Treatments." Transportation Research Record: Journal of the Transportation Research Board, No. 2636, 1–8.

Kay Fitzpatrick, Ph.D., P.E., PMP. Kay is a Senior Research Engineer with the Texas A&M Transportation Institute. Her research interests include pedestrian treatments, roadway geometric design, traffic control devices, roadway safety, and the effects of posted speed limits on safety and operations. Her work has influenced changes in the Texas Department of Transportation *Roadway Design Manual*, FHWA *Manual on Uniform Traffic Control Devices* and the AASHTO *Green Book*.

Mike Cynecki, PE, PTOE. Mike is a Project Manager with Lee Engineering in Phoenix, Arizona. He has 40 years of traffic operations and safety experience, including 26 years with the City of Phoenix Street Transportation Department. Mr. Cynecki installed the first nine Pedestrian Hybrid Beacons in Phoenix and serves as an instructor for the FHWA Designing for Pedestrian Safety Workshops.

Eun Sug Park, Ph.D. Eun Sug is a statistician with the Texas A&M Transportation Institute. Park began her career with TTI in 2001, accumulating numerous awards and recognitions leading modeling and data analysis tasks in many transportation research studies for numerous state and federal agencies. Her expertise includes modeling and analysis of transportation data, safety analysis, Bayesian modeling, uncertainty assessment, high dimensional data analysis, environmental modeling, source apportionment/identification, and assessment of health effects of multiple air pollutants.

Michael P. Pratt, PE, PTOE. Mike is an assistant research engineer with the Texas A&M Transportation Institute. Pratt's main areas of interest include operational effects of roadway geometry and traffic control devices, traffic signal operations, and roadway safety.

Michelle Beckley, PE. Michelle is a project engineer with Y2K Engineering, LLC. Her professional experience has been focused in active transportation planning projects, safety studies, traffic impact analyses, parking studies, and ITS field deployments and evaluation projects.

Journal of Safety Research 78 (2021) 138–145

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigation of speed and trajectory of motorcycle riders at curved road sections of two-lane rural roads under diverse lighting conditions

Panagiotis Lemonakis *, Nikolaos Eliou, Theodoros Karakasidis

Department of Civil Engineering, University of Thessaly, Volos, Greece

ARTICLE INFO

Article history:
Received 15 September 2020
Received in revised form 15 February 2021
Accepted 25 May 2021
Available online 17 June 2021

Keywords:
Naturalistic riding
Speed
Trajectory
Deviation
Motorcycle
Horizontal curves

ABSTRACT

Introduction: Vehicular accidents at horizontal curves are over-represented compared to accidents that occur at tangent sections. Investigations have been conducted aimed at identifying the major causes that result in higher accident risk, both in terms of severity and rate, at curved road sections. Excessive or abrupt changes in speeding and improper vertical position are cited as major factors of lane departure, whereas other factors (either human or environmental) have also been documented. However, most research involves 4-wheel vehicles rather than other modes of transport that behave differently. More specifically, while motorcyclist fatalities occur more frequently than passenger vehicles, when accounting for vehicle distance traveled only a limited number of research studies address their behavior at curved road sections. *Method:* This paper presents the findings of field operational tests carried out by motorcyclists along two-lane rural roads with a wide range of horizontal curves using an instrumented motorcycle. Key objectives of the research included the conditions under which the motorcyclists differentiate their trajectory in regards to the direction of the horizontal curves, the correlation between the trajectory and the geometry of the road, and the impact of the lighting conditions on riders' behavior. *Results:* The research showed that motorcyclists tend to ride closer to the centerline of the road, neglect the hazards associated with dim lighting conditions, and maintain constant speed in the left hand and the right-hand horizontal curves.

1. Introduction

Driving a vehicle is a complex activity in which multiple factors are engaged, for example, user's psychology, vehicle's characteristics, presence of other road users, and road geometry (Bella, 2005). In particular, the negotiation of horizontal curves by motorcycle riders requires enhanced riding skills compared to tangent sections, especially along rural roads where single motorcycle crashes frequently occur. Inappropriate lateral position, speeding, high entry speed, and inattention are considered crucial parameters for motorcyclists' safety. However, the construction and maintenance of a safe road environment from the motorcyclist point of view do not always result in fewer motorcycle crashes because motorcycle riders tend to adopt hazardous attitudes when they feel safe (homeostasis; e.g. travel with excessive speed). Therefore, they compensate for any benefits originally gained and hence there are indications that the increase of risk awareness might be an effective measure against motorcycle accidents (Wang et al., 2018).

As discussed in a study that evaluated the safety adjacent to horizontal curves, the curved road sections are of particular importance because they are overrepresented in crashes. Indeed, in the United States, there are as many as three times the number of more severe crashes along horizontal curves compared to crashes at tangent sections (Khan, 2012). Smaiah et al. reports that more than 50% of the powered two-wheeler crashes occur in bends due to loss of control. They also mention that in France almost a third of the total powered two-wheeler crashes in curves are fatal (Smaiah, 2018). It is noteworthy that vulnerable road users, which also includes powered-two-wheeler riders, account for half of all fatalities in the EU (Morris et al., 2018).

The higher likelihood of accidents on curves is also mentioned in another study that investigates single-vehicle motorcycle crashes on horizontal curves in Florida. It is reported that 57% of the total fatal single motorcycle crashes occur on curves, although they represent only about 6% of the total road network (Wang et al., 2018; Xin et al., 2017a). The importance of single-vehicle accidents was also revealed in another study which found that this type of accident is notably widespread among motorcyclists and usually results in death and seriously injured casualties (Morris et al., 2018). Especially at road curves, the behavior of the motor-

* Corresponding author at: University of Thessaly, Faculty of Civil Engineering, Pedion Areos, 38334, Greece.

E-mail address: plemonak@uth.gr (P. Lemonakis).

https://doi.org/10.1016/j.jsr.2021.05.009
0022-4375/

cyclist is crucial in preventing or causing an accident (Wang et al., 2018).

The investigation of motorcyclists' trajectories is vital to provide comprehensive insights into steering corrections due to rider's failure to perceive the geometry of a curve or due to inappropriate speed in the road environment. Moreover, riders must be capable of sufficiently coping with hazardous situations that other road users create. This study aims to investigate motorcyclists' behavior, focusing on their speed and trajectory along two-lane rural using an instrumented motorcycle. The primary goals of the experiments and the post-process of the raw data were to investigate the conditions under which the motorcyclists differentiate their trajectory in regards to the direction of the horizontal curves, the correlation between the trajectory and the geometry of the road, and the impact of the lighting conditions on riders' behavior. The findings revealed that travel speed does not depend on the lighting conditions of the road environment or the direction of the horizontal curve, whereas the analysis of the trajectories showed that the riders travel closer to the centerline of the road instead of its right boundary.

2. Literature review

Taking into account the accident statistics, the efforts to implement measures to mitigate motorcycle accidents is disproportional low. For instance, in the United States a motorcyclist is 28–35 times more likely to experience a fatal accident compared to a passenger car driver when accounting for the traveled vehicle distance, whereas in 2016 51% more motorcyclists were killed compared to 2015 (Casanova-Powell, 2018; NHTSA, 2007). The interaction between geometric design and the road users' trajectory has been researched extensively since the 1970s (Spacek, 2005). The operation of a motorcycle is a much more challenging and complex activity compared to a passenger car but, nevertheless, the knowledge behind the interaction between road geometry and riding behavior is not clear-cut. For instance, a powered-two-wheeler rider has to perform a series of actions, some of them counterintuitive, in strict order sequential to effectively negotiate a curve, a lot more than a passenger car driver. The manipulation of the brakes and the steering between motorcycles and cars is different and, as a result, falls or running wide on a curve due to overbraking among motorcyclists is quite common (Wang et al., 2018).

Among the various factors that influence the frequency and severity of motorcycle accidents on rural curves, two are the most dominant: speeding and lateral position of the vehicles. These two factors are not independent of each other since the interactions between them determine the behavior of the drivers as they traverse a curve and therefore, taking also into account the different track that each rider follows, the forces applied on the motorcycle may vary considerably (Spacek, 2005; Wang et al., 2018). Moreover, as discussed by Xu et al, the curvature, velocity, and lateral position at the entry of a curve determine the shape of the vehicle's track, which also depends on the type of the vehicle. Consequently, not all of the vehicles will identically negotiate a curve even if they travel under the same traffic conditions (Xu, Luo, & Shao, 2018).

The effect of the curve type on rural two-lane motorcycle accidents was the subject of a recent study that concluded that as the sharpness of the curve increases, the likelihood of motorcycle crashes decreases as shown in Fig. 1 (Wang et al., 2018).

2.1. Speed as a contributing factor to road accidents

Speed behavior is highly correlated with the driving behavior throughout the horizontal curve sections of the roads and hence the speed profiles of the road users are of paramount importance

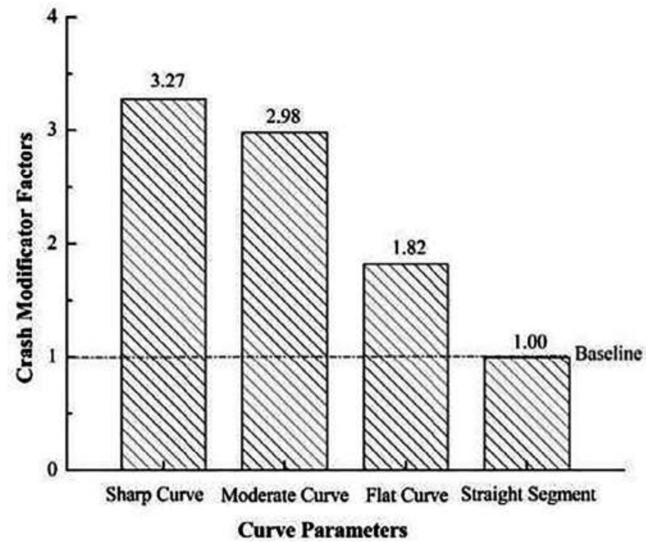

Fig. 1. Correlation of curvature and crash modification factors for motorcycle crashes.

to the design of the road alignment (Spacek, 2005) while, conversely, driving behavior depends on the road geometry (Said, Abd El Halim, & Yasser, 2010). Safety and comfort are both associated with drivers' speed throughout curves since speed determines the amount of lateral acceleration that will be applied to the driver (Lauffenburger, Basset, & Gissinger, 2005). Eustace and Indupuru analyzed the factors that contribute to motorcyclists' fatal injuries and concluded that excessive speeding increases the risk of motorcycling fatal crash occurrence (Eustace & Indupuru, 2011).

Wang et al quantified the impact of speeding on the occurrence of single motorcycle crashes, which constitute the majority of fatal motorcyclist accidents in the United States (Casanova-Powell, 2018) and found an increase of 10.84% when speeding is involved in motorcycle accidents (Wang, Lee, & Lin, 2013). Several studies cited in a study that investigates the correlation between injury severity and horizontal curve design show that speeding is the primary contributory factor to the severity of injuries in motorcycle crashes. Various researchers also quantified the impact of speeding on injury severity and reported that an increase in powered-two-wheeler accidents ranging from 16.78% to 24.11% must be expected when speeding is present (Wang et al., 2018; Xin et al., 2017a). Savolainen and Mannering concluded with a much higher percentage, stating that the likelihood of single motorcycle accidents is increased by 212% due to speeding (Savolainen & Mannering, 2006). Consequently, the implementation of speed control measures is an efficient action to counteract single motorcycle crashes (Xin et al., 2019).

The travel speed of the drivers depends on various parameters (e.g., road geometry, pavement, vehicle's characteristics, physiology), and is usually different from the design speed, whereas it does not have a fixed value through a curve (Xu, Lin, & Shao, 2017). In the EU about one out of four motorcycle accidents occur because the riders traveled with disproportional excessive speed for the prevailing conditions (Morris et al., 2018). Wang et al also concluded that excessive speed is the most important factor for fatal and seriously injured riders involved in single motorcycle accidents on rural curves (Wang et al., 2018). In particular, the speed at the entry of the curve is critical as its magnitude determines to a great extent whether the rider will encroach the centerline or the shoulder of the road (Xu et al., 2018). Therefore, by reducing the entry speed the track of the vehicles is consistent with the expectations of the driver, and hence the possibility of accidents decreases (Xu et al., 2018).

2.2. Evaluation of track behavior at curved road segments

Although speeding is the most crucial factor leading to crashes, there is evidence that the appropriate lateral position of the vehicles is an efficient countermeasure against single-vehicle accidents that represent a big proportion of total crashes. Indeed, as mentioned in a study that evaluated the trajectory variability, the loss of control type of accidents is observed in 40% of fatal crashes in France (Rosey & Auberlet, 2012). Jacob and Violette found that the real trajectory of a vehicle is the outcome of the interaction "driver–vehicle–road" system and can be used effectively to improve road safety (Jacob & Violette, 2012). Rosey and Auberlet also concluded that the trajectory variability is a tool that can be used to identify road sections where improvements might be deemed necessary to reduce the likelihood of accidents (Rosey & Auberlet, 2012).

It is amply documented that only a small proportion of the drivers precisely follow the middle of the travel lane (Das Vivek, Jayashree, & Rahul, 2016; Xu et al., 2018). Instead, the majority of them either cut-the-curve or perform steering corrections due to false judgments to bring their vehicle back to the desired trajectory. If these steering corrections require sharp maneuvers and/or the pavement is wet and/or the driver simultaneously applies their brakes then the vehicle becomes particularly unstable and it is quite possible to depart the travel lane due to the increased demands in lateral friction. Indeed, the centrifugal acceleration in specific sections of a curve could be particularly higher than the one expected according to the design standards (Spacek, 2005).

As discussed in a recently published study, single motorcycle crashes frequency decreases by 0.74% for an increase in curve radius by 1% (Xin et al., 2019). Negotiating small radii curves requires special riding skills since the available space is limited and, hence, in such curves the likelihood of single motorcycle accidents is increased (Wang et al., 2018). Besides, sharp curves tend to attract sensation seeker riders who perform risky activities resulting in higher crash risk (Xin et al., 2017a; Xu et al., 2018). In fact, 5% of motorcycle accidents in the EU involve risky riders (Morris et al., 2018). In general, smoother curves are safer compared to tight ones for motorcyclists (Wang et al., 2018); (Xin et al., 2017b) and the more the curve radius the less the likelihood of a motorcyclist accident due to less speed variation and increase of sight distance (Wang et al., 2018).

All of the trajectories that deviate from the center of the traffic lane are considered incorrect and increase the likelihood of crashes. Therefore the less the variability of the lateral position the more uniform and safe is the trajectory of the vehicles (Das Vivek et al., 2016), and hence lateral position variability can be used as an indicator of design consistency evaluation (Rosey & Auberlet, 2012). Lin et al found that longer spiral transition curves are associated with less lateral deviation and vice versa, and they suggested that the length of the transition curve should be equal to a certain percentage of the curve radius (Lin, Yang, & Pan, 2011). The steering length of the driver who precisely follows the geometry of the curve is about half of the steering length of the driver who cut-the-curve and consequently the latter begins to steer their vehicle long before the TC point (Shu, Shao, & Xu, 2016). Nevertheless, curves with short lengths or small radius increase the likelihood of motorcycle crashes (Xin et al., February 2019). However, the finding of Wang et al suggests the opposite conclusion and, more specifically, they found that an increase in curve length is linked to more motorcycle crashes (Wang et al., 2018).

Four research studies presented specific track patterns based on experimental data (Shu et al., 2016; Spacek, 2005; Xu et al., 2017). All of them concluded that the proportion of drivers who follow the center of the traffic lane is negligible (varying from 0% to 8%) and, therefore, the value of the curve radius that is used to calculate a series of geometric features does not reflect the real driving curvature; hence, the trajectory radius should be used instead. Lastly, the findings of other studies also confirm that curves increase the likelihood of fatal motorcycle accidents, especially during nighttime travel (Eustace & Indupuru, 2011; Xin et al., 2017a).

2.3. Literature review conclusions

A research study that investigated the tracking behavior in curved areas cites various papers suggesting that the link between road geometry and accidents, in terms solely of speed, is deemed to fail and makes a reference to a dissertation conducted by Friedinger who concluded that the corrections to the traveling direction (not speeding) is the primary cause of curve accidents (Spacek, 2005). Spacek also agrees that the root cause of single-vehicle accidents due to loss of control that often occur along horizontal curves cannot be attributed solely to the inappropriate high speed (Spacek, 2005), whereas Yuen et al. found that riders' behavior is strongly affected by the riding experience and the annual distance traveled (Yuen, Karim, & Saifizul, 2014). Another conclusion is that the correlation between horizontal curves and motorcycle injury severity is not well documented yet since the relevant studies do not determine how the injury severity is influenced by the characteristics of the curves (Xin et al., 2017a).

3. Data collection

An in-depth analysis of motorcycle riders' behavior, who probably constitute the most vulnerable road users, contributes to the improvement of riders' safety and traffic management in general (Barmpounakis, Vlahogianni, & Golias, 2015). Therefore, a cost-effective methodology to record and process reliable and accurate behavioral data, such as speed and lateral placement of the vehicles, is the key to this direction. Aiming at this goal and for the needs of this paper, 18 recruited riders registered to the motorcycle club of Volos City in Greece traveled along a two-lane rural road under natural riding conditions on a motorcycle properly instrumented with a camera and a GPS receiver of high position accuracy during daytime and nighttime.

The principal benefit of this experiment over other methods aimed at obtaining riding performance data is that it allows the recording of data under real riding conditions without the limitations of riding simulators or the subjective approach of questionnaires. In addition, instrumented motorcycles record detailed data of a limited number of riders, and consequently for a limited number of trajectories (Jacob & Violette, 2012). By choosing long enough experimental routes, the total database allows the researchers to draw firm conclusions. In line with this, the measurements took place along a road section of approximately 10.38 km consisting of 21 horizontal curves, which is depicted in Fig. 2.

The riders were not aware of the goals of the measurements and were instructed to travel as naturally as possible. One-half of them were employed during the daytime measurements and the other half during the nighttime measurements. Their mean age was 40.39 years (SD = 9.90 years) and they were all experienced riders with a mean valid riding license of 19.53 years (SD = 11.90 years). They were all owners of at least one registered motorcycle while their riding hours ranged from 1 to 8 per day (Mean = 2.83 h, SD = 1.99 h).

All of them traveled in both directions of a rural road segment, starting from the point 39°18′08.2″N 22°54′47.6″E (point A in Fig. 2) and making a U-turn at the point 39°17′01.8″N 22°50′15.9″E (point B in figure B) and driving back to the starting point. For practical reasons all riders started from point A and com-

Fig. 2. Map, detail and geometric characteristics of the horizontal curves of the experimental route.

pleted their measurement in point A after performing a U-turn in point B. It has to be noted that the literature review regarding the planning and execution of naturalistic driving studies did not reveal any order effects concerns that should be taken under consideration. Consequently, a dataset of 756 traverses through the curves recorded and processed, which is considered adequate enough. Only a few lighting poles were installed along the testing route, most of which were out of order and, hence, the artificial light during the nighttime measurements emanated from the instrumented vehicle's headlamps.

The measurements were carried out with the use of a medium-capacity motorcycle namely a DL V-Strom 650 cc Suzuki, particularly popular on the Greek rural roads. The day and time of the measurements were selected in such a way that the traffic volume from other road users was very limited. Thus, the behavior of the riders was unaffected by the presence and the behavior of other vehicles, and hence the recorded data reflect the unobstructed travel speed and trajectory of the participants.

As already implied, the research study focused on the behavior of the riders expressed by their travel speed and trajectory at horizontal curves, with main geometrical characteristics presented in Table 1. The first row of the Table concerns the code name of the curves, whereas the second, third, and fourth row the Radius, Length, and Direction (either Left-hand or Right-hand) of each curve, respectively.

After the execution of the measurements, the raw data was converted to *csv* and *dxf* files for further analysis using more widespread software. In this way, the trajectory and speed profile of each measurement were accurately known since the equipment that was installed on the bike recorded the position and speed of the rider with a frequency of 5 Hz. The investigation, among other goals, attempted to identify any rider tendencies to travel closer to the centerline of the road or not. For that purpose, the centerline of the travel lane was drawn based on which deviation area (either to the right or to the left) was calculated. Fig. 3 illustrates a random example of this step of the process. The upper part of the Figure presents the boundaries of a curve, whereas the lower part presents a detail of the start/end of the curve in which the various symbols (i.e., the travel direction, the road axis, the centerline of the traffic lanes, two random trajectories (one for each travel direction)) and the area of deviation (either to the right or to the left of the centerline of the traffic lane) are explained.

Figs. 4–7 present the mean velocities and the total deviation from the centerline of the traffic lane of the riders per curve and lighting conditions (daylight–nighttime) of the road sections Volos – Anchialos and vice versa, respectively. It must be noted that the deviation from the centerline is expressed as the area between the centerline of the travel lane and each trajectory corresponds to the total area of deviation regardless of how many times the two lines intersect.

The data behind these figures were then copied to the Statistical Package for the Social Sciences (SPSS) for further analysis. Normality tests (both Shapiro-Wilk and Kolmogorov-Smirnov tests) were conducted before processing the data, which showed that they do not significantly deviate from a normal distribution. The investigation oriented toward the speed differential with relation to the prevailing environmental conditions (e.g., lighting conditions, the direction of the horizontal curve, deviation from the centerline of the travel lane).

4. Data process

4.1. Route Volos – Anchialos

The null hypothesis is that there is no statistically significant difference between the mean speed of riders who travel under different lighting conditions along horizontal curves. Therefore, the independent variable is the lighting conditions (two groups: daytime and nighttime), while the dependent variable is the traveled mean speed. The null hypothesis was tested by conducting a paired two-tailed *t*-test to determine if there is a difference in the mean scores of the two groups. The results of the test revealed that there is no statistically significant difference between the mean speed along horizontal curves of daytime riders ($n = 21$, $M = 99.63$,

Table 1
Geometric characteristics of the horizontal curves of the experimental route.

Curve	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	K21
R_{curve} (m)	900	250	330	200	900	800	200	195	225	380	270	220	200	600	370	400	300	300	220	220	250
L_{curve} (m)	217	261	257	275	269	236	406	380	144	236	257	200	241	176	330	389	212	132	139	166	177
Direction	Left	Left	Right	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Right	Left	Right	Left	Right	Left	Left

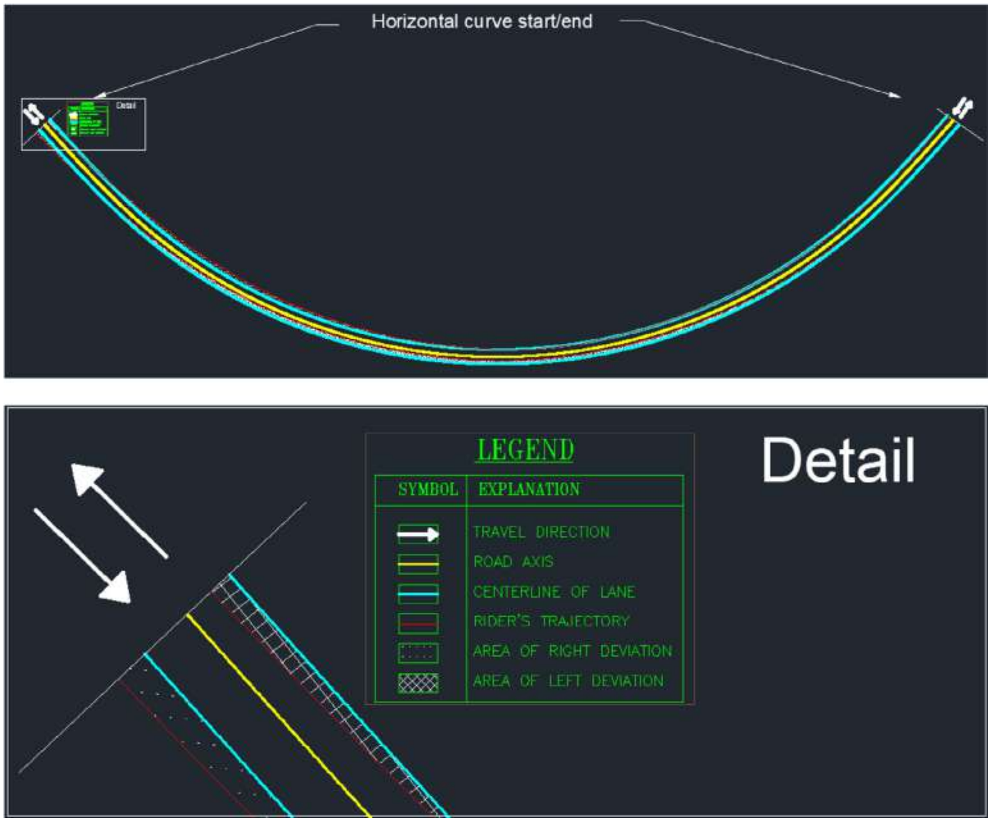

Fig. 3. Right/Left deviation from the centerline of the travel lane.

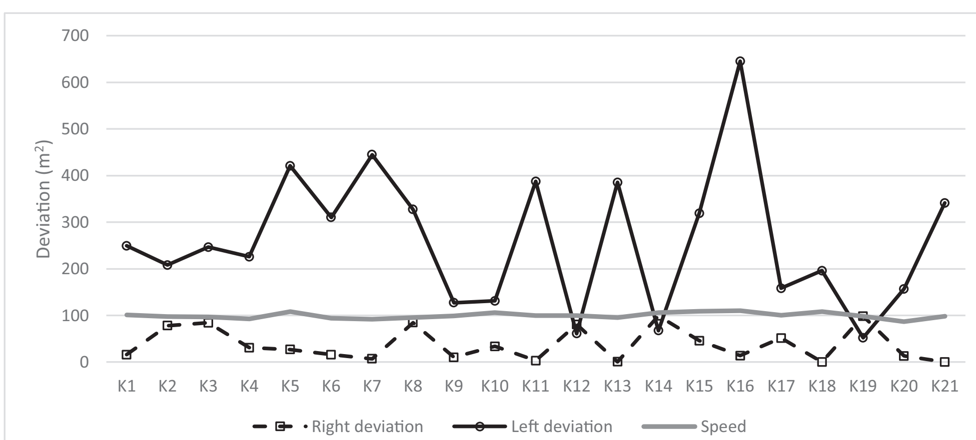

Fig. 4. Average speed and deviation from the travel lane's centerline per curve: Daytime Volos - Anchialos.

SD = 6.28) and nighttime riders (*n = 21, M = 101.41, SD = 5.03*), *t(20) = –1.69, p = .11*.

Another aim of the research is to investigate whether the direction of the curves (left-hand or right-hand) affects the magnitude of the travel speed regardless of the prevailing light conditions. Therefore, the null hypothesis, in this case, is that the travel speed does not differentiate between left-hand and right-hand horizontal curves. By implementing a similar approach, an independent two-tailed *t*-test was conducted that revealed that there is no significant difference between the mean speed of the riders traveling along left-hand (*n = 22, M = 100.42, SD = 6.10*) and right-hand curves (*n = 20, M = 100.63, SD = 5.36*), *t(40) = –0.12, p = .0.90*.

Evidence suggests that the vehicles tend to travel closer to the centerline compared to the edge of the road (Taragin, 1945). This finding is also confirmed in the framework of the present research study. More specifically the deviation to the left side (*n = 42*, $M = 269.03\ m^2$, *SD = 168.50*) as opposed to the right side (*n = 42*, $M = 34.01\ m^2$, *SD = 36.20*) of the travel lane was predominant, meaning that the riders tend to ride closer to the axis of the road during both the nighttime or the daylight measurements (*t(41) = 8.01, p < .001*).

Moreover, to investigate whether there is a substantial difference regarding the magnitude of the left deviation against the lighting conditions (daylight or nighttime) and the direction of

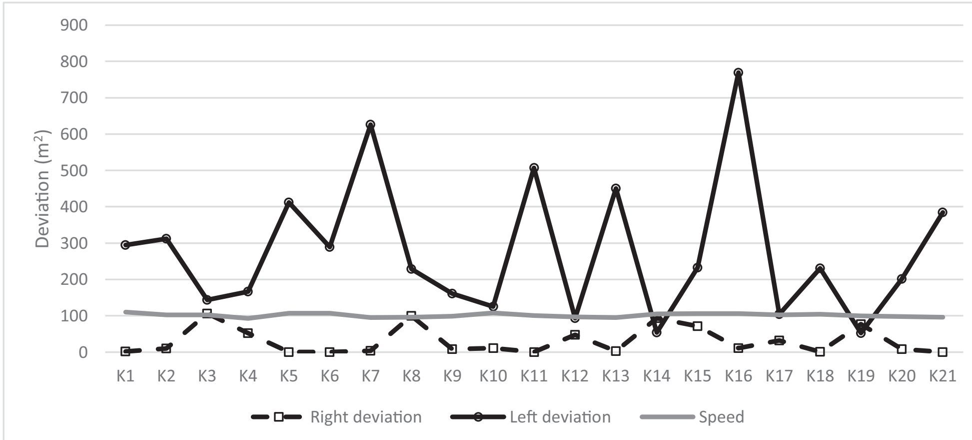

Fig. 5. Average speed and deviation from the travel lane's centerline per curve: Nighttime Volos – Anchialos.

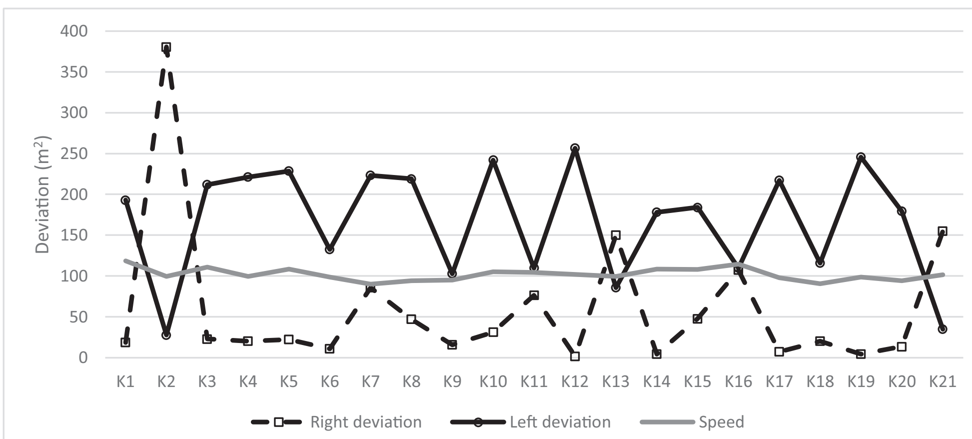

Fig. 6. Average speed and deviation from the travel lane's centerline per curve: Daytime Anchialos – Volos.

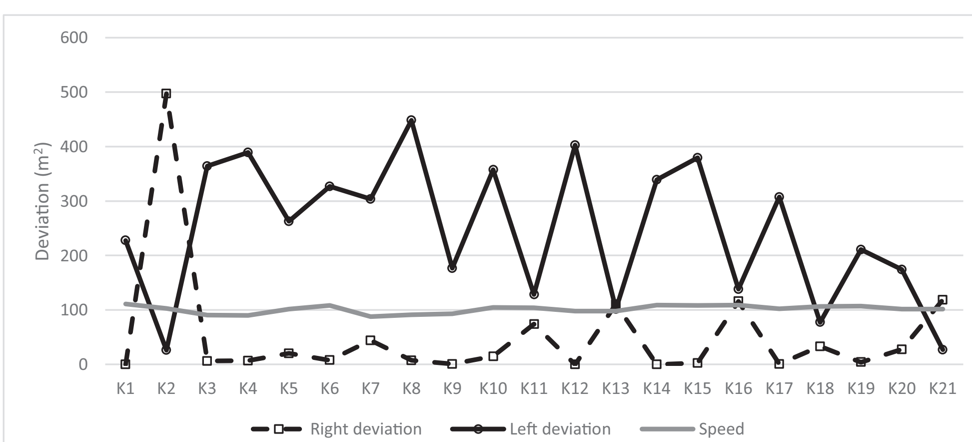

Fig. 7. Average speed and deviation from the travel lane's centerline per curve: Nighttime Anchialos – Volos.

the horizontal curves (either left-hand or right-hand) a paired two-taileded and an independent *t*-test were conducted, respectively. The null hypotheses were that the trajectory of the riders toward the centerline of the road is independent of the status of the ambient light and the direction of the horizontal curves. The results showed that there is no significant difference in the scores for the nighttime measurements ($n = 21$, $M = 277.96$, $SD = 189.75$) and daylight measurements ($n = 21$, $M = 260.09$, $SD = 148.42$), $t(20) = -1.08$, $p = 0.293$ along the horizontal curves of the direction Volos – Anchialos. However, statistical significant difference was identified between the left-hand ($n = 22$, $M = 359.51$, $SD = 170.50$) and the right-hand ($n = 20$, $M = 169.49$, $SD = 96.04$), $t(34) = 4.50$, $p < .001$ horizontal curves.

4.2. Route Anchialos – Volos

A similar approach was implemented in the opposite direction as well by performing independent or paired two-tailed *t*-tests accordingly. Some of the results were also confirmed based on the measurements of this direction, while others did not. More specifically:

- No statistically significant difference between the mean speeds in horizontal curves of daytime riders ($n = 21$, $M = 101.89$, $SD = 7.61$) and nighttime riders ($n = 21$, $M = 101.27$, $SD = 7.12$), $t(20) = 0.43$, $p = .0.67$ was established.
- The implementation of an independent two-tailed t-test did not reveal a statistically significant difference between the mean speed of the riders traveling along left-hand ($n = 20$, $M = 101.61$, $SD = 6.66$) and right-hand curves ($n = 22$, $M = 101.46$, $SD = 7.98$), $t(40) = 0.07$, $p = .94$.
- The deviation to the left side ($n = 42$, $M = 206.84\ m^2$, $SD = 110.01$) rather than to the right side ($n = 42$, $M = 142.12\ m^2$, $SD = 97.70$), $t(41) = 5.29$, $p < .001$ of the travel lane was also predominant for both the nighttime and daylight measurements.
- Contrary to the opposite direction, a statistically significant difference was found in regards to the impact of the ambient lighting on the magnitude of the left deviation. Particularly, the results showed that the difference between the nighttime measurements ($n = 21$, $M = 246.15$, $SD = 129.42$) and daylight measurements ($n = 21$, $M = 167.53$, $SD = 69.36$), $t(20) = -4.35$, $p < .001$ is indeed statistically significant.
- Lastly, the impact of the direction of the horizontal curves to the tendency of the riders to drive closer to the centerline of the travel lane is also statistically significant since the p-value for the deviation of the left-hand ($n = 20$, $M = 281.65$, $SD = 88.73$) and the right-hand ($n = 22$, $M = 138.82$, $SD = 79.54$) horizontal curves is less than 0.05 ($t(38) = 5.47$, $p < .001$).

5. Conclusions

Each driver accepts a certain risk level proportional to their driving abilities, attempting to minimize the travel time. If the perceived risk is less than the acceptable one then the driver increases the travel speed and changes their vertical position striving to maximize the cost-benefit ratio (Bella, 2005). Safe riders adapt their speed and lateral position to the prevailing driving conditions.

Although the reaction time and, consequently, the critical stopping distance is much increased during the night, the riders do not reduce their traveling speed when the light conditions are dimmed. That poses a threat to their safety since nighttime single-vehicle accidents are over-represented in rural horizontal curves (Green, 2003) and hence it should be further investigated. In addition, the lack of overhead artificial lighting reduces the visual perception ability of the riders to such an extent that they might not be able to stop within the illuminated distance created by their headlamps.

Therefore, either the riders are not aware of the hazards that are associated with nighttime riding in regards to the extended critical stopping distance, or other benefits exist that eliminate this threat. Indeed, during nighttime trips, the riders might more efficiently perceive oncoming traffic and adjust their trajectory and travel speed accordingly. However, oncoming traffic is just one of the reasons causing emergency braking, while other hazards (e.g., stray animals crossing the road, obstacles on the pavement, potholes) are probably neglected by the riders.

On the left-hand horizontal curves, the riders have the opportunity to traverse a trajectory of less curvature and, consequently, increase their comfort and travel speed by exploiting a greater portion of the pavement width. Moreover, as discussed in a study by Lemonakis et al., drivers differentiate their trajectory based on the direction and the curvature of the horizontal curves (Lemonakis, 2014). However, according to the results of the analysis, the riders do not alter their travel speed between right-hand and left-hand horizontal curves regardless of their curvature.

The investigation in both directions of either daylight or nighttime measurements revealed that riders' trajectory in horizontal curves is closer to the centerline of the road to a great extent. This observation is probably justified by the fact that riders feel more comfortable when they keep a safe clearance from the shoulder where they assume that greater hazards might exist compared to the oncoming traffic (e.g., pedestrians, stationary vehicles, debris). Further investigation is recommended on this topic since oncoming traffic of heavy or overtaking vehicles increases the driving workload and the manipulation skills of the riders.

6. Recommendations for future research

Single motorcycle accidents on a wet road surface, with poor pavement conditions on non-access controlled-curves, are less likely to occur (Wang et al., 2018; Xin et al., 2017a). This surprising finding implies that single motorcycle accidents can be prevented to a great extent if the riders have a timely warning about the imminent hazards. Since the root cause of the vast majority of traffic accidents is human error, as discussed in a recent study (Raipuria, 2017), more light must be shed on the way that riders perceive the traffic conditions and how they adjust their speed and lateral position afterward. This would be the first step to propose countermeasures against motorcyclist accidents, whereas the second step would be the proper education and training of the riders (Xin et al., 2019), both considered currently insufficient (Casanova-Powell, 2018; Smaiah, 2018). As Wang et al concluded, on the one hand motorcyclists take all the precaution measures against the hazards that they are made aware of in a timely and proper fashion, while on the other hand they are not properly trained about the complexity and special skills required while riding (Wang et al., 2018).

7. Availability of data and materials

All data generated or analyzed during this study are included in this published article. For further information please contact the corresponding author.

Funding

This article is funded by Stavros Niarchos Foundation and materialized by the University of Thessaly. The authors would like to thank the Stavros Niarchos Foundation for the scholarship of this study and the University of Thessaly for the generous support.

Acknowledgements

Not applicable.

References

Barmpounakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2015). Intelligent transportation systems and powered two wheelers traffic. *IEEE Transactions*

on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2015.2497406.
Bella, F. (2005). Speeds and lateral placements on two-lane rural roads: Analysis at the driving simulator. In 13th international conference "road safety on four continents". Warsaw, Poland.
Casanova-Powell. (2018). Tara motorcyclist traffic fatalities by state. Washington, DC: Governors Highway Safety Association.
Das Vivek, R., Jayashree, M., & Rahul, S. (2016). Lateral placement of vehicles on horizontal curves. In *11th transportation planning and implementation methodologies for developing countries, 2016* (pp. 43–51). Mumbai, India: Transportation Research Procedia.
Eustace, D., & Indupuru, Vamsi Krishna (2011). *A simplified method for analyzing factors contributing to motorcyclists' fatal injuries in Ohio* https://ecommons.udayton.edu/cee_fac_pub/3. Civil and Environmental Engineering and Engineering Mechanics Faculty Publications.
Green, E. R. (2003). *Roadway lighting and driver safety: KTC-03-12/SPR247-02-1F*. University of Kentucky - College of Engineering - Kentucky Transportation Center.
Jacob, B., & Violette, E. (2012). Vehicle trajectory analysis: an advanced tool for road safety. *Procedia - Social and Behavioral Sciences, 48*, 1805–1814. Transport Research Arena 2012.
Khan, G. et al. (2012). Horizontal curves, signs, and safety. *Transportation Research Record: Journal of the Transportation Research Board, 2279*. https://doi.org/10.3141/2279-15.
Lauffenburger, J. P., Basset, M., & Gissinger, G. L. (2005). Design of a coupled longitudinal-lateral trajectographic driver model. *16th Triennial world congress*. Prague, Czech Republic: IFAC.
Lemonakis, P. V. et al. (2014). A new methodology for approaching motorcycle riders' behavior at curved road sections. *European Transport Research Review, Springer*. https://doi.org/10.1007/s12544-014-0132-6.
Lin, Y., Yang, Z., & Pan, X. (2011). Effect of spiral transition curve length on vehicle path. *Xinan Jiaotong Daxue Xuebao: Journal of Southwest Jiaotong University, 46* (2), 200–204.
Morris, A. P. et al. (2018). *Safer Wheels study on powered two-wheeler and bicycle accidents in the EU Final report*. Luxembourg: Publications Office of the European Union.
NHTSA Traffic Safety Facts [Report]. (2007). DOT HS 810 990: National Center for Statistics and Analysis.
Raipuria, G. (2017). *Vehicle trajectory prediction using road structure*. Faculty of Mechanical, Maritime and Materials Engineering: Delft University of Technology.
Rosey, F., & Auberlet, J.-M. (2012). Trajectory variability: Road geometry difficulty indicator. *Safety Science, 50*(9), 1818–1828. https://doi.org/10.1016/j.ssci.2012.04.003.
Said, D., Abd El Halim, O., & Yasser, H. (2010). Methodology for driver behaviour data collection and analysis for integration in geometric design of highways. *4th International Symposium on Highway Geometric Design. Valencia, Spain*.
Savolainen, P., & Mannering, F. (2006). Probabilistic models of motorcyclists' injury severities in single-and multi-vehicle crashes. *Accident Analysis and Prevention, 39*, 955–963.
Shu, H., Shao, Y.-M., & Xu, J. (2016). Effects of geometric features of highway horizontal alignment on steering behavior. *Journal of Vibroengineering, 18*(16). ISSN 1392-8716.
Smaiah, S. et al. (2018). A practical approach for high precision reconstruction of a motorcycle trajectory using a low-cost multi-sensor system. *Sensor, 18*, 2282. https://doi.org/10.3390/s18072282.
Spacek, P. (2005). Track behavior in curve areas: An attempt at typology. *Journal of Transportation Engineering, 9*(669). https://doi.org/10.1061/(ASCE)0733-947X(2005)131.
Taragin, A. (1945). Public roads a journal of highway research – Effect of roadway Wodth on vehicle operation. *Public Roads Administration: Federal Works Agency, 24*(6).
Wang, Z. et al. (2018). *Study on motorcycle safety in negotiation with horizontal curves in Florida and development of crash modification factors [Report]*. Center for Urban Transportation Research - University of South Florida: Florida Department of Transportation.
Wang, Z., Lee, C., Lin, P.-S. (2013). Modeling injury severity of single-motorcycle crashes on curved roadway segments. Washington DC: 2104 TRB 93rd Annual Meeting, Vol. 14=5123.
Xin, C., et al. (2019). Development of crash modification factors of horizontal curve design features for single-motorcycle crashes on rural two-lane highways: A matched case-control study. Accident Analysis & Prevention. 123:51–59; doi: 10.1016/j.aap.2018.11.008.
Xin, C., et al. (2017a). Modeling safety effects of horizontal curve design on injury severity of single-motorcycle crashes with mixed-effects logistic model. Transportation Research Record: Journal of the Transportation Research Board, 2637; 38–46.
Xin, C. et al. (2017b). Safety effects of horizontal curve design on motorcycle crash frequency on rural, two-lane, undivided highways in Florida. *Transportation Research Record Journal of the Transportation Research Board, 2637*, 1–8. https://doi.org/10.3141/2637-01.
Xu, J., Lin, W., & Shao, Y. (2017). New design method for horizontal alignment of complex mountain highways based on "trajectory speed" collaborative decision. *Advances in Mechanical Engineering, 9*(4), 1–18. https://doi.org/10.1177/1687814017695437.
Xu, J., Luo, X., & Shao, Y.-M. (2018). Vehicle trajectory at curved sections of two-lane mountain roads: a field study under natural driving conditions. *European Transport Research Review, 2018*. https://doi.org/10.1007/s12544-018-0284-x.
Yuen, C. W., Karim, M. R., & Saifizul, A. (2014). Investigation on motorcyclist riding behaviour at curve entry using instrumented motorcycle. *The Scientific World Journal - Hindawi Publishing Corporation, 2014*. https://doi.org/10.1155/2014/968946.

Panagiotis Lemonakis I am the co-author of more than 32 scientific research papers that are published in national and international scientific journals and conferences proceedings whereas I have contributed to the publication of a book chapter. Moreover I have actively participated to the European Project 2-Be-Safe (http://www.2besafe.eu/) as member of the Highway Engineering Lab of Department of Civil Engineering in Volos Greece. On 2018 I qualified to receive a scholarship from Stavros Niarchos Foundation for post-doc research in the field of Highways Safety. At the moment I am teaching Highway Engineering as a lecturer to the students of the University of Thessaly.

Dr. Nikolaos Eliou is currently Professor in Highway Design and Road Safety and Professor in Postgraduate Department "Transportation projects Design and Management" at the University of Thessaly, Greece. His primary research and professional interests are: Road Safety Audits and Roadside Safety, Road Design Guidelines revision and Road Safety Expert. Dr. Eliou received his B.S. in Civil Engineering at the Aristotle University of Thessaloniki, Greece (1984) and his PhD in Transportation Engineering (1990). He has taught several undergraduate and graduate courses in the areas of Transportation and Road Safety. He has been involved as investigator in several research programs funded by the European Union and the Greek Government. Dr. Eliou is the author of 33 publications in international peer reviewed scientific journals, and 133 international and national conference papers and presentations. He has supervised several PhD students in the area of highway engineering and road safety, as well as a significant number of Master Theses.

Dr. Theodoros Karakasidis is currently Professor of Applied Physics at the University of Thessaly, Greece. Before joining the University of Thessaly. His primary research interests are: computational materials science, statistical physics, non-linear dynamics, time series analysis and machine learning. Dr. Karakasidis received his B.S. in Physics at the Aristotle University of Thessaloniki, Greece (1989), his M.S. (1991) and his PhD in Physics (1995) from University Pierre et Marie Curie (Paris 6), France. He has taught several undergraduate and graduate courses in the areas of Physics, numerical analysis and time series analysis. He has been involved as senior investigator and principal investigator in several research programs funded by the European Union and the Greek Government. Dr. Karakasidis is the author of nearly 100 publications in international peer reviewed scientific journals, and 80 international and national conference papers and presentations. He has supervised several PhD students in the area of computational materials modeling and time series analysis, as well as a significant number of Master Theses. He has served as a reviewer in international scientific journals like Scientific Reports, Physical Review E, Physical Review Letters, Chaos, Computational materials Science. He has participated as guest editor in several special issues and also as a member of the organizing committee of international conferences and workshops.

Journal of Safety Research 78 (2021) 292–302

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

“Like it’s wrong, but it’s not that wrong:” Exploring the normalization of risk-compensatory strategies among young drivers engaging in illegal smartphone use

Fareed Kaviani [a,*], Kristie L. Young [c], Brady Robards [b], Sjaan Koppel [c]

[a] Monash Sustainable Development Institute, Monash University, Australia
[b] School of Social Sciences, Monash University, Australia
[c] Monash University Accident Research Centre, Monash University, Australia

ARTICLE INFO

Article history:
Received 25 January 2021
Received in revised form 19 March 2021
Accepted 16 June 2021
Available online 3 July 2021

Keywords:
Smartphone use while driving
Social norms
Risk compensation
Young drivers
Road safety

ABSTRACT

Introduction: Young drivers are the most vulnerable road users and most likely to use a smartphone illegally while driving. Although when compared with drink-driving, attitudes to illegal smartphone risk are nearly identical, smartphone use among young drivers continues to increase. *Method:* Four in-depth focus groups were conducted with 13 young (18–25 years) drivers to gain insight into their perceptions of the risks associated with the behavior. Our aim was to determine how drivers navigate that risk and if their behavior shapes and informs perceptions of norms. *Results:* Three key themes emerged: (a) participants perceived illegal smartphone use as commonplace, easy, and benign; (b) self-regulatory behaviors that compensate for risk are pervasive among illegal smartphone users; and (c) risk-compensation strategies rationalize risks and perceived norms, reducing the seriousness of transgression when compared with drink-driving. Young drivers rationalized their own use by comparing their selfregulatory smartphone and driving skills with those of “bad drivers,” not law abiders. *Practical Applications:* These findings suggest that smartphone behaviors shape attitudes to risk, highlighting the importance for any countermeasure aimed at reducing illegal use to acknowledge how a young person’s continued engagement in illegal smartphone use is justified by the dynamic composition of use, risk assessment and the perceived norms.

1. Introduction

Young (18–25) newly-licensed drivers are most at risk of crashing (Regev et al., 2018), making road crashes the leading cause of death for 15–29 year olds across the globe (World Health Organization, 2018). Their inexperience (Scott-Parker et al., 2012), poor risk analysis (Simons-Morton et al., 2014; White et al., 2011), and inclination to take greater risks on the road all contribute to an increased likelihood of crashing. These factors also contribute to young drivers being more susceptible to distraction-related crashes (Buckley et al., 2014). Although they are the age group least likely to drive at least weekly in Victoria, Australia (where this study is conducted) (Transport Accident Commission, 2019), they are most likely to use a smartphone for longer periods of time per day (Kaviani, Robards, et al., 2020), including while driving (Kaviani, Young, et al., 2020).

Their reasons for use are varied, with many using their devices beyond merely call and text functionality; navigation, music, and social media applications such as Snapchat have emerged as popular smartphone behaviors young drivers are engaging with illegally while driving (George et al., 2018; Truelove et al., 2019).

In an effort to reduce distraction caused injuries or deaths, licensing jurisdictions around the world have introduced Graduated Licensing System (GDL) laws that gradually lift licensing restrictions (e.g., peer passenger restrictions, zero BAC) as the experience of the novice driver increases[1]. For example, in Victoria, Australia, all learner and probationary drivers are forbidden from using a mobile phone for any function (including hands-free and smart-assist functions such as Apple Carplay) while driving (including while stationary but not parked). Victorian full license holders can use a phone to make or receive a call and use the audio/music functions or GPS if the device is secured in a commercial cradle or

* Corresponding author.
E-mail address: fareed.kaviani@monash.edu (F. Kaviani).

[1] For more information on Victoria’s Graduated Licensing System please see: https://www.vicroads.vic.gov.au/safety-and-road-rules/driver-safety/young-and-new-drivers/victorias-graduated-licensing-system

https://doi.org/10.1016/j.jsr.2021.06.010
0022-4375/

can be operated hands-free and is not touching the driver's body (VicRoads, 2019).

Despite these laws, self-reported use remains high. In 2019, 37.1% of all Victorian drivers admitted to using a smartphone illegally while driving in the past month (Kaviani, Young, et al., 2020a), with 66.4% of 18–25 year-olds admitting to illegal use. Although the behavior is not declining in frequency, perceptions of risks associated with smartphone use while driving are decreasing when compared to previous reports (Transport Accident Commission, 2018). Although still high, young drivers rate the dangers of smartphone use slightly lower than those associated with driving over the legal blood alcohol limit (9.1/10 and 9.5/10, respectively; Transport Accident Commission, 2018).

To effectively counter illegal use it is necessary to understand why young drivers continue to engage with smartphones, while acknowledging the behavior is almost as risky as drink-driving, and why those perceptions of risks are declining. In this paper we report on focus group findings on three key themes: (1) how illegal smartphone use is perceived as benign; (2) the ways smartphone behaviors are linked to confidence, risk perception, road conditions, and task demand; and (3) how the risk-compensation strategies drivers describe are linked to perceptions of norms and the seriousness of illegal smartphone use while driving.

1.1. Norms and risk

Human behavior is guided by social norms (Deutsch & Gerard, 1955), and those norms—a collection of informal or formal rules—are believed to govern acceptable risks (Bicchieri et al., 2018). Several studies have attempted to evaluate the impact norms have on the decision to engage with a smartphones while driving (Atchley et al., 2011; Carter et al., 2014; Merrikhpour & Donmez, 2017; Nelson et al., 2009; Nemme & White, 2010). Atchley et al. (2012) suggested that while younger drivers recognize the risks associated with distracted driving, they perceive it to be a normative behavior. Indeed, studies have shown most young drivers engaging in distracted driving believed their peers and parents—their most important social referents for driving behaviors (Merrikhpour & Donmez, 2017)—participated in distracted driving more frequently than themselves, directly shaping their own understandings of norms around distracted driving (Carter et al., 2014). To demonstrate the impact of this attitude on perceptions of smartphone risk, Atchley et al. (2012) employed attitudes to drink-driving as a baseline comparison. In the research, younger drivers were asked to rate the responsibility of drivers across different crash scenarios and, although participants assigned the same level of culpability to drink-drivers as with texting drivers, they were more likely to fine drink-drivers and give them longer jail sentences (Atchley et al., 2012). Additionally, drivers on a hand-held phone call, despite receiving a higher level of culpability than the drunk driver, escaped punishment altogether, commensurate with a sober driver. These findings suggest that younger adults perceive the risks associated with smartphones as a natural and acceptable component of driving, showing a difference of norms compared with drink-driving being unacceptable after a shift from decades of anti-drink-driving countermeasures and increased enforcement such as randomized breath testing (Davey & Freeman, 2011). Current attitudes toward smartphone use reflect those regarding drink-driving extant in the 1970s, leading Atchley et al. (2012, p. 283) to conclude that, "distracted driving is not a problem of lack of perceived risk, but rather a disconnection between the norms underlying the behavior and knowledge of risks."

It is important to note, however, there has been a general decline in drinking behaviors among younger people in Australia (Australian Institute of Health and Welfare, 2020), including drink driving, while smartphone use and use while driving has increased (Australian Transport Council, 2011). Although the years of negative imagery, public information campaigns, and policy shifts young people have lived through may be driving the decline in drinking for those under 30, recent research suggests this may be unlikely (Foxcroft & Tsertsvadze, 2012). Indeed, shifting cultural trends, especially around social media, may be having a greater impact on attitudes to drinking, with alcohol no longer central to facilitating connection and communication between young people (Livingston & Pennay, 2021). The prevalence of the device have shaped norms supporting use as predominantly social and beneficial, yet recent research demonstrates that the more one engages with their device, the more likely they are to engage in antisocial and risky smartphone behavior such as prohibited, dangerous, or dependent use (Kaviani, Robards, et al., 2020).

The risks associated with smartphone use and driving, however, are dynamic, influencing perceptions around associated norms. Research demonstrates that drivers using their phones engage in a continuous process of self-regulation and risk compensation (Oviedo-Trespalacios et al., 2019; Zhou et al., 2016). By self-regulating their driving and/or smartphone behaviors, drivers believe they can compensate for risks (Zhou et al., 2016) and reduce the likelihood of detection by police (Oviedo-Trespalacios, King, et al., 2017; Oviedo-Trespalacios et al., 2018). Studies have shown the decision to engage in smartphone tasks, and choice of self-regulatory strategy, is conditional and adapted to three factors: driver characteristics, road traffic conditions, and task demands (Hancox et al., 2013; Oviedo-Trespalacios et al., 2018, 2019; Tivesten & Dozza, 2015; Young & Lenné, 2010). Driver characteristics such as crash risk perception (the appetite and interpretation of risk; Oviedo-Trespalacios, King, et al., 2017) and confidence in multi-tasking or driving (Hancox et al., 2013) can determine engagement and choice of behavioral adjustments. Smartphone use is also conditional and adaptive to traffic density, road type, and conditions (Christoph et al., 2019; Tivesten & Dozza, 2015). For instance, drivers may slow down while using a smartphone to reduce the likelihood of crash or severe injury (Oviedo-Trespalacios, Haque, et al., 2017a). Similarly, drivers may choose to only engage with phone tasks while driving on quiet, simple, or familiar roads, or at traffic lights (Oviedo-Trespalacios, Haque, et al., 2017b). Depending on the demands of the driving and smartphone task, a driver may engage or adjust their method of communication (i.e., choose between a text or call; Oviedo-Trespalacios, King, et al., 2017; Tractinsky et al., 2013).

These risk-compensatory strategies undoubtedly shape how young drivers interpret, navigate, and interact with risk, and studies that focus on norms must consider how such behavior can similarly shape attitudes to risk (Atchley et al., 2011). By examining the correlative relationship between self-regulatory behaviors and perceptions of risk and norms, this current research considers these issues. We seek to understand the relationship between socially constructed norms (i.e., perceptions of what other people do, such as friends and family) and individual smartphone use and driving practices, which include a range of risk-compensatory strategies that we explore here.

1.2. The current study

In this study, current attitudes to illegal smartphone use while driving were canvassed, as well as the extent to which drivers reported engaging in self-regulatory behaviors. As with Atchley et al.'s (2012) use of drink-driving norms as a baseline to compare attitudes with smartphone use, our study similarly posited questions comparing the two behaviors to illustrate how current self-regulatory behavior might frame risk and shape attitudes to illegal phone use while driving. We expect the norms that illegal smartphone users perceive in their environment will frame the belief that risk-compensatory strategies such as self-regulating driving behaviors and illegal use are safe and acceptable, ameliorating the seriousness of transgressing smartphone laws.

Knowing how self-regulatory behaviors inform perceptions of danger and attitudes to illegal smartphone use will allow for the development of countermeasures that consider and challenge norms around "acceptable" or "mitigated" risk. Focusing campaigns and interventions on norms without considering the active strategies young drivers employ to use a phone while driving in their perceived "safe way" isolates those drivers most at risk. As we later demonstrate, it is important to acknowledge their adaptive processes in order to speak directly to their lived experience. Addressing this gap in the literature will equip policy makers with a unique vantage point to reduce distraction-related deaths and injuries.

2. Method

2.1. Participants

This study involved four in-depth focus groups with 13 young drivers (18–25 years old) from Victoria, Australia, conducted between May and June 2020. Due to the specific focus on norms around illegal smartphone behavior while driving, anticipating participant responses occurred early during the data collection process. Additionally, stratifying factors such as age, income, and gender were not included in the analysis, therefore the kinds of views did not provide a wide range of features that further focus groups would necessarily capture. Another contributing factor to terminating further data collection was the overwhelming amount of relevant data, something research (Reid & Reid, 2005) suggests is common in the use of online methods. As such, data collection was terminated after the fourth focus group. Smaller group sizes were deemed appropriate as research suggests (Morgan, 1998) that they are more conducive to encouraging personal accounts of an illegal, controversial behavior. This also informed the decision to conduct focus groups, not one-on-one interviews; research suggests smartphone use while driving has become so habitual that drivers are often unaware they are doing it (Hansma et al., 2020; Oulasvirta et al., 2012), therefore, it was hoped the group dynamic would elicit self-reflection and motivate discussion about what is often an unconscious, private behavior.

The focus groups were designed to explore specific quantitative results that emerged from 2,774 responses to a Smartphone Use and Driving Survey (SUDS) the authors conducted between June and August 2019 (Kaviani, Young, et al., 2020b). The 13 young drivers were recruited for focus groups through the initial survey (SUDS). Volunteers were required to have: (a) held a valid Victorian driver license; (b) driven at least once a week in the prior months; (c) have used their smartphones over three hours a day on average; and (d) used a smartphone illegally while driving in the past 31 days of taking the original survey. A $35 gift card was offered to possible focus group volunteers to incentivize participation.

It is important to note how the sample selection criteria may skew the data and affect future reproducibility. Our sample represent high use smartphone owners, which has been shown to increase the likelihood of risk-taking while driving (Kaviani, Young, et al., 2020b). Additionally, nearly all participants are progressing through the Graduated Licensing Program, meaning their inexperience may produce cavalier attitudes (White et al., 2011). These particularities may significantly render the tone of discussion.

2.2. Procedure and materials

Prior to each focus group, prospective participants were emailed an explanatory statement that outlined confidentiality, storage of data, and complaint procedures; a consent form (signed and returned); and a short demographic and smartphone use survey (to confirm they still used their smartphones illegally). There were four focus groups of varying sizes (Table 1), taking approximately one hour each. Due to COVID-19 restrictions in Victoria at the time, focus groups were conducted over Zoom (a video platform) and were moderated by the first author. The discussion was guided by structured questions (included in Appendix A, Supplementary material). The questions explored smartphone use and driving behaviors, legal and non-legal deterrents for illegal use while driving, and justifications and potential countermeasures for illegal use.

2.3. Data analysis

The first author transcribed the data verbatim. Participants were de-identified and prescribed aliases. Constant comparison analysis (Onwuegbuzie et al., 2009) was employed using NVIVO, whereby the data were coded, then those codes were grouped into emergent and a priori categories. From those categories emerged the themes reported on here (Braun & Clarke, 2006).

3. Results and discussion

3.1. Theme A: Participants perceive illegal smartphone use as commonplace, easy, and benign

"They think it's normal, so it is normal."
Grace [P2[2]]

Participants believed that all road users are engaging with smartphones illegally. Additionally, the behavior is perceived as benign and easy to do. The comments below illustrate a view of smartphones as natural, common, and unremarkable fixtures of contemporary life that provide perpetual availability and functionality—even after entering a vehicle:

Stuart [P1]: We have these devices by our sides virtually the whole day, it doesn't seem as if we're doing anything different or explicitly out of the ordinary from our everyday life to be using it while we're driving. It's just the same as what we're doing throughout the rest of our life.
Krista [P2]: The more we're able to do with our phones, the more we never want to put them down. Like, when do we ever go anywhere without our phones nowadays? It's just something constantly attached to us and the more we can do within them then the more we want to use them.

Stuart and Krista's use of 'we' and 'us' demonstrate the universality of their beliefs, which informs norms and provides justification for their own use. Similarly contributing to norms, one's immediate social environment (usually peers and family members) determines the acceptability of certain driving behaviors (Allen et al., 2017). All participants perceived that parents and peers engaged in smartphone use while driving, with several individuals who indicated they had been caught by police.

Derek [F]: It's quite socially acceptable to do it.
Grace [P2]: I've had people send me videos while they're driving. So, I feel like it's just like, they think it's normal, so it is normal.
Liam [P2]: When I was learning to drive, and before that, my par-

[2] Participant license type.

Table 1
Participant license type and phone use.

FG	Name (alias)	License Type	Hours on phone per day	Freq of illegal use	Types of use
1	Stuart	P1**	4–5	Always	Texting (HH); calls (hands-free); music; navigation
	Derek	F	4–5	Half the time	Texting (HH, HF); calls (HF); email; music; internet; navigation; video; other
	Lola	P2	6–7	Always	Texting (HH); calls (HH, HF); social media; email; music; navigation
	Joanne	P1	5–6	ST	Navigation
2	Jody	F	6–7	Half the time	Texting (HH, HF); calls (HF); social media; music; navigation
	Krista	P2	4–5	Always	Navigation; music
	Blake	P1	9+	Always	Texting (HH); calls (HF); social media; music; dating; internet; navigation; other
3	Lula	F	5–6	Always	Calls (HH); social media; internet; navigation
	Judith	P2	4–5	ST	Calls (HF); music
4	Liam	P2	3–4	Always	Texting (HH, HF); calls (HF); social media; music
	Lexie	P2	4–5	Half the time	Calls (HF); music
	Grace	P2	4–5	Most of the time	Calls (HF); music; navigation
	Simsala	L	5–6	Most of the time	Texting (HF); calls (HF); music; navigation

Note: ST = Sometimes; HH = Hand-held; HF = Hands-free; **L = Learner permit; P1 = Probationary P1 license; P2 = Probationary P2 license, F = Full license.

ents would use their phones as well. So that was the role model.

Lula [F]: It's also amongst my peer group, everyone does it. So, I guess it's kind of the norm in a sense, and that I don't know any different... Even my mum, she's usually the one that I am calling when I'm driving

There is a strong desire among young people to adopt similar attitudes to their peers and comply with their expectations and behaviors (Mattern & Neighbors, 2004). Similarly, a parent's strong social ties, authority, and ability to withhold privileges motivate young people to comply with road rules (Allen et al., 2017). Yet, participants in this study believed all road users engage in illegal use and regularly witnessed peers and family members use a smartphone while driving. The combination of these beliefs and observations contribute to the notion that illegal smartphone use is a normal, benign, and unordinary component of driving. This belief was regularly evoked to justify use, contributing to a misperception regarding the acceptability of the behavior.

Indeed, participants routinely thought of their smartphone use as habitual (Oulasvirta et al., 2012), not once alluding to smartphone addiction or dependency as contributing to illegal use. Of course, 'habitual' may tactically lack the stigma or deviance associated with 'addiction;' however, their observations are concurrent with previous research that found merely spending more time on a smartphone per day, even in lieu of psychological dependency, statistically significantly predicts the likelihood of illegal use (Kaviani, Young, et al., 2020b). This is important to consider when shifting norms: any injunctive messaging should be supported by encouraging *positive* smartphone habits (Pinder et al., 2018), both in and outside of the vehicle. This reacquaints the driver with a sense of agency and control over their phone use (Hansma et al., 2020) while combating loss aversion (Delgado et al., 2018) so drivers feel they are gaining—not sacrificing—from their changed behavior.

A successful countermeasure must shift risk perceptions and encourage illegal users to measure the acceptability of their habit against drivers that abstain from use all together. Focusing on how uncommon a behavior is could reduce use among illegal users; however, our participants noted that even if the behavior were uncommon, their illegal use would continue. Additionally, research reveals between 43% (Transport Accident Commission, 2019) to 66.4% (Kaviani, Young, et al., 2020b) of young drivers are engaging in illegal use, meaning the behavior *is* actually common. Therefore, campaigns employing descriptive messaging aimed at shifting norms must also incorporate a strong injunctive message challenging the acceptability of smartphone use (Atchley et al., 2012).

Highlighting the role that smartphone manufacturers have in aiding and normalizing use (Galitz, 2018), participants repeatedly alluded to the devices' ease of use and access. Technological features such as facial recognition to unlock the phone provide quick hands-free access, while familiarity with the phone's application grid and keyboard layout allow for easy engagement. Advancements in mobile technology and accessibility may ameliorate cognitive demand and facilitate surreptitious use; however, research shows that even with eyes on the driving environment cognitive demand from secondary tasks can result in impaired driving performance and increased crash risk (Strayer et al. (2017)). Yet, as demonstrated by Lula's comment, ease of use can facilitate dangerous use by giving the impression that merely having one's eyes on the road is sufficient to driving safely:

Lula [F]: I've gotten to the point where half the time I don't even have to look at my phone because I just have it and I'll know without even looking at it... I know exactly in what place all my apps are, where the buttons are, where everything is... like, my brain is off the road, but my eyes are not.

Even while rationalizing use because of its ease, Lula concedes her 'brain is off the road.' The simplicity with which participants can engage with their smartphones, its perpetual presence, and increasingly diverse functionality, facilitate what many participants defined as 'habitual' use. Our interpretation of their definition aligns with the understanding of habitual to mean use as every day, typical, and without much thought (Hansma et al., 2020). The belief that habits are hard to break was ubiquitously expressed; however, it became evident that ease of use and access and functionality were observations loosely employed to justify use and deflect, or completely abdicate, personal responsibility.

The common remark among learner and P1 and P2 license holders was the device is an essential and harmless tool. The main types of illegal use among these participants were GPS for navigation and Spotify via Bluetooth for music, although several used their phones for calling, texting, or browsing the internet (Table 1). Regarding their reasons for illegal use, all participants provided varied responses. These included wanting to make practical use of time while driving long distances or while stationary or slow-moving in heavy traffic, to listen to one's own music or podcasts rather than the radio, answer calls or respond to texts, or avoid the anxiety of not knowing directions to a particular location (navigation and music were not a reason for illegal use among Full License holders due to the legal status of such use).

The desire to stop illegal smartphone use was palpably absent. Contrary to the impact of descriptive norms, some participants admitted that even if the behavior were uncommon their usage would continue—albeit more surreptitiously—because the behavior was benign. For young people, norms that create external obligations (such as peer social pressure) can strongly impact behavior, even among individuals for whom attitudes to compliance are not aligned with personal attitudes or beliefs. Therefore, parents, peers, and other role models should establish and communicate desired norms to promote and motivate safer use (Buckley et al., 2014). As our research highlights, however, illegal smartphone use is usually committed in the absence of a passenger; therefore, motivating an individual to see the validity of the norm (i.e., abiding by the law) itself would be more effective (Legros & Cislaghi, 2020). On the surface, the lax attitude toward smartphone use while driving and the casual nature with which our participants engage the device may appear wanton at best and reckless at worst. Throughout our discussions, however, it became evident their responses and beliefs were predicated on the assumption that risk had been compensated for before choosing to engage. Importantly, for a driver to internalize obligations, the efficacy and acceptability of self-regulatory strategies must be challenged in order to speak to their lived experience of engaging with risk-compensatory behaviors.

3.2. Theme B: Self-regulatory behaviors are pervasive among illegal smartphone users

> *"You draw your own boundaries and make your own rules"*
> *—Lexie [P2]*

This study provides further evidence that, by adjusting behavior accordingly, drivers felt they could compensate for, or ameliorate, negative effects of smartphone impairment on safety (Zhou et al., 2016) and reduce the likelihood of police detection (Oviedo-Trespalacios et al., 2018). As illegal users perceive the behavior as normative, drivers use inaccurate social comparison (Mattern & Neighbors, 2004) information to justify the risks associated with self-regulated engagement (i.e., they believe their behavior is safe and conscientious, especially when compared to drivers that did not effectively regulate their smartphone use and driving). These strategies were presented as a personal set of rules for which transgressing would be irresponsible, selfish, and dangerous. In their opinion, drivers that failed to compensate for risk were the problem—not smartphone use. Instead of comparing their behavior with law abiders, this misperception informs a certain level of risk-taking as socially acceptable, the prevalence of which has led to the normalization of risks associated with self-regulatory behaviors. This may explain why perceptions of risk associated with smartphone use—although almost as high as drink-driving—are declining (Transport Accident Commission, 2018).

All three conditional factors mentioned in the introduction—driver characteristics (confidence and risk perception), road traffic conditions, and secondary task demands—were pertinent throughout the focus groups and are discussed below.

3.2.1. Driver characteristics (risk perception)

All participants were aware of distraction risks involving smartphone use; however, as the following exchange demonstrates, they attributed the problem to driver incompetence, not smartphone use:

> *Blake [P1]: I think there are bad drivers everywhere. And I would worry just as much about a really bad driver, not on their phone, just driving. It's more driving quality than phone use that worries me.*
>
> *Jody [F]: Yeah. If someone's using their phone but they're managing to stay within their lines and not swerving, they're keeping a straight line, or they're turning whenever they need to, but they're still maintaining as if they weren't on their phone, that's okay. Cause I'm kinda like 'look, they're much less likely to cause anything compared to the ones that are swerving everywhere, they make me uncomfortable'. People that managed to like, you can't tell that they're using their phone, I'm like, 'look, they've got it figured out'.*
>
> *Blake [P1]: Yeah. You gotta more judge the driving than you do what's caused the driving.*

Although aware smartphones can be distracting, their own use is unlikely to change while successfully complying with their own self-regulatory/risk-compensatory rules. Their comments may reflect an optimism bias (White et al., 2011), in addition to an unwillingness to admit their risky driving could also be harmful. Further, it illustrates a high opinion of one's own driving. This is consistent with research that demonstrates most drivers consider themselves more skillful and less risky than other road users (Svenson, 1981). A telling quote from Blake neatly illustrates this point: "*just because other people can't use it right, doesn't mean I can't use it right.*" The problem, then, becomes *other* road users. This distinction is further borne out when participants discussed being a passenger of an offending driver. Lula and Krista, both having admitted to "always" using their smartphones illegally while driving, had no difficulty perceiving the risks associated with another driver's illegal use:

> *Lula [F]: But yeah, if I watch someone do it and I'm in the car I'm like, 'what are you doing? Put it down, you're not watching the road, you're not paying attention, you're being distracted'.*
>
> *Krista [P2]: I personally am not comfortable when I see somebody who is on their phone because I think, 'you might be all right for this second, what if you hit an animal? I feel like that's very contradictory considering what I've said about my own phone use.*

Thus, their own use is not a problem of perceived risk but, rather, a disconnection between norms and risk (Atchley et al., 2012). Or, as we argue, a connection between misperceived norms and the normalization of risk. Shifting perceptions of norms requires challenging risk-compensatory strategies. If young, inexperienced drivers are adapting their smartphone and driving to mitigate risk, then the dangers associated with those strategies need to be demonstrated. Throughout the Victorian Graduated Licensing System, license holders could be regularly exposed to the laws, risks, and consequences of smartphone use through regular communications. License tests could include hazard questions or simulated smartphone use in hazards training (Chan et al., 2010). In addition to top-down dissemination of information, young drivers should be encouraged and rewarded when sharing educational resources with peers and family through social media, a process that can become attractive or fashionable if a celebrity is marketed as a compliant role model (Lapinski & Rimal, 2005).

3.2.2. Driver characteristics (confidence)

Participants frequently qualified smartphone engagement by their level of confidence in multitasking, driving safely, and avoiding apprehension. Whether they could accomplish the task safely, and without being caught, determined the type and likelihood of engagement.

Simsala, a learner driver that uses her phone for GPS and music, noted she limits her texting and calls to hands-free (which is still illegal for learner drivers in Victoria); however, it is confidence—not the law—that circumscribes the behavior:

> *Simsala [L]: I'll be like, okay, definitely I cannot use my phone on the road because I am not confident in driving because I'm an L and I'm still under supervision.*

Confidence also mediated the more advanced license holder's engagement. Joanne, a P1 license holder, reported she limits use to GPS voice function; however, other more confident participants in her focus group (a P1, P2, and full license holder respectively) frequently engaged with a variety of smartphone applications.

While all uses are illegal under Victorian law for the probationary license holders, there was clearly a hierarchy of risk-taking in how these participants understood their practices, modulated by a sense of self-confidence. Confidence, however, can give the impression of competence (Lesch & Hancock, 2004), which can result in dangerous scenarios, as Derek explains:

Derek [F]: I definitely feel like I'm a confident driver, but there's certainly times where I, you know, sort of get a bit of a shock when I sort of swerve into a different lane or something.

Throughout discussions participants noted that dangerous experiences of smartphone distraction can affect their confidence; however, their ability to regain confidence was commonly alluded to.

Lula [F]: I think if I had something to shake my confidence, like an experience or a near miss or something like that, it'd be enough to definitely shake my confidence. And then, you know, probably for however long it took me to move on from that I'd like to think that I wouldn't go anywhere near my phone whilst I'm driving. But once I learned to deal with that and my confidence built back up, then I'd probably fall back into the same old habit.

Confidence in avoiding apprehension was similarly checked when confronted with a "close call" or actual police apprehension, the effects of which are also ephemeral:

Liam [P2]: When it happens [getting caught by police], you'll stop doing it. 'Cause like I've lost my license from speeding, so I didn't speed after that for a while. And then it kind of just builds up again and then I think you'd get a bit more confident and start doing it again.

Lula [F]: If I've just been on [my smartphone] doing something I know I shouldn't be, and I come over the top of the hill and I'm like, 'ugh, that's a police officer', you know my heart will race and I'll get this pit in my stomach being like, 'Oh God, if I had been two seconds too late I would have got caught.' But then five minutes later I'll do it again.

The participants went further to admit that even after a near crash or close call with police—or after actual apprehension—their confidence will return. To counter recidivism, drivers could be reminded of past transgressions via electronic or paper communication. These could include a compulsory test or interactive information session to remind drivers of smartphone law, the dangers associated with all forms of use, and practical personal interventions for breaking habits and reducing use.

Their level of engagement was also predicated upon their confidence in multitasking. As Lola revealed earlier and elucidates upon below, smartphones are increasingly thought of as easy to use in the context of driving; confidence with the device can reduce the sense of cognitive demand and increase engagement with the device:

Lola [P2]: I mean, if I don't feel confident I won't, but I have texted while driving…but I'm also familiar with like the layout on my phone, so I don't always rely on vision, I just kind of go by feel and autocorrect.

The cognitive and physical demands are such that engagement is unlikely if they feel incapable. Additionally, drivers may allocate different smartphone tasks for specific driving conditions such as when stopped at an intersection (as discussed in sections 3.2.3 and 3.2.4). However, it is important to understand that smartphone use and driving compete for the driver's visual, physical, and cognitive attention (Oviedo-Trespalacios et al., 2016). This means that when resources are allocated to one task, the resources required for the other tasks are not available and therefore results in impaired performance (Strayer & Johnston, 2001). Our research suggests a driver's confidence in their multi-tasking and avoiding police apprehension moderates the likelihood they will engage in risk compensatory behaviors. Unfortunately, the longer a driver gets away with breaking the law (Stafford & Warr, 1993) and avoids crashing the more confident they will be in continuing their phone use while driving (Oviedo-Trespalacios, 2018). Therefore, messaging targeting the perceived acceptability and efficacy of multitasking may have an impact (Sherman et al., 1997).

For optimal efficacy, any message-based approach should be supported by increasing deterrent mechanisms, such as enforcement or fines, or regular and on-going educational programs defined by a multi-action and interactive approach. For instance, advertising could show mourners engaging with their smartphones during a somber eulogy, the memorial photo then opening to a scene where the victim uses a smartphone before crashing, the takeaway being, "*respect the road if you respect your life*." Similarly, strategies to end illegal self-regulatory smartphone use must shift perceptions among illegal users that "out of sight" use, including Bluetooth or hands-free, reduces the seriousness of the offense or certainty of apprehension.

3.2.3. Road traffic conditions

Aligning with previous research, participants noted they were more likely to engage in smartphone use in familiar environments, on quiet roads, or while stopped at traffic lights (Tivesten & Dozza, 2015). These scenarios were considered less dangerous or likely for apprehension:

Blake [P1]: You know, if you're alone on the road and there's no one near you, no one's going to see you. It's pretty low risk in your mind. That's one scenario which you wouldn't even think about it.
Facilitator: So, managing the risk?
Blake [P1]: Yeah. Not only the safety risk but the risk of getting caught.

Indeed, waiting at an intersection was viewed as an innocuous and acceptable moment to engage with smartphones. Alternatively, participants reported avoiding use in situations that required greater concentration such as near schools, at high speeds, in the city, or on windy roads (Oviedo-Trespalacios, Haque, et al., 2017a; Young & Lenné, 2010). The following exchange between Stuart and Lola, both heavy users while driving, demonstrate this:

Stuart [P1]: There are times when it seems as if you're more safe to use your phone when you're driving. Like if you were stopped at a traffic light for instance, I would say that it's not controversial that there's far less risk in doing that and using a phone there than if you are driving along a country road at faster speeds—that's a very different circumstance.
Lola [P2]: Yeah. Cause if you're already stopped, you'd think that there'd be less risk. You might just piss someone off behind you cause you're not going at the green light. But if you're moving and you're distracted, I think there's a much higher risk of you coming into oncoming traffic or something.

Additionally, a frequent consensus among all participants was the decision not to use a smartphone when driving with passengers. This highlighted an interesting and previously researched point where one's appetite for risk was almost effaced in consideration of their passenger's safety (Huemer et al., 2018). A passenger renders the fear of hurting others salient (Kaviani, Young, et al., 2020a); however, participants mitigated concern for the safety of other road users by compensating for risk:

Blake [P1]: I don't do it if I'm driving my friends or family. That's not an embarrassment thing. I think that's more of like a safety *thing. Like you're more accountable for other people when you're driving people.*

Lula [F]: I kind of, it sounds horrible, I justify it in the sense that if I'm the only one in my car, that's one less person that I'm putting at risk. I know that everyone else on the road that I'm passing at the same time, I'm putting them in just as much risk. But… if someone else is in my car, then I won't even look at it, won't even touch it because I'm responsible for those people.

Lula's social connection and physical proximity to her passengers increases the sense of responsibility for their safety, whereas the risks to road users beyond her vehicle are de-personalized and less salient. Having a passenger may also reduce the desire for using a phone to facilitate social connection, while navigation, music, or other phone tasks can be fulfilled by the passenger allowing the driver to concentrate on the road.

The need for constant connection and maintaining a sense of belonging means young people are increasingly unable to sit with their own company without turning to their smartphones. Merely applying fines and demerit points to admonish illegal use fails to address the psychosocial attraction and impact smartphones have on behavior, especially with regard to habitual and dangerous use (Kaviani, Robards, et al., 2020). Treating illegal smartphone use as a health issue and not just road safety concern addresses research that shows more time spent on smartphones per day greatly increases problematic dangerous, prohibited, and dependent use (Kaviani, Robards, et al., 2020).

3.2.4. Secondary (smartphone) task demand

As previous research highlights, young drivers are using their smartphones for various tasks while driving (George et al., 2018). Our findings reveal how certain tasks are dictated, and dictate, types of driving behaviors; participants employed specific compensatory strategies to mitigate cognitive demand and risk of apprehension. Allowing only a brief glance at the phone's interface was a typical refrain from participants, showing they understand the safety risks associated with long periods of eyes off the road (Simons-Morton et al., 2014), while techniques such as using hands-free, restricting the types of applications they engage with or the amount of time they engage visually, helped to give a sense of concealment (Gauld et al., 2014).

Jody [F]: Usually I try and avoid using it outside of maps and music or phone calls. So, when I do, I'll be like extra vigilant, looking all around me and just sort of got the adrenaline going with it.

Lola [P2]: I might just look away from the road to unlock my phone or something and then just kind of skip a song or something. I don't feel like I'm looking away from the road for that long. And so my judgment's not that impaired. Like it's wrong, but it's not that wrong.

Derek [F]: You kind of pick your moment to use your phone when you don't think there's police around or, you know, when you think there's less danger.

Participants incorrectly believed (Oviedo-Trespalacios et al., 2018) using applications such as GPS or music to be low risk and therefore justified, comparing the behavior to just "*touching parts of your car.*"

Interventions could address norms (to promote a message that most drivers do not use a phone while driving), as well as provide psychological insight into the ineffectiveness of multitasking and other deleterious effects of smartphones on cognition and well-being.

3.3. Theme C: Risk-compensation strategies rationalize risks and perceived norms, reducing the seriousness of transgressions when compared with drink-driving

"I don't know if it's real, but it feels like the risk [when drink-driving] is a lot more real"

—Lola [P2]

Research into distinguishing between the influence of norms and risks associated with distracted driving has highlighted the usefulness of comparing the behavior with another road safety issue that young drivers find generally unacceptable (i.e., drunk driving; Atchley et al., 2012). Atchley et al.'s (2012) study revealed that while drivers recognized the risks of smartphone use as equal to drunk driving, their perceptions of the behavior as normative reduced the seriousness of transgressions. In an attempt to understand the factors influencing that perception, our study initiated a discussion comparing attitudes toward drink-driving and smartphone use.

The young participants emphatically agreed that drink-driving is unacceptable, yet the comparison with smartphone use had an extraordinary impact as participants attempted to articulate justifications for engaging in illegal phone use while simultaneously acknowledging the risks of both behaviors. At times, participants struggled to rationalize why smartphone use was more acceptable, the palpable cognitive dissonance highlighting the impact norms have on acceptable levels of risk. For instance, after participants acknowledged the similarity of risks but down-played smartphone severity due to its prevalence, the facilitator asked how justifiable drink-driving would be if it too were prevalent:

Krista [P2]: You've completely "dad moved us" here [a colloquial term for reductio ad absurdum].

Jody [F]: Because it [drink-driving] does seem like a much more serious thing than phone usage and because the phone usage also has the varying degrees [regulation] and all that, I still feel like it's very different.

Here, the availability of self-regulatory risk-compensation strategies reduced the seriousness of transgressions and normalized the prevalence of use when compared with drink-driving. Indeed, it became apparent that each aspect informing why drink-driving was socially unacceptable was related to the driver's inability to self-regulate their inebriation. By having access to risk-compensatory strategies the perceived severity and negative consequences of the behavior were reduced and normalized:

Krista [P2]: There's stuff that you can't mitigate [when driving drunk], it's risk you can't mitigate. You cannot have a blood alcohol test and be like, okay, I'm at this whatever reading and that means that these things are what's inhibited now—I can control that now. Whereas you can be like, I need to put the phone down, I need to stop, and you can continue on driving perfectly normal. You can't do that with drinking. You're inebriated, you're affected, doesn't matter whether you want to or not, if you sit behind that wheel and you've had a drink, then that's it. Game over essentially, there's nothing else you can do apart from not drive and eliminate the risk completely.

For participants, the inability to regulate drink-driving meant the risks and performance decrements "felt" more real to them when comparing the behavior to smartphone use. Being affected the "entire trip," as opposed to "just some points," rendered the dangers, risks, and consequences more immediate and impactful, and therefore the decision to drink-drive loftier. Also shaping attitudes to smartphone use was the belief that, when compared with drink-driving, it is a victimless crime:

Blake [P1]: Like, there's no headline that's going to be 'phone user killed someone in a crash.' Cause you can't verify that, and that's probably a reason that the reporting is so different.
Jody [F]: It's not as serious, like a lot of the consequences aren't necessarily as serious as the drink-driving ones... in my mind it pretty much seems like the consequence of driving and using your phone is you get caught and you get fined demerit points.

Supporting the notion that drink-driving is more unacceptable than smartphone use was the severity of punishment. It was widely known that if caught drink-driving the punishment requires attendance at a behavior change program and the installment of an alcohol interlock device, which, for Liam, sounded like a "*really big hassle*."

As such, punishment could include a compulsory Behavioral Change Program (Buckley et al., 2014) such as current programs—which Liam noted as a "*really big hassle*" worth avoiding—aimed at reducing re-offending among drink-drivers (VicRoads, 2020). These programs could incorporate learnings around phone dependency (Kaviani, Young, et al., 2020b) and offer solutions for managing risk (Oviedo-Trespalacios et al., 2020).

All participants agreed media campaigns have raised awareness of the risks and humanized the consequences of drink-driving, but they felt the public safety message around smartphone use—besides being sparse and tepid in comparison to drink-driving—did not consider their successful risk-compensation strategies.

Krista [P2]: I feel like there is an ad at the moment. I don't know if you guys have seen it, it's like 'a glance at your phone is like this and there's someone sitting in the backseat covering the driver's eyes'. I feel like that's not serious enough though, do you know what I mean?
Jody [F]: Yeah...there was one ages ago that they had where every time they looked at their phone everything would just go black, for however long it was. And then the child going across the school crossing something.
Krista [P2]: It doesn't sound as graphic as like drink-driving.
Jody [F]: Yeah. It's nowhere near as graphic. And the other thing, if I was going anywhere near a school crossing and all that it wouldn't be an area that I would use my phone because that's high risk, so it still doesn't really relate to me in any way.
Blake [P1]: To be honest, and this is even more morbid, I don't think there's anything that would change the way that I use my phone.
Jody [F]: Yeah look me neither.
Blake [P1]: I couldn't think of anything that would, you know, change the way I do it, because the way I do it now is the way that I've consciously thought mitigates the most risk. And even though I know it is dangerous and you shouldn't do, I still think I would.

The candid manner in which Blake and Jody confess their unyielding illegal phone use illustrates the impact self-regulating risk has on their perceptions of the behavior and internal moral compass. The scenario presenting an individual unsuccessfully compensating for risk (driving near a school) becomes the baseline for comparison, illustrating the relationship between misperceived norms and the normalization of risk.

When compared with smartphone use, young drivers perceived the risks, dangers, and victims associated with drink-driving as more "*real*" and "*serious*" because the behavior could not be regulated. As such, because they were regulating their smartphone use while driving and compensating for perceived risks, participants did not believe they were engaging in a dangerous activity. Shifting norms around smartphone use requires addressing these misperceptions. In doing so, stigma around smartphone use may increase, shifting attitudes toward the behavior. Smartphone use and driving campaigns may benefit from adopting explicit messaging around dangers of use and victims of the crime by associating illegal use with anti-social outcomes such as social ostracization due to injuring or killing others, imprisonment, or shame. This may increase the seriousness of the offense. A focus on juxtaposing identical crash scenarios where one is caused by self-regulatory smartphone use and the other by drink-driving could address the paradox of "acceptable risk" while challenging the belief that a driver need only be sober to be safe and "celebrated."

Although recent evidence evaluating the impact of a fear-based approach found positively framed films decreased self-reported risky-driving significantly greater than films appealing to fear (Cutello et al., n.d.), this study demonstrates young drivers have not been exposed to any narratives convincing them that their smartphone use could result in serious consequences. Although a fear-based approach may not instantly change behavior, fear within the context of narrative persuasion has been shown to increase stigma around texting while driving (Tamul et al., 2021). Providing more media coverage of smartphone caused fatalities can increase social pressure and decrease the likelihood a driver will continue to illegally engage with their smartphones.

Conversely, likely a testament to the inculcating effect decades of anti-drink-driving messaging has had on shaping public and personal sentiment, there exists a strong moral code against drink-driving. This has increased stigma and reduced any expected utility to getting into a vehicle intoxicated.

Jody [F]: Drink-driving is something that people will feel very strongly on. Whereas using your phone not necessarily.

Lola [P2]: I think drink-driving has a much bigger stigma. Like, not just in public but amongst my friends as well... It's celebrated if you're the designated driver, but it's not if you're abstaining from phone use.

Lola's comment, besides highlighting how being a designated driver is considered a venerable designation, also demonstrates how drink-driving is perceived as a choice. Participants viewed the behavior as pre-meditated, while smartphone use was more a response or impulse; as Blake extenuates, a drunk driver enters their vehicle knowing they are about to drink-drive, whereas, "*no one says I'm going to drive and then go on my phone.*" Framing smartphone use this way has the effect of rendering the driver passive to their device and likely less culpable if breaking the law.

The attitudes and opinions of participants corroborate the impact of decades of anti-drink-driving messaging on social norms (Atchley et al., 2012). Advertising and media campaigns, alongside other multifaceted countermeasures such as proper enforcement, community based programs and education programs, have manufactured moral outrage at the behavior, making it synonymous with idiocy in the form of the pithy slogan, "*if you drink, then drive, you're a bloody idiot*" (Transport Accident Commission, 2020). The slogan has become a catchphrase admonishing and reminding drivers of the unacceptability and stigma associated with drink-driving.

As with any media messaging, it is important to investigate the target population's response; however, leveraging upon the success of TAC's drink-driving messaging could expedite attitudes and norms around smartphone use. While we acknowledge the parallels between drink-driving and smartphone use while driving are not exact, and are shaped differently by cultural factors, we *can* learn from the campaigns against drink-driving. For instance, the ease and candor with which our participants evoked justifications for illegal phone use highlight an absence of stigma associated with admitting to engaging in self-regulatory risk-compensatory techniques. Messaging must challenge the acceptability of "mitigated" risk within the parameters of an illegal behavior, yet participants admitted nothing would impact their use because they can do it safely—save "*pissing someone off*" at traffic lights. Additionally, they routinely extolled their aptitude in risk-compensation strategies, shifted the accountability of negative consequences onto inept drivers, admitted their confidence can easily rejuvenate, and preference their own set of rules over the law. A message needs to isolate this behavior as both wrong and anomalous. For instance, "*if you drive for your phone, you're bloody selfish,*" or "*if your phone drives you, you're bloody selfish,*" capture how self-regulatory behavior centers the phone over the safety of all road users, suggesting the behavior is anti-social and unacceptable. Imbuing the behavior with negative social value could increase guilt or shame, which, aside from being an effective form of messaging to change behavior (Duhachek et al., 2012), has been shown to predict the likelihood of engaging in illegal use (Kaviani, Young, et al., 2020a). An educational campaign that highlights and elucidates on the risks of these self-regulatory behaviors by emphasizing the need to drive for the road, not phone, can directly address those most at risk of perceiving the behavior as normal. Additionally, communication could be presented at hot spots such as traffic lights or on highway boards during rush hour while traffic is slow, as long as it is tested to ensure additional driver distraction is not created. It is important to note that PR messaging is one element of what should be a multifaceted approach to reducing illegal use.

Finally, highlighting that smartphone use is a choice as much as drink-driving may place the onus of responsibility back on drivers. Supplementary behaviors or alternatives need to be presented in a manner that acknowledges smartphone use can be a habit and that drivers are engaging in self-regulatory strategies they believe compensates for risk. Avoiding all forms of illegal use should be presented as the most effective way to compensate for risks to oneself and others. For example, as Simsala [L] noted:

> *I always make sure that I've queued up enough [music] so it'll go for the ride… and I'll text the person, 'I'm just leaving my house now should be 30 minutes depending on traffic'… then I don't have to send that quick message in traffic if I'm running late.*

By planning ahead Simsala was able to engage in *lawful* risk-compensatory behaviors.

4. Limitations and further research

As with any research, this study's limitations should be noted. Focus groups present both advantages and challenges in the collection of data. Focus groups occurred via online videoconferencing platform Zoom, which may have predisposed some to either temper their views or inflate their experiences to match the group dynamic or present a more extreme, individualistic perspective, although this can also be a limitation in face-to-face focus groups (Boateng, 2012). From the facilitator's point of view, however, participants were extremely candid and at ease during discussions, as the data likely demonstrates.

As mentioned in the introduction, this study overlooked stratifying factors in the exploration of norms, a limitation that if addressed in future research may offer a more nuanced perspective. In terms of peer-pressure and risk-taking around smartphone use while driving, it would be valuable to better understand the gendered dimensions of these processes, and to interrogate potential links between masculinity and risk-taking, if such dynamics do indeed contribute to the behavior. Additionally, at risk of oversimplifying norms around smartphone use, external factors that may impact the likelihood one would engage with their device were deliberately omitted. For instance, the pressure to respond and remain perpetually available to peers via smartphones has been shown to influence use while driving (Seiler, 2015; Seiler & Kidwell, 2016; Seiler & Kirby, 2017), however, this study was primarily focused on norms around risk-taking, not norms governing communication and social connection.

Further research may benefit from evaluating the perceptions of illegal smartphone use among young drivers that abide by the law. Particularly, it would be interesting to explore the reasons and techniques for avoiding illegal use, and if they can be leveraged to dissuade all drivers from engaging in the behavior.

Conflicts of interest

We have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2021.06.010.

References

Allen, S., Murphy, K., & Bates, L. (2017). What drives compliance? The effect of deterrence and shame emotions on young drivers' compliance with road laws. *Policing and Society, 27*(8), 884–898. https://doi.org/10.1080/10439463.2015.1115502.

Atchley, P., Atwood, S., & Boulton, A. (2011). The choice to text and drive in younger drivers: Behavior may shape attitude. *Accident Analysis & Prevention, 43*(1), 134–142. https://doi.org/10.1016/j.aap.2010.08.003.

Atchley, P., Hadlock, C., & Lane, S. (2012). Stuck in the 70s: The role of social norms in distracted driving. *Accident Analysis & Prevention, 48*, 279–284. https://doi.org/10.1016/j.aap.2012.01.026.

Australian Institute of Health and Welfare. (2020). National Drug Strategy Household Survey 2019. 10.25816/E42P-A447.

Australian Transport Council (2011). *National road safety strategy 2011-2020*. Australian Transport Council.

Bicchieri, C., Muldoon, R., & Sontuoso, A. (2018). Social norms. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2018)*. Metaphysics Research Lab: Stanford University.

Boateng, W. (2012). Evaluating the efficacy of focus group discussion (FGD) in qualitative. *Social Research, 3*(7), 4.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.

Buckley, L., Chapman, R. L., & Sheehan, M. (2014). Young driver distraction: State of the evidence and directions for behavior change programs. *Journal of Adolescent Health, 54*(5), S16–S21. https://doi.org/10.1016/j.jadohealth.2013.12.021.

Carter, P. M., Bingham, C. R., Zakrajsek, J. S., Shope, J. T., & Sayer, T. B. (2014). Social norms and risk perception: Predictors of distracted driving behavior among novice adolescent drivers. *Journal of Adolescent Health, 54*(5, Supplement), S32–S41. https://doi.org/10.1016/j.jadohealth.2014.01.008.

Chan, E., Pradhan, A. K., Pollatsek, A., Knodler, M. A., & Fisher, D. L. (2010). Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills. *Transportation Research. Part F, Traffic Psychology and Behavior, 13*(5), 343–353. https://doi.org/10.1016/j.trf.2010.04.001.

Christoph, M., Wesseling, S., & van Nes, N. (2019). Self-regulation of drivers' mobile phone use: The influence of driving context. *Transportation Research Part F: Traffic Psychology and Behavior, 66*, 262–272. https://doi.org/10.1016/j.trf.2019.09.012.

Cutello, C. A., Gummerum, M., Hanoch, Y., & Hellier, E. (n.d.). Evaluating an intervention to reduce risky driving behaviors: Taking the fear out of virtual reality. Risk Analysis, n/a(n/a). 10.1111/risa.13643.

Davey, J. D., & Freeman, J. E. (2011). Improving Road Safety through Deterrence-Based Initiatives. *SQU Med J, 11*(1), 29–37.

Delgado, M. K., McDonald, C. C., Winston, F. K., Halpern, S. D., Buttenheim, A. M., Setubal, C., ... Lee, Y.-C. (2018). Attitudes on technological, social, and behavioral economic strategies to reduce cellphone use among teens while driving. *Traffic Injury Prevention, 19*(6), 569–576. https://doi.org/10.1080/15389588.2018.1458100.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629–636. https://doi.org/10.1037/h0046408.

Duhachek, A., Agrawal, N., & Han, D. (2012). Guilt versus Shame: Coping, Fluency, and Framing in the Effectiveness of Responsible Drinking Messages. *Journal of Marketing Research, 49*(6), 928–941. https://doi.org/10.1509/jmr.10.0244.

Foxcroft, D. R., & Tsertsvadze, A. (2012). Cochrane Review: Universal school-based prevention programs for alcohol misuse in young people. *Evidence-Based Child Health: A Cochrane Review Journal, 7*(2), 450–575. https://doi.org/10.1002/ebch.v7.210.1002/ebch.1829.

Galitz, S. (2018). Killer Cell Phones and Complacent Companies: How Apple Fails to Cure Distracted Driving Fatalities. *University of Miami Law Review, 72*(3).

Gauld, C. S., Lewis, I., & White, K. M. (2014). Concealing their communication: Exploring psychosocial predictors of young drivers' intentions and engagement in concealed texting. *Accident Analysis & Prevention, 62*, 285–293. https://doi.org/10.1016/j.aap.2013.10.016.

George, A. M., Brown, P. M., Scholz, B., Scott-Parker, B., & Rickwood, D. (2018). "I need to skip a song because it sucks": Exploring mobile phone use while driving among young adults. *Transportation Research Part F: Traffic Psychology and Behavior, 58*, 382–391. https://doi.org/10.1016/j.trf.2018.06.014.

Hancox, G., Richardson, J., & Morris, A. (2013). Drivers' willingness to engage with their mobile phone: The influence of phone function and road demand. *IET Intelligent Transport Systems, 7*(2), 215–222. https://doi.org/10.1049/iet-its.2012.0133.

Hansma, B. J., Marulanda, S., Chen, H.-Y., & Donmez, B. (2020). Role of habits in cell phone-related driver distractions. *Transportation Research Record*. https://doi.org/10.1177/0361198120953157.

Huemer, A. K., Schumacher, M., Mennecke, M., & Vollrath, M. (2018). Systematic review of observational studies on secondary task engagement while driving. *Accident Analysis & Prevention, 119*, 225–236. https://doi.org/10.1016/j.aap.2018.07.017.

Kaviani, F., Robards, B., Young, K. L., & Koppel, S. (2020). Nomophobia: Is the fear of being without a smartphone associated with problematic use? *International Journal of Environmental Research and Public Health, 17*(17), 6024. https://doi.org/10.3390/ijerph17176024.

Kaviani, F., Young, K. L., Robards, B., & Koppel, S. (2020a). Understanding the deterrent impact formal and informal sanctions have on illegal smartphone use while driving. *Accident Analysis & Prevention, 145*, 105706. https://doi.org/10.1016/j.aap.2020.105706.

Kaviani, F., Young, K. L., Robards, B., & Koppel, S. (2020b). Nomophobia and self-reported smartphone use while driving: An investigation into whether nomophobia can increase the likelihood of illegal smartphone use while driving. *Transportation Research Part F: Traffic Psychology and Behavior, 74*, 212–224. https://doi.org/10.1016/j.trf.2020.08.024.

Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory, 15*(2), 127–147. https://doi.org/10.1111/j.1468-2885.2005.tb00329.x.

Legros, S., & Cislaghi, B. (2020). Mapping the social-norms literature: An overview of reviews. *Perspectives on Psychological Science, 15*(1), 62–80. https://doi.org/10.1177/1745691619866455.

Lesch, M. F., & Hancock, P. A. (2004). Driving performance during concurrent cell-phone use: Are drivers aware of their performance decrements? *Accident Analysis & Prevention, 36*(3), 471–480. https://doi.org/10.1016/S0001-4575(03)00042-3.

Livingston, M., & Pennay, A. (2021). Don't believe the hype, teens are drinking less than they used to. *National Drug & Alcohol Research Centre* https://ndarc.med.unsw.edu.au/blog/dont-believe-hype-teens-are-drinking-less-they-used.

Mattern, J. L., & Neighbors, C. (2004). Social norms campaigns: Examining the relationship between changes in perceived norms and changes in drinking levels. *Journal of Studies on Alcohol, 65*(4). Gale Academic OneFile. http://link.gale.com/apps/doc/A121401894/AONE?u=monash&sid=zotero&xid=96608b25.

Merrikhpour, M., & Donmez, B. (2017). Designing feedback to mitigate teen distracted driving: A social norms approach. *Accident Analysis & Prevention, 104*, 185–194. https://doi.org/10.1016/j.aap.2017.04.016.

Morgan, D. L. (1998). *Planning focus groups*. Sage.

Nelson, E., Atchley, P., & Little, T. D. (2009). The effects of perception of risk and importance of answering and initiating a cellular phone call while driving. *Accident Analysis & Prevention, 41*(3), 438–444. https://doi.org/10.1016/j.aap.2009.01.006.

Nemme, H. E., & White, K. M. (2010). Texting while driving: Psychosocial influences on young people's texting intentions and behavior. *Accident Analysis & Prevention, 42*(4), 1257–1265. https://doi.org/10.1016/j.aap.2010.01.019.

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A qualitative framework for collecting and analyzing data in focus group research. *International Journal of Qualitative Methods, 8*(3), 1–21. https://doi.org/10.1177/160940690900800301.

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing, 16*(1), 105–114. https://doi.org/10.1007/s00779-011-0412-2.

Oviedo-Trespalacios, O. (2018). Getting away with texting: Behavioral adaptation of drivers engaging in visual-manual tasks while driving. *Transportation Research Part A: Policy and Practice, 116*, 112–121. https://doi.org/10.1016/j.tra.2018.05.006.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Demmel, S. (2018). Driving behavior while self-regulating mobile phone interactions: A human-machine system approach. *Accident Analysis & Prevention, 118*, 253–262. https://doi.org/10.1016/j.aap.2018.03.020.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2017). Self-regulation of driving speed among distracted drivers: An application of driver behavioral adaptation theory. *Traffic Injury Prevention, 18*(6), 599–605. https://doi.org/10.1080/15389588.2017.1278628.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2016). Understanding the impacts of mobile phone distraction on driving performance: A systematic review. *Transportation Research Part C: Emerging Technologies, 72*, 360–380. https://doi.org/10.1016/j.trc.2016.10.006.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2017). Effects of road infrastructure and traffic complexity in speed adaptation behavior of distracted drivers. *Accident Analysis & Prevention, 101*, 67–77. https://doi.org/10.1016/j.aap.2017.01.018.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2019). "Mate! I'm running 10 min late": An investigation into the self-regulation of mobile phone tasks while driving. *Accident Analysis & Prevention, 122*, 134–142. https://doi.org/10.1016/j.aap.2018.09.020.

Oviedo-Trespalacios, O., King, M., Haque, M. M., Washington, S., & Podobnik, B. (2017). Risk factors of mobile phone use while driving in Queensland: Prevalence, attitudes, crash risk perception, and task-management strategies. *PLOS ONE, 12*(9), e0183361. https://doi.org/10.1371/journal.pone.0183361.

Oviedo-Trespalacios, O., Vaezipour, A., Truelove, V., Kaye, S.-A., & King, M. (2020). "They would call me, and I would need to know because it is like life and death": A qualitative examination of the acceptability of smartphone applications designed to reduce mobile phone use while driving. *Transportation Research Part F: Traffic Psychology and Behavior, 73*, 499–513. https://doi.org/10.1016/j.trf.2020.06.007.

Pinder, C., Vermeulen, J., Cowan, B. R., & Beale, R. (2018). Digital behavior change interventions to break and form habits. ACM Transactions on Computer-Human Interaction, 25(3), 15:1-15:66. 10.1145/3196830.

Regev, S., Rolison, J. J., & Moutari, S. (2018). Crash risk by driver age, gender, and time of day using a new exposure methodology. *Journal of Safety Research, 66*, 131–140. https://doi.org/10.1016/j.jsr.2018.07.002.

Reid, D. J., & Reid, F. J. M. (2005). Online focus groups: An in-depth comparison of computer-mediated and conventional focus group discussions. *International Journal of Market Research, 47*(2), 131–162. https://doi.org/10.1177/147078530504700204.

Scott-Parker, B., Watson, B., King, M. J., & Hyde, M. K. (2012). Young, inexperienced, and on the road: Do novice drivers comply with road rules? *Transportation Research Record, 2318*(1), 98–106. https://doi.org/10.3141/2318-12.

Seiler, S. J. (2015). Hand on the wheel, mind on the mobile: An analysis of social factors contributing to texting while driving. *Cyberpsychology, Behavior, and Social Networking, 18*(2), 72–78. https://doi.org/10.1089/cyber.2014.0535.

Seiler, S. J., & Kidwell, J. (2016). Defining the mobile-digital situation within perpetual telecopresence: An exploratory study of the impact of mobile phones on the life-world. *Sociological Spectrum, 36*(4), 208–235. https://doi.org/10.1080/02732173.2015.1137254.

Seiler, S. J., & Kirby, R. (2017). Carefree, connected, and driving: A study of social-structural factors contributing to texting while driving among teenagers. 9, 27.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. D. (1997). *Preventing crime: What works, what doesn't, what's promising*. University of Maryland, Department of Criminology and Criminal Justice.

Simons-Morton, B. G., Guo, F., Klauer, S. G., Ehsani, J. P., & Pradhan, A. K. (2014). Keep your eyes on the road: Young driver crash risk increases according to duration of distraction. *Journal of Adolescent Health, 54*(5), S61–S67. https://doi.org/10.1016/j.jadohealth.2013.11.021.

Stafford, M. C., & WARR, M. (1993). A reconceptualization of general and specific deterrence. *Journal of Research in Crime and Delinquency, 30*(2), 123–135.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2017). The smartphone and the driver's cognitive workload: A comparison of Apple, Google, and Microsoft's intelligent personal assistants. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 71*(2), 93–110. https://doi.org/10.1037/cep0000104.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*(6), 462–466. https://doi.org/10.1111/1467-9280.00386.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 47*(2), 143–148. https://doi.org/10.1016/0001-6918(81)90005-6.

Tamul, D., Einstein, C., Hotter, J., Lanier, M., Purcell, L., & Wolf, J. (2021). Narrative persuasion and stigma: Using news accounts to denormalize texting while driving. *Accident Analysis & Prevention, 151*, 105876. https://doi.org/10.1016/j.aap.2020.105876.

Tivesten, E., & Dozza, M. (2015). Driving context influences drivers' decision to engage in visual–manual phone tasks: Evidence from a naturalistic driving study. *Journal of Safety Research, 53*, 87–96. https://doi.org/10.1016/j.jsr.2015.03.010.

Tractinsky, N., Ram, E. S., & Shinar, D. (2013). To call or not to call—That is the question (while driving). *Accident Analysis & Prevention, 56*, 59–70. https://doi.org/10.1016/j.aap.2013.03.017.

Transport Accident Commission. (2018). Road Safety Monitor 2018. http://www.tac.vic.gov.au/__data/assets/pdf_file/0018/417420/Road-Safety-Monitor-2018-Final-Report-23.01.20.pdf.

Transport Accident Commission. (2019). Road Safety Monitor 2019. https://www.tac.vic.gov.au/__data/assets/pdf_file/0011/445781/Road-Safety-Monitor-2019-Final-19.05.20.pdf.

Transport Accident Commission. (2020). TAC campaign targets Bloody Idiots who Drink Drive this September. TAC. https://www.tac.vic.gov.au/about-the-tac/media-room/news-and-events/2013-media-releases/tac-campaign-targets-bloody-idiots-who-drink-drive-this-september.

Truelove, V., Freeman, J., & Davey, J. (2019). "I Snapchat and Drive!" A mixed methods approach examining snapchat use while driving and deterrent perceptions among young adults. *Accident Analysis & Prevention, 131*, 146–156. https://doi.org/10.1016/j.aap.2019.06.008.

VicRoads. (2019). Mobile phones, technology & driving (Victoria) [Text]. VicRoads. https://www.vicroads.vic.gov.au:443/safety-and-road-rules/driver-safety/mobile-phones-and-driving.

VicRoads. (2020). Behavior change programs (Victoria) [Text]. VicRoads. https://www.vicroads.vic.gov.au:443/licenses/demerit-points-and-offences/behavior-change-programs.

White, M. J., Cunningham, L. C., & Titchener, K. (2011). Young drivers' optimism bias for accident risk and driving skill: Accountability and insight experience manipulations. *Accident Analysis & Prevention, 43*(4), 1309–1315. https://doi.org/10.1016/j.aap.2011.01.013.
World Health Organization. (2018). Global status report on road safety 2018. https://apps.who.int/iris/bitstream/handle/10665/277370/WHO-NMH-NVI-18.20-eng.pdf?ua=1.
Young, K. L., & Lenné, M. G. (2010). Driver engagement in distracting activities and the strategies used to minimise risk. *Safety Science, 48*(3), 326–332. https://doi.org/10.1016/j.ssci.2009.10.008.
Zhou, R., Yu, M., Wang, X., & Allen, P. (2016). Why do drivers use mobile phones while driving? The contribution of compensatory beliefs. *PLoS ONE, 11*(8), e0160288. https://doi.org/10.1371/journal.pone.0160288.

Fareed Kaviani is a Doctoral Researcher at BehaviourWorks, in the Monash Sustainable Development Institute.

Kristie Young is a Senior Research Fellow and an ARC Discovery Early Career Researcher Award (DECRA) Fellow with the Monash University Accident Research Centre (MUARC) in Melbourne, Australia.

Brady Robards is a Senior Lecturer in Sociology at Monash University in the School of Social Sciences.

Dr Sjaan Koppel holds a PhD in Psychophysiology and is an Associate Professor at the Monash University Accident Research Centre (MUARC).

Journal of Safety Research 78 (2021) 210–220

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Linking emotional intelligence to safety performance: The roles of situational awareness and safety training

Zhongmin Wang [a,*], Zhou Jiang [b], Anna Blackman [a]

[a] College of Business, Law and Governance, James Cook University, Australia
[b] College of Business, Government and Law, Flinders University, Australia

A R T I C L E I N F O

Article history:
Received 23 November 2020
Received in revised form 20 February 2021
Accepted 8 June 2021
Available online 22 June 2021

Keywords:
Emotional intelligence
Situational awareness
Safety
Training

A B S T R A C T

Introduction: Safety outcomes in the workplace require individual employees to perform (behave) safely in everyday duties. While the literature suggests that emotional management capabilities or traits can be positively related to individual performance in certain conditions, it is not clear how they can influence safety-related performance in high-risk work contexts. Drawing upon trait activation theory, this paper aims to examine when emotional intelligence (EI) benefits employees' safety performance. We propose that when employees receive inadequate safety training, EI is more likely to trigger their situational awareness and consequently promote their safety performance. *Method:* We collected time-lagged data from 133 full-time airplane pilots working in commercial aviation industry. Hierarchical moderated regression analysis was conducted to test the moderating effect of safety training inadequacy on the EI–situational awareness relationship. The moderated mediation model, which involves conditional indirect effects of EI on safety performance via situational awareness across different levels of safety training inadequacy, was tested using the PROCESS-based bootstrap confidence interval. *Results:* Safety training inadequacy negatively moderated the relationship between EI and situational awareness, such that EI was significantly related to situational awareness only when safety training inadequacy was more salient. The more inadequate safety training was, the greater the indirect effect of EI on safety performance via situational awareness was. *Conclusions:* Inadequate safety training, as a negative situational cue, can activate individuals' EI to drive their safety-related cognitions (e.g., situational awareness) and behaviors. Effective safety training may be able to complement employees' low EI in shaping their situational awareness and safety behaviors. *Practical Applications*: Aviation managers should monitor the adequacy and effectiveness of safety training; this could make pilots' situational awareness and safety performance depend less on personal attributes (e.g., EI), which organizations are less able to control. When training capacity is temporarily limited, priority might be given to those with low EI.

Emotional Intelligence (EI) has attracted substantial scholarly attention due to its significant implications on workplace outcomes, including safety-related results (Karimi, Leggat, Bartram, & Rada, 2020; O'Boyle, Humphrey, Pollack, Hawver, & Story, 2011; Sunindijo & Zou, 2013). A large number of empirical studies have demonstrated that EI has a positive impact on individuals' work performance (Karimi et al., 2020; Law, Wong, & Song, 2004). For example, some researchers found that EI-specific attributes, including one's understanding of his/her own and others' emotions, and regulation and utilization of emotions, collectively enhance task and contextual performance (Bozionelos & Singh, 2017; O'Boyle et al., 2011). However, other studies have found that EI has only a marginal or non-significant effect on employee performance (Tu, Guo, Hatcher, & Kaufman, 2018). The inconsistent results suggest that the relationship between EI and work-related performance might be subject to certain conditions. For example, while the effect of EI on important work outcomes such as performance has been emphasized, both theoretically and empirically, some EI researchers argue that the study of EI is better situated in a specific context, particularly where emotions are likely to cause undesirable feelings or psychological states (Miao, Humphrey, & Qian, 2017).

In line with this contention, scholars (e.g., Sunindijo & Zou, 2013) have attempted to investigate the influence of EI in typical safety–critical work contexts (e.g., the construction setting), showing that EI facilitates workers in implementing safety management

* Corresponding author.
E-mail addresses: aaron.zhongmingwang@gmail.com, zhongmin.wang@my.jcu.edu.au (Z. Wang).

https://doi.org/10.1016/j.jsr.2021.06.005
0022-4375/

tasks. Safety–critical situations are prone to raising emotional reactions (Leung, Chan, & Yuen, 2010); thus, the traits or abilities that enable better control over and regulation of undesirable emotional experiences should help individuals keep safe (Wang, Zou, & Li, 2016). Despite being theoretically meaningful, as we will discuss shortly, empirical research regarding the relationship between EI and safety performance is underdeveloped in multiple ways. To advance our knowledge in this regard, the current paper tests a situation-incurred, conditionally-mediated process underlying the EI–safety performance linkage.

While prior research has assisted with our understanding or forecast of the role of EI in predicting various types of work performance (e.g., safety performance), some areas warrant further investigation. For example, increasing numbers of voices in the literature argue that the influence of EI on performance is more indirect than direct (Ingram, Peake, Stewart, & Watson, 2019; Rode et al., 2007). Within these voices, researchers have argued for the potentiality that the insignificant EI–performance relationship observed in past research (e.g., Tu et al., 2018) could have been due to the ambiguity of EI's ability in explaining more proximate enablers of work performance. This indicates the need to place a focus on the core, proximate indicators of individual performance to explore their variations that are attributed to EI or similar emotional traits or abilities. In the safety-related domain, Endsley (1988) claims that situational awareness is the most important and proximal indicator of individuals' safety performance, especially in industries such as aviation, gas and mining, and construction. This claim is supported by the perspective that situational awareness denotes safety-oriented cognitions or abilities (Endsley, 2000). Based on Endsley, situational awareness is characterized by being aware of what is happening around the workplace, and this involves the capability of appraising critical environmental cues, processing vital safety information, forecasting near-future occurrences, and finding solutions to manage emerging risks. That is, safety performance is largely embedded in individuals' situational awareness (Caponecchia, Zheng, & Regan, 2018; Endsley & Robertson, 2000). As such, we first focus on situational awareness, which is argued to be a direct manifestation of safety performance, and its relationship with EI.

As stated above, prior research suggests inconsistent findings regarding the EI–performance relationship, indicating the roles of boundary conditions. While situational awareness is not a type of performance itself, it denotes a critical, immediate indicator of safety performance (Sneddon, Mearns, & Flin, 2013); thus, we expect that the relationship between EI and situational awareness may also be subject to certain boundary conditions. In this paper, we consider training-related conditions that may intervene with EI to affect employees' psychological and behavioral reactions. This consideration is based on the view that external cues are important to drive how people use personal resources to guide their understanding of relevant contexts (Tett & Burnett, 2003), and that this importance becomes more salient when external situations become more challenging (Farh, Seo, & Tesluk, 2012). Specifically, for this study, EI is a personal resource that people might rely on more under more challenging circumstances to shape their understanding of the associated environment (e.g., situational awareness) and thus to direct subsequent behaviors (e.g., safety behaviors). The literature has highlighted safety training inadequacy as a significant concern in safety–critical industries and organizations (Chan, Wong, Hon, Javed, & Lyu, 2017), as evidenced in research suggesting that ineffective or insufficient safety training represents significant challenges that produce heightened anxiety and stress (Huber, Hill, & Merritt, 2015). Therefore, we examine safety training inadequacy as a moderator (i.e., boundary condition) for the effect of EI.

Integrating these ideas, this paper develops and tests a model (Fig. 1), which proposes that safety training inadequacy will interact with EI to influence situational awareness, which in turn influences safety performance. We draw on trait activation theory (TAT) (Tett & Burnett, 2003) to conceptualize the moderating effect of training inadequacy on the EI–situational awareness relationship. According to TAT, trait-relevant external cues (e.g., task, social, and organizational demands or stressors) are likely to strengthen or weaken the relationship between traits and individuals' cognitive and behavioral outcomes, for traits can be activated by these cues to guide individuals' thoughts and actions. TAT and related research highlights that these traits are broadly defined as personal attributes that may not change rapidly in time (Farh et al., 2012). In the present study, EI is such an attribute; it denotes an individual's "emotions-related behavioral dispositions and self-perceived abilities" (Sanchez-Ruiz, Mavroveli, & Poullis, 2013, p. 658). Following Farh et al.'s (2012) application of TAT in EI research, as we will theorize later, this paper argues that EI will be activated by challenging external cues (e.g., safety training inadequacy) to affect situational awareness. Since situational awareness is an immediate enabler of safety performance (Endsley, 2000), we also expect that when activated by training inadequacy, the role of EI will extend to impact safety performance through situational awareness.

This paper contributes to the literature in different ways. First, it sheds light on an indirect approach to exploring the EI–performance link, for which prior research has generated inconclusive results. Specifically, by extending this link to the safety context, we emphasize the importance of focusing on a core, proximal indicator of safety performance (i.e., situational awareness) to appreciate the role played by EI. Second and relatedly, our research advances the EI–performance literature by identifying safety training inadequacy as a boundary condition that triggers emotional competence to function in building employees' situational awareness. Extending TAT (Tett & Burnett, 2003) into the safety performance setting to explain the effect of EI, we verify that EI, a self-perceived ability operationalized through a trait approach (i.e., self-rated EI; Wong & Law, 2002), when activated by the inadequacy of training (an opposing work demand), contributes to shaping situational awareness. Third, it enriches the EI and work performance theories by verifying the moderating mediation model of EI–safety performance involving situational awareness

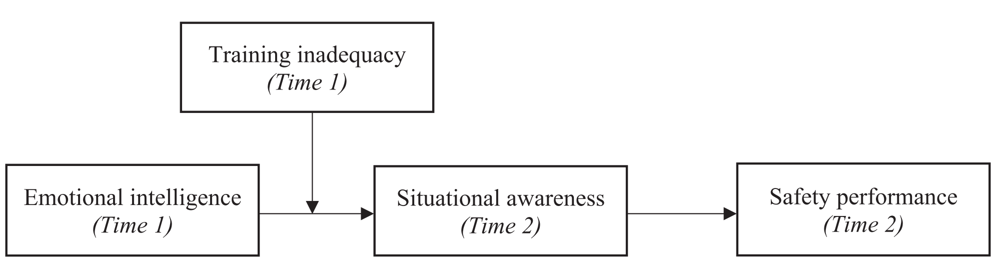

Fig. 1. Research model.

as a mediator and safety training inadequacy as a boundary condition, thereby providing a new explanation of when and how EI might matter in employee performance.

1. Literature review and hypothesis development

According to Mayer and Geher (1996), EI refers to individuals' ability to regulate and use emotional competence to guide thinking and improve performance. In the daily workplace, which Khalili (2012) calls an emotion-eliciting environment, there would be a variety of emotional distractions from intrapersonal factors (e.g., personal health and family issues), interpersonal factors (e.g., the relationship with colleagues and supervisors), and other factors (e.g., stressful company or environment). These distractions can perturb people's moods and cause unpleasant feelings, and the planned work scheme may thus be interrupted when individuals are less able to control these emotional experiences (Cardenas, Major, & Bernas, 2004). It is important to cast aside the distractions and focus on the most urgent/important tasks to avoid unnecessary consumption of mental energy if one is to be effective at work.

The literature suggests the relevance of EI-related attributes in one's awareness of or attention to one's affiliated situation. For example, Brackett, Rivers, and Salovey (2011) have summarized that employees with high emotional regulation skills tend to be more vigilant to emerging information from colleagues and the work environment, be more proficient in distinguishing various types of information, and shift attention from less significant tasks to focus on the more critical. In the course of performing work, situational awareness plays a crucial role during the decision-making process (Endsley & Robertson, 2000). According to Endsley (1988, p. 97), situational awareness refers to "the perception of elements of the environment, the comprehension of their meaning, and the projection of their future status." The process of accomplishing a specific task involves effective information analysis, planning, decision-making and action. As Durso, Hackworth, Truitt, Crutchfield, and Nikolic (1999) emphasized, situational awareness establishes the foundation of decision-making that requires significant attention to contextual cues.

Research indicates that it is important to maintain a comfortable stable mood in the workplace without the interference of negative emotions (Muchinsky, 2000). This requires a person to possess effective emotional management abilities or traits (e.g., EI). Emotions bias people's thinking and behaviors (Miner & Glomb, 2010). For instance, individuals are likely to overestimate their capabilities and neglect the distal details when in a good mood. In contrast, people tend to underestimate their abilities and lack confidence when in a bad mood. Either way, intense emotions could alter people's way of thinking, interfere with their attention, and lead them to make decisions that violate their original purpose (Beal, Weiss, Barros, & MacDermid, 2005). These arguments assert that EI may be positively related to situational awareness because EI involves one using emotions to reason about situational information and to help undertake rational behaviors. However, as discussed earlier, in safety contexts, this potential relationship between EI and situational awareness might be subject to the level of safety training inadequacy. Below, we theorize how safety training inadequacy might alter the strength of this relationship.

2. The moderating role of safety training inadequacy

This paper draws upon the trait-activation perspective to discuss the role of training inadequacy. Previous research has proved that many personal traits are stable dispositions that can predict job performance (Barrick, Mount, & Judge, 2001; Mount, Barrick, & Stewart, 1998) and verified that whether and to what extent they can influence one's behaviors and performance is subject to situational cues (Kell, Rittmayer, Crook, & Motowidlo, 2010). Based on a personality–job fit perspective combined with a traits–situation interaction perspective, Tett and Burnett (2003) introduced TAT to explain how traits or stable individual attributes can drive one's psychological and behavioral reactions by considering contextual interferences.

TAT presents a person–situation interactionist perspective suggesting that trait-relevant situations (or situational cues) can activate an individual's way of expressing their traits, therefore influencing job performance (Tett & Burnett, 2003). In the workplace, employees prefer to seek jobs that can easily express their instinctive traits or use their attributes to pursue satisfaction and achieve success (Farh et al., 2012). For example, an introverted person will not look for a sales or customer service job because he/she knows the challenges will put him/her in an unfavorable situation. In contrast, an extraverted person may easily use their strengths to gain satisfaction in the above jobs (Barrick, Stewart, & Piotrowski, 2002). This is consistent with Johns' (2006) perspective that work contexts can be situational opportunities or constraints to enlarge or suppress the influence of traits on work performance, depending on whether the situation or associated cues are relevant/important to the traits.

In the present study, EI is a trait-like attribute that reflects one's self-perceived capability (Wong & Law, 2002). Applying TAT to this case, certain situational or context cues might activate EI to manifest its function on employee outcomes. TAT states that traits are activated by task, social, or organizational cues to affect performance (Tett & Burnett, 2003), partly because activated traits or relatively-stable attributes (e.g., trait-like abilities) elicit cognitions or psychological states before changing behaviors or performance (Blickle, Schütte, & Genau, 2018). Situational awareness, as noted earlier, reflects such psychological cognitions relevant to one's environment, and it is a precursor of job performance (e.g., safety performance) (Irwin, Caruso, & Tone, 2019). In line with TAT, we propose that the EI and situational awareness relationship will be stronger when safety training inadequacy becomes more salient.

Training equips individuals with essential skills and relevant capabilities to perform their tasks and duties in an effective manner (Bartel, 1995). In a safety–critical work context, inadequate safety training may result in consequences harmful to the individual (e.g., stress, errors, and mistakes) and the organization (e.g., turnover, low productivity, and jeopardized safety patterns) (Dysvik & Kuvaas, 2008; Elnaga & Imran, 2013; Zhao, Hwang, & Gao, 2016). Importantly, when employees are aware that the company offers insufficient training, which hinders them from gaining relevant resources that are needed to navigate or explore the safety–critical environment (Givehchi, Hemmativaghef, & Hoveidi, 2017), they tend to seek from inside (i.e., within the individual) and use related personal abilities/resources to deal with the situation (Rhee, Hur, & Kim, 2017).

A lack of safety training may represent risky situational constraints (Zhao, Wu, & Wang, 2018), which to some extent signal the cue that employees may not immediately gain resources from the organization before the situation is improved. Research indicates that inadequate training impairs employees' confidence, raises concerns, and incurs negative emotions (e.g., fear, depression, and anxiety, Huber et al., 2015). These negative psychological and emotional experiences would need individuals to access and utilize effective emotional management abilities or attributes to maintain their cognitive, attitudinal, or behavioral engagement at work (Rothbard & Wilk, 2011).

This line of reasoning suggests that when safety training is not adequate, it sends out the contextual cue that the lack of knowledge/information to evaluate or forecast the safety–critical situation will lead to emotionally challenging circumstances. In this case, EI, which can assist to manage these challenging circumstances, is likely to be activated to help deal with associated unpleasant emotions (e.g., frustration, fear, confusion, or anxiety) so as to maintain situational awareness, which is a key to ensuring safety (Endsley, 2000). Therefore, when there is a lack of training, EI will play a stronger role in promoting and maintaining situational awareness. Conversely, when training is abundant, the role of EI might be less prominent. Thus, we propose:

Hypothesis 1: Safety training inadequacy will strengthen the relationship between EI and situational awareness.

3. The moderated mediation effect on safety performance

In addition to the above direct implications for situational awareness, we also propose that EI and safety training adequacy have downstream implications for safety performance. In fact, as mentioned earlier, past research has verified that situational awareness is an important antecedent of safety performance (Endsley, 1999; Fernandes & Braarud, 2015). In the case of the complex work environment, a lack of situational awareness can cause severe consequences that incur risks and unsafe behaviors (Nazir, Colombo, & Manca, 2012). Some researchers hold that situational awareness reflects cognitive motivations for information processing (van Winsen, Henriqson, Schuler, & Dekker, 2015), which is needed to direct safety actions. Based on the workplace safety research literature, these motivational characteristics associated with situational awareness can directly enable employees to perform tasks safely (Christian, Bradley, Wallace, & Burke, 2009). Integrating these arguments with the aforementioned discussions of the interactive effect of EI and safety training inadequacy on situational awareness, we propose that safety training adequacy will moderate the mediated effect of situational awareness on the relationship between EI and safety performance. Specifically, under high levels of safety training inadequacy, individuals will rely more on EI to maintain situational awareness, and in turn achieve better safety performance. In contrast, when safety training is abundant (i.e., low levels of training inadequacy), individuals become more confident in using gained safety-related knowledge and skills through training to conduct their tasks; thus, their EI tends to have less of an impact on situational awareness and in turn on safety performance. Therefore, we posit the following moderated mediation effects:

Hypothesis 2: Safety training inadequacy moderates the mediated relationship between EI and safety performance through situational awareness, such that the mediated relationship will be stronger when safety training is more inadequate.

4. Methods

4.1. Sample and procedure

We collected a convenience sample of full-time pilots working in the commercial aviation industry in China. With the assistance of a fleet manager, paper and online questionnaires were administered among frontline pilots in four airlines in mainland China. To reduce common method bias that is associated with self-reported data (Podsakoff, MacKenzie, & Podsakoff, 2012), participants were asked to complete a two-wave survey at two separate time points. In the first questionnaire (Time 1), pilots provided demographic information and answered questions regarding EI and safety training inadequacy. Approximately one month later (Time 2), they were asked to respond to the second questionnaire, which included questions regarding situational awareness and safety performance.

In Time 1, a total of 211 participants (response rate = 79.3%) returned useable responses. Out of these participants, 161 returned valid questionnaires (response rate = 76.3%) in Time 2. For each individual, his or her two questionnaires were matched using a self-created, unique code that mixed numbers and letters. Finally, questionnaires were successfully matched for 133 pilots. In this sample, 97.7% of the respondents were male. The average age was 28.66 years (SD = 4.11), and all of them had completed tertiary education. The average job tenure was 38.86 months (SD = 41.27).

4.2. Measures

Measurement items were originally written in English and then translated into Chinese following a back-translation procedure (Brislin, 1980). Participants rated all these items using a Likert-type scale (1 = strongly disagree; 5 = strongly agree).

EI. The 16-item scale developed by Wong and Law (2002) measured emotional intelligence. These items collectively capture individuals' abilities in understanding their own and others' feelings, regulating emotions, and using emotions for motivation purposes. Example items are "I am sensitive to the feelings and emotions of others" and "I am able to control my temper and handle difficulties rationally." The Cronbach's α for emotional intelligence was 0.89.

Safety training inadequacy. Four items developed by Evans, Glendon, and Creed (2007) were employed to measure the inadequacy of safety training in the organization. An example item is "Company training provided adequate skills and experience to carry out normal operations safely" (reverse coded). The Cronbach's α for safety training inadequacy was 0.85.

Situational awareness. The 13-item scale created by Sætrevik (2013) was adopted to measure pilots' situational awareness. Example items are "I plan ahead in order to handle various and adverse incident that may arise" and "I usually know what's going to happen next with regard to safety." The Cronbach's α for situational awareness was 0.62.

Safety performance. The seven-item instrument developed by Griffin and Neal (2000) was employed to measure pilots' safety performance. This instrument captures individuals' compliance and participation in safety procedures and behaviors in the workplace setting. Example items are "I carry out work in a safe manner" and "I help my co-workers when they are working under risky or hazardous conditions." The Cronbach's α for safety performance was 0.93.

Control variables. As prior research indicated that age, gender, and job tenure could potentially impact employees' safety performance (DeJoy, Schaffer, Wilson, Vandenberg, & Butts, 2004; Siu, Phillips, & Leung, 2003), as in other studies (e.g., Griffin & Hu, 2013), these variables were controlled in the data analyses. Because safety performance was self-reported, the supplementary analyses controlled for social desirability, which, as some researchers have argued, could potentially lead participants to offer slightly, if not to a large extent, more favorable performance ratings (e.g., Xie, Roy, & Chen, 2006). If the results remain similar regardless of controlling for social desirability, or the expected relationship becomes more prominent, or the social desirability is not significantly related to performance/behaviors, it is less likely that the results are distorted by self-rated measures of performance or behaviors (e.g., Cheng, Yen, & Chen, 2012; Crant, 1995). Social desirability was measured by five items developed by Hays, Hayashi, and Stewart (1989). An example item is "I am always courteous even to people who are disagreeable." The Cronbach's α for social desirability was 0.63.

4.3. Data analysis

Missing values were dealt with by multiple imputation (Bernaards & Sijtsma, 2000). Specifically, if a participant had missing data on several, but not all, items on the scale for a variable, we followed Wanberg, Zhu, and Van Hooft (2010) in applying the full information maximum likelihood (FIML) to replace these missing values. Those participants that had all items missing on a scale were excluded from subsequent analyses, for there was no information/data on this scale that could be used for imputation (Bernaards & Sijtsma, 2000). Three cases had missing values for age and/or job tenure and were excluded from the main analyses. This process resulted in a sample of 130 for confirmatory factor analysis (CFA), and the listwise deletion resulted in a sample of 126 for correlational and regression analyses.

Before the hypotheses were tested, CFA was conducted to examine the discriminant validity of the four focal study variables (i.e., EI, safety training inadequacy, situational awareness, and safety performance) measured by multiple items. Considering the relatively small sample size and the large number of measurement items, we followed the recommendation of Little, Cunningham, Shahar, and Widaman (2002) in creating item parcels to reduce the inflation errors. EI and situational awareness were each represented by four parcels and safety performance by two parcels. The moderation hypothesis (Hypothesis 1) was examined using hierarchical regression analysis. The moderated mediation hypothesis (Hypothesis 2) was tested using the PROCESS code for SPSS (Hayes, 2013).

5. Results

5.1. Confirmatory factor analysis (CFA)

To check if the key variables featured in the research model (Fig. 1) could be distinguished from one another, we conducted CFA to test the measurement model. The hypothesized baseline model was a four-factor model, in which EI, safety training inadequacy, situational awareness, and safety performance were loaded on four separate factors. We compared this model with six three-factor models, three two-factor models, and a one-factor model. The results of CFA are shown in Table 1.

Following prior researchers (e.g., Franco-Santos, Nalick, Rivera-Torres, & Gomez-Mejia, 2017; Millner et al., 2020), we used three commonly used fit indexes to assess model fit: the root mean square error of approximation (RMSEA), the standardized root means square residual (SRMR), and comparative fit index (CFI). If a model's fit indexes meet the cut-off criteria (RMSEA <0.08, SRMR <0.08, and CFI >0.90), this model is regarded to fit the data well (Hu & Bentler, 1999). As shown in Table 1, the four-factor measurement model met all these criteria, while the other models did not. In addition, the chi-square difference ($\Delta\chi^2$) test further supported that the four-factor measurement model achieved a fit better than those of all other alternative models. This result suggested that EI, safety training inadequacy, situational awareness, and safety performance were empirically distinct constructs in the current research, sufficiently justifying the treatment of them as separate variables in the following analyses.

5.2. Descriptive statistics

The means, standard deviations, and correlations of the demographic controls and the four main variables are shown in Table 2. EI was significantly related to safety performance but not situational awareness. Consistent with several scholars' (Farh et al., 2012) view that the influence of EI on employee states or outcomes to some extent depends on boundary conditions, this result offered initial insights on the necessity of testing moderators of EI regarding its relationship with situational awareness. Detailed results regarding the moderation follow.

5.3. Results of hypothesis testing

Hypothesis 1, the moderation hypothesis, predicted that safety training inadequacy would strengthen the relationship between EI and situational awareness. It required a test of the interaction effect of EI and safety training inadequacy on situational awareness. We followed the well-established and most widely used procedure for moderation testing (Hayes, Montoya, & Rockwood, 2017), hierarchical regression analysis, in examining this hypothesis. The results are reported in Table 3.

As can be seen from Table 3, we followed our predecessors (e.g., Oliver, Hausdorf, Lievens, & Conlon, 2016) in performing a three-step hierarchical regression. In Step 1, control variables including gender, age, and job tenure were entered. In Step 2, the main effect step, the independent (i.e., EI) and moderating (i.e., safety training inadequacy) variables were entered. In Step 3 was entered the interaction term, which was equal to the product of IE and safety training inadequacy. In accordance with the advice of Aiken, West, and Reno (1991), both EI and safety training inadequacy

Table 1
CFA results.

Models	χ^2	*df*	$\Delta\chi^2$	Δdf	χ^2/df	SRMR	RMSEA	CFI
4-factor model	**99.32***	**71**	**–**	**–**	**1.40**	**0.07**	**0.06**	**0.96**
3-factor model A	173.91***	74	74.60***	3	2.35	0.10	0.10	0.84
3-factor model B	169.83***	74	70.51***	3	2.29	0.10	0.10	0.85
3-factor model C	197.74***	74	98.42***	3	2.67	0.11	0.11	0.81
3-factor model D	258.40***	74	159.09***	3	3.49	0.11	0.14	0.71
3-factor model E	243.46***	74	144.14***	3	3.29	0.10	0.13	0.73
3-factor model F	181.26***	74	81.94***	3	2.45	0.11	0.11	0.83
2-factor model A	267.06***	76	167.74***	5	3.51	0.13	0.14	0.70
2-factor model B	329.79***	76	230.47***	5	4.34	0.13	0.16	0.60
2-factor model C	322.66***	76	223.35***	5	4.25	0.13	0.16	0.61
1-factor model	423.29***	77	323.97***	6	5.50	0.16	0.19	0.46

Note. $N = 130$. Three cases were excluded because of missing values. 4-factor model: each variable was treated as a single factor; 3-factor model A: EI and situational awareness were combined; 3-factor model B: situational awareness and safety performance were combined; 3-factor model C: EI and training inadequacy were combined; 3-factor model D: Training inadequacy and safety performance were combined; 3-factor model E: EI and safety performance were combined; 3-factor model F: training inadequacy and situational awareness were combined; 2-factor model A: EI and training inadequacy were combined; situational awareness and safety performance were combined; 2-factor model B: EI and situational awareness were combined; training inadequacy and safety performance were combined; 2-factor model C: EI and safety performance were combined; situational awareness and training inadequacy were combined; 1-factor: all variables were combined.

Table 2
Means, standard deviations, and correlations.

Variables	Mean	SD	1	2	3	4	5	6
1. Gender	0.98	0.15						
2. Age	28.75	4.12	0.00					
3. Job tenure (months)	39.31	41.90	0.00	0.66**				
4. Emotional intelligence	3.89	0.50	−0.02	−0.17	−0.17			
5. Training inadequacy	1.65	0.66	−0.04	0.17	0.12	−0.29**		
6. Situational awareness	3.61	0.37	0.010	0.16	0.12	0.13	−0.09	
7. Safety performance	4.32	0.72	0.03	−0.06	−0.03	0.26**	−0.26**	0.21*

Note. N = 126. Listwise deletion was applied. Gender was dummy coded (male = 1 and female = 0).
* $p < .05$.
** $p < .01$.

Table 3
Results of moderated regression analyses (Hypothesis 1).

	Situational awareness		
	Step 1	Step 2	Step 3
Gender	0.10	0.09	0.10
Age	0.14	0.17	0.20
Job tenure (months)	0.02	0.04	0.04
Emotional intelligence (EI)		0.14	0.17+
Training inadequacy		−0.08	−0.09
EI × training inadequacy			0.21*
R^2	0.03	0.07	0.11*
ΔR^2		0.03 [a]	0.04*[,b]

Note. N = 126. Listwise deletion was applied. Standardized estimates are reported.
\+ $p < 0.10$
* $p < 0.05$.

were mean-centered before the interaction term was calculated to reduce multicollinearity. According to Baron and Kenny (1986), testing a moderating effect does not require the main effects of the independent and moderating variables to be statistically significant. When the interaction term is significant in predicting the dependent variable (i.e., situational awareness for this particular moderation analysis), it reveals the existence of a moderation.

As Table 3 presents, the interaction between EI and safety training inadequacy was significant in predicting situational awareness ($\beta = 0.21$, $p < .05$), providing initial support for the moderating role of safety training inadequacy. To verify if the moderation was in the expected direction, we followed Dawson (2014) in creating a graphic presentation of the interaction effect. Specifically, we plotted the simple slope of the effect of EI on situational awareness one standard deviation above and below the mean of safety training inadequacy (see Fig. 2). Results of simple slope analysis showed that EI had a significant, positive effect on situational awareness only when safety training inadequacy was high (simple slope = 0.28, $t = 2.76$, $p < .01$), and there was not a significant effect when safety training inadequacy was low (simple slope = −0.02, $t = -0.255$, $p > .10$). Therefore, the relationship between EI and situational awareness was stronger when there was a higher level of safety training inadequacy, supporting Hypothesis 1.

Hypothesis 2 predicted that training inadequacy would suppress the indirect relationship between EI and safety performance through situational awareness. This hypothesis represented a first-stage moderated mediation. According to Hayes (2013), the test of a first-stage moderated mediation requires (1) testing the moderation of the first stage of the mediation and (2) testing the conditional indirect effects. Specifically, for the current research, Hypothesis 1 has supported that safety training inadequacy moderated the first stage of the mediation (i.e., the link between EI and situational awareness). The conditional indirect effects were examined with Hayes' (2013) PROCESS code for SPSS with 5000 bootstrap samples. Following Belogolovsky, Bamberger, and Bacharach (2012), we used a 90% bias-corrected confidence interval to test the significance of conditional indirect effects in our main analyses. A confidence interval not including zero indicates that an indirect effect is statistically significant (Hayes, 2013).

Table 4 presents the conditional indirect effects across low and high levels of training inadequacy. Results showed that the conditional indirect effect of EI on safety performance through situational awareness was significant under high-level training inadequacy ($B = 0.10$, Boot $SE = 0.06$, 90%CI = [0.02, 0.21]) but non-significant under low-level training inadequacy ($B = -0.01$, Boot $SE = 0.04$, 90%CI = [−0.09, 0.06]). The index of moderated mediation was significant (index = 0.08, Boot $SE = 0.06$, 90%CI = [0.01, 0.21]), suggesting that these two conditional indirect effects were significantly different from each other. These results demonstrated that only under high levels of safety training inadequacy could situational awareness mediate the relationship between EI and safety performance. Therefore, the mediated relationship was stronger when safety training inadequacy was high rather than low, and Hypothesis 2 was supported.

5.4. Supplementary analysis

As noted earlier, since some researchers argue that self-rated performance may be inflated by the respondents because of the socially desirable orientation (Schriesheim, 1980), we conducted additional analyses to test the proposed hypotheses by controlling for social desirability. Specifically, in these supplementary analyses, social desirability was added to the group of control variables in the hierarchical regression analysis for the moderation (for Hypothesis 1) and to the group of covariates controlled for both situational awareness and safety performance in the PROCESS analysis (for Hypothesis 2).

Results showed that the hypotheses remained supported. Similar to the main analyses reported above, the interaction term of EI and safety training inadequacy in predicting situational awareness was positive and significant ($\beta = 0.21$, $p < .05$). PROCESS results, with social desirability controlled for and based on 5000 bootstrap samples, demonstrated that the conditional indirect effect of EI on safety performance through situational awareness was stronger when safety training inadequacy was high ($B = 0.10$, Boot $SE = 0.06$, 95%CI = [0.01, 0.27]) rather than low ($B = -0.02$, Boot $SE = 0.04$, 95%CI = [−0.12, 0.06]). The index of moderated mediation was significant (Index = 0.09, Boot $SE = 0.07$, 95%CI = [0.001, 0.26]). Altogether, these results further consolidated the support of the proposed hypotheses.

6. Discussion

Researchers have begun investigating the role of EI in safety contexts (Sunindijo & Zou, 2013). However, our knowledge of how and when EI matters in promoting employees' safety performance is still limited. Inconsistent findings in the literature regard-

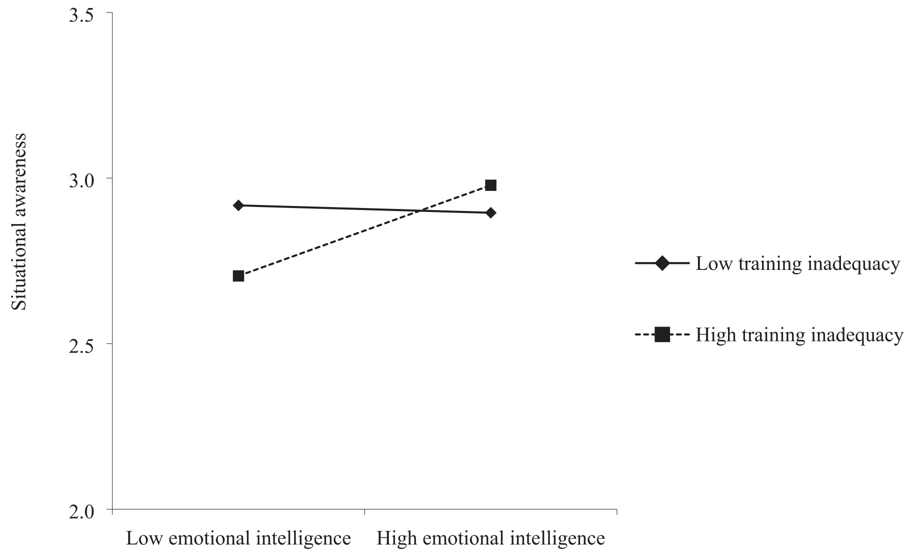

Fig. 2. The moderating effect of training inadequacy on the relationship between emotional intelligence and situational awareness.

Table 4
Conditional indirect effects of emotional intelligence on safety performance via situational awareness (Hypothesis 2).

	B	Boot SE	Boot LLCI	Boot ULCI
Conditional Indirect Effects				
Low-level training inadequacy	−0.01	0.04	−0.09	0.06
High-level training inadequacy	0.10	0.06	0.02	0.21
Index of Moderated Mediation	0.08	0.06	0.01	0.21

Note. LLCI = lower limit confidence interval (CI). ULCI = upper limit CI. 90%CI based on 5000 bootstrap samples. Low-level and high-level training inadequacy equals one standard deviation below and above the mean of training inadequacy.

ing the EI–performance linkage have led researchers to see the value of focusing on the indirect effect of EI on performance and the associated boundary conditions (i.e., moderators) to explore how EI influences performance (Rode et al., 2007). To this end, we investigated the relationship between EI and safety performance through the mechanism of situational awareness, a proximal enabler of safety performance, and the moderating role of safety training inadequacy, which is an important but underestimated boundary condition in the effects of EI. Our findings suggested that safety training inadequacy strengthened the link between EI and situational awareness, meaning that for those individuals receiving inadequate safety training, EI was more likely to contribute positively to their situational awareness, as compared to those receiving adequate training. In addition, we found that the moderating effect of safety training inadequacy could extend to moderate the indirect effect of EI on safety performance through situational awareness. In the following sections, we discuss the theoretical and practical implications of these findings, as well as the limitations of this study and suggestions for future research.

6.1. Theoretical implications

Drawing on the trait-activation perspective (Tett & Burnett, 2003), we proposed and found that EI would interact with safety training inadequacy to affect situational awareness, and in turn influence safety performance. Our study has important theoretical implications. First, we empirically contribute to the theoretical indication that the EI–performance relationship is more indirect through extending the focus to a proximal antecedent of safety performance (i.e., situational awareness) and probing the boundary condition underlying its relationship with EI. According to our results, the influence of EI on situational awareness is conditional on safety training inadequacy. Specifically, this influence tends to be consolidated when training is more inadequate. Therefore, training inadequacy, as a contextual factor, can trigger the role of EI in promoting situational awareness. This finding is consistent with TAT (Tett & Burnett, 2003), which emphasizes that personal attributes of a relatively stable nature can be activated by certain contexts to exert influences on one's cognitions, psychological states, and behaviors (Farh et al., 2012; Judge & Zapata, 2015). In our case, when safety training is less adequate, EI is activated to play a role. For example, when a company cannot provide sufficient safety training, from which employees may benefit little, if any, to improve their situational awareness at work where safety needs particular attention. In this case, as per our findings, employees with high-level EI may be better able to conquer the challenges associated with lack of training to maintain a certain level of situational awareness needed in safety contexts.

This observation regarding employees whose EI is activated is also in accordance with previous studies that show that, when facing a disadvantaged or challenging situation, individuals with higher levels of EI are more likely to take proactive actions than to be passive (Kim, Cable, Kim, & Wang, 2009). For example, high-EI individuals usually actively seek advice from experienced colleagues or supervisors, or they search for useful material online to increase their safety knowledge and boost safety-related situational awareness. In the workplace, individuals with high levels of EI will be more vigilant, such as becoming more cautious about the emerging information, paying attention to the details, and discreetly forecasting the changes (Brackett et al., 2011). These tendencies characterize one's mindfulness or heedfulness toward his or her associated environments, and thus they are an explicit symbolization of situational awareness (Brackett et al., 2011). In contrast, instead of being active in looking for solutions to tackle the challenges (e.g., inadequate training), low-level EI employees may respond passively and negatively by complaining, losing confidence, and misinterpreting contextual information, which can lead to or be the embodiments of poor situational awareness (Jordan et al., 2002).

Second, we extended the moderation model to test a moderated mediation that represents a more complex process explaining the EI–safety performance relationship. Specifically, we have verified the downstream implications of situational awareness, the level of which varies with the interaction of EI and safety training inadequacy, on safety performance. By doing so, we broaden the impact of EI on performance through an expanded application to safety performance at work, an underexplored area in both EI and safety domains, as well as supplement the emerging but limited studies (e.g., Sunindijo & Zou, 2013) on the EI–safety performance linkage, which have neglected the boundary conditions of how EI matters in boosting safety behaviors. The findings of the present study indicate that only when safety training is inadequate can situational awareness serve as a mediation mechanism to transmit the effect of EI to safety performance. When employees receive adequate training, the mediating effect of situational awareness disappears.

These findings provide evidence for Rode et al.'s (2007) contention that EI's influence on employees' attitudes and behaviors may not be overt or direct and that it depends on work contexts and conditions. In this research, safety training inadequacy, serving as an adverse condition, boosts a greater effect of EI to facilitate the development of situational awareness, which enables superior performance in a safety–critical environment. What is implied is that while negative conditions may strike individuals' mindsets and mental/cognitive models and lead to lower performance, personal attributes like EI are likely to help alleviate this strike and thus maintain the situational awareness-based process underlying the EI–safety performance relationship. This implication accords with the TAT theory, which indicates that organization-level constraints can activate less-changeable personal traits or abilities to initiate a process that improves performance (Tett & Burnett, 2003). In the present study, we have highlighted that when employees are aware that the company offers insufficient training that is needed for safety effectiveness, they tend to turn to the self and utilize related abilities or traits to guide their cognitions so as to achieve the desired safety performance. From a broader perspective, this is supportive of the view that the conflict between the demands of performing well and the lack of training make it important for individuals to rely on personal resources (e.g., EI) to master the situation and achieve better performance (Brackett, Rivers, Shiffman, Lerner, & Salovey, 2006; Rhee et al., 2017).

6.2. Practical implications

Our study has important implications for managerial practice. For example, our findings provide organizations and managers, particularly those in safety–critical industries, with the knowledge that employees' EI does matter in persuading them to comply with safety procedures and participate in building a safe workplace climate. However, to what extent managers can expect to rely on or improve employees' EI to ensure desirable safety performance depends on specific situations, some of which are under the organization's or the manager's control. Our results indicate that whether employees receive enough safety-related training can help guide managers' emphasis on employees' EI.

From the safety management perspective, it is ideal that organizations can provide comprehensive, quality safety training to employees. However, it is often the case that the organization lacks resources or has unintentionally neglected certain aspects, damaging the quality of training development or delivery (Analoui, 2000); alternatively, due to individual differences, the content may not be well received and/or digested by trainees even when quality development/delivery is assured from the organization side (Brown, 2001). Our results suggest that if the organization does not provide, or employees do not perceive, enough safety training, those with higher levels of EI may be better able to keep on track with safety performance, for their EI is more likely to help them maintain situational awareness, which is key to ensuring safe behaviors. Therefore, when managers are aware of problems with safety training, they might pay more attention to employees who are less emotionally intelligent, especially when these training problems may not be fixed in a rapid and/or effective manner. In this case, work unit managers may consider implementing strategies to improve employees' EI through, for example, coaching, mentoring, and peer support (Mattingly & Kraiger, 2019). It might also be useful if employees' EI is first assessed by professional experts (e.g., researchers and/or management consults) using one or more appropriate methods (e.g., surveys, test banks, and interviews). The information generated from such assessments could help managers make effective decisions regarding who should be prioritized (i.e., those with lower levels of EI) when EI-enhancing strategies are to be implemented.

When training ineffectiveness is an existing shortcoming, our findings also have implications on the recruitment and selection process. As mentioned earlier, we found that inadequate training makes the role of EI more salient in improving situational awareness and thus safety performance. Based on this finding, managers, aware of the organization's weakness in training, may consider incorporating EI assessments in the recruitment and selection of new staff for safety–critical positions. For example, it could be practical that an appropriate threshold is predetermined for the results of an EI assessment to exclude applicants with low levels of EI.

7. Limitations and future research

This study has a few limitations that future research could address. First, we focused specifically on the training (in)adequacy when exploring the boundary conditions of the effects of EI on situational awareness, and subsequently on safety performance. However, training (in)adequacy may only be considered a subcomponent of training (in)effectiveness, which may also contain elements such as uselessness/usefulness of training (Bell, Tannenbaum, Ford, NOE, & Kraiger, 2017). As such, it is uncertain whether the overall quality of safety training could moderate the influence of EI. Future research should consider the role of overall safety training effectiveness, which can more comprehensively capture the training-related situation, when testing the relationship between EI, situational awareness, and safety performance.

Second, related to the boundary condition, we exclusively concentrated on the interaction effect of the training-related context and EI, having neglected other possible contextual features. Indeed, the literature suggests that contextual variables at the job and organization levels might also serve as triggers to promote the functions of EI in employee outcomes. For example, a lack of job autonomy may require employees to handle barriers to satisfying important, basic psychological needs (e.g., need for autonomy) and thus may activate their EI to regulate negative feelings caused by relevant barriers (Kim et al., 2009). At the organization level, the clarity of safety policy may intervene in the effects of EI. When safety policy is more ambiguous, EI should be more likely to be activated because employees with higher levels of EI may be motivated to proactively seek meaning out of the ambiguous situation and cognitively pursue safety control (Huang, Chen, Krauss, & Rogers, 2004). Future research may consider additional moderators, both task- and organization-focused, to identify the conditions that promote or hinder EI from influencing safety-related outcomes.

Third, we did not control for some variables that may have confounded our proposed relationships. For example, existing research suggests that safety-specific orientations such as safety motivation (i.e., one's willingness to commit to safety behaviors; Neal & Griffin, 2006) and risk-taking orientation (Westaby & Lowe, 2005) affect safety performance. It has also been reported that characteristics such as trait mindfulness can influence situational awareness (Zhang, Ding, Li, & Wu, 2013). Future research should consider controlling for some of these confounders to explore whether EI or similar constructs can incrementally explain the variation of situational awareness and safety performance. To do this, it might be worth exploring available longitudinal panel data, which may have provided opportunities to explore and rule out potential confounders (Cheng, Guo, Hayward, Smyth, & Wang, 2020).

Fourth, we focused on situational awareness as a single conditional, mediation mechanism of the EI–safety performance linkage. Although this focus has verified the view that the link between EI and safety performance tends to be indirect, situational awareness might not be the only path through which EI can influence performance. The literature suggests that EI may promote positive and alleviate negative psychological states that are related to employees' wellbeing and thereby affect their performance (Mattingly & Kraiger, 2019; Sánchez-Álvarez, Extremera, & Fernández-Berrocal, 2016). It might be worthwhile for future research to examine wellbeing-related constructs such as burnout, engagement, and thriving at work as mediation mechanisms, as well as explore the associated boundary conditions.

Fifth, although our participants were commercial pilots from four different airlines in mainland China, the relatively small sample size may still have limited the generalizability of our findings considering that there are approximately 60,000 pilots in Chinese commercial aviation industry ("Statistics Bulletin on Civil," 2020). Future research may gather a larger sample from more airlines to further validate our findings. To generalize our findings to the broader population, future researchers may consider retesting our model in other safety–critical fields (e.g., mining, oil exploration, and manufacturing) with a greater sample size.

Finally, our sample was gender biased. Female pilots accounted for only 2.30% of our respondents, and thus our results might be more applicable for male pilots and not be generalized to females specifically. However, the gender bias was within the expectation, for female commercial pilots only account for 1.28% of the total pilot population in China (Brenda, 2018). Given that our sample size was relatively small and had a greater proportion of females than the national average, we conclude that the sample is representative in terms of gender. Considering that this female population is very small, future research may consider qualitative approaches to investigating the phenomena related to EI and safety if the experience of Chinese female commercial pilots is to be explored.

8. Conclusion

Drawing on data collected from pilots in commercial airlines, this study examined the influence of EI on safety performance through the mediating mechanism of situational awareness, with a focus on safety training inadequacy as a boundary condition. Our empirical findings suggest that EI is critical for employees to maintain a reasonable level of situational awareness that is needed to perform safety tasks, and this role of EI is more prominent when there is a lack of safety training. Our study contributes to the workplace safety literature by introducing trait activation theory to advance our understanding of how EI can be triggered in adverse work contexts (e.g., lack of safety training) to enhance situational awareness and safety performance. From a practitioner perspective, our findings suggest that, in addition to ensuring the effectiveness of safety training, companies may include EI tests when recruiting suitable employees for safety–critical positions, for high-EI individuals are more capable to conquer adverse situations.

References

Statistics Bulletin on Civil, 2020. (2020). Statistics bulletin on civil aviation industry development in 2019. Retrieved from https://worldbusinessanalytics.com/2020/06/10/statistics-bulletin-on-civil-aviation-industry-development-in-2019/.

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Analoui, F. (2000). What motivates senior managers? *Journal of Managerial Psychology, 15*(4), 324–340. https://doi.org/10.1108/02683940010330984.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*(1–2), 9–30. https://doi.org/10.1111/ijsa:2001.9.issue-1&210.1111/1468-2389.00160.

Barrick, M. R., Stewart, G. L., & Piotrowski, M. (2002). Personality and job performance: Test of the mediating effects of motivation among sales representatives. *Journal of Applied Psychology, 87*(1), 43–51. https://doi.org/10.1037//0021-9010.87.1.000.

Bartel, A. P. (1995). Training, wage growth, and job performance: Evidence from a company database. *Journal of Labor Economics, 13*(3), 401–425. https://doi.org/10.3386/w4027.

Beal, D. J., Weiss, H. M., Barros, E., & MacDermid, S. M. (2005). An episodic process model of affective influences on performance. *Journal of Applied Psychology, 90*(6), 1054–1068. https://doi.org/10.1037/0021-9010.90.6.1054.

Bell, B. S., Tannenbaum, S. I., Ford, J. K., Noe, R. A., & Kraiger, K. (2017). 100 years of training and development research: What we know and where we should go. *Journal of Applied Psychology, 102*(3), 305–323. https://doi.org/10.1037/apl0000142.

Belogolovsky, E., Bamberger, P. A., & Bacharach, S. B. (2012). Workforce disengagement stressors and retiree alcohol misuse: The mediating effects of sleep problems and the moderating effects of gender. *Human relations, 65*(6), 705–728. https://doi.org/10.1177/0018726711435250.

Bernaards, C. A., & Sijtsma, K. (2000). Influence of imputation and em methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research, 35*(3), 321–364. https://doi.org/10.1207/s15327906mbr3503_03.

Blickle, G., Schütte, N., & Genau, H. A. (2018). Manager psychopathy, trait activation, and job performance: A multi-source study. *European Journal of Work and Organizational Psychology, 27*(4), 450–461. https://doi.org/10.1080/1359432x.2018.1475354.

Bozionelos, N., & Singh, S. K. (2017). The relationship of emotional intelligence with task and contextual performance: More than it meets the linear eye. *Personality and Individual Differences, 116*, 206–211. https://doi.org/10.1016/j.paid.2017.04.059.

Brackett, M. A., Rivers, S. E., & Salovey, P. (2011). Emotional intelligence: Implications for personal, social, academic, and workplace success. *Social and Personality Psychology Compass, 5*(1), 88–103. https://doi.org/10.1111/j.1751-9004.2010.00334.x.

Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology, 91*(4), 780–795. https://doi.org/10.1037/0022-3514.91.4.780.

Brenda, G. (2018). In China, female pilots strain to hold up half the sky Retrieved from https://www.reuters.com/article/us-china-airshow-women-pilots/in-china-female-pilots-strain-to-hold-up-half-the-sky-idUSKCN1NA2SP.

Brislin, R. W. (1980). *Translation and content analysis of oral and written materials*. Boston MA: Allyn and Bacon.

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology, 54*(2), 271–296. https://doi.org/10.1111/j.1744-6570.2001.tb00093.x.

Caponecchia, C., Zheng, W. Y., & Regan, M. A. (2018). Selecting trainee pilots: Predictive validity of the wombat situational awareness pilot selection test. *Applied Ergonomics, 73*, 100–107. https://doi.org/10.1016/J.APERGO.2018.06.004.

Cardenas, R. A., Major, D. A., & Bernas, K. H. (2004). Exploring work and family distractions: Antecedents and outcomes. *International Journal of Stress Management, 11*(4), 346–365. https://doi.org/10.1037/1072-5245.11.4.346.

Chan, A. P. C., Wong, F. K. W., Hon, C. K. H., Ali Javed, A., & Lyu, S. (2017). Construction safety and health problems of ethnic minority workers in hong kong. *Engineering, Construction and Architectural Management, 24*(6), 901–919. https://doi.org/10.1108/ECAM-09-2015-0143.

Cheng, Y.-N., Yen, C.-L., & Chen, L. H. (2012). Transformational leadership and job involvement: The moderation of emotional contagion. *Military Psychology, 24* (4), 382–396. https://doi.org/10.1080/08995605.2012.695261.

Cheng, Z., Guo, W., Hayward, M., Smyth, R., & Wang, H. (2021). Childhood adversity and the propensity for entrepreneurship: A quasi-experimental study of the great Chinese famine. *Journal of Business Venturing, 36*(1), 106063. https://doi.org/10.1016/j.jbusvent.2020.106063.

Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology, 94*(5), 1103–1127. https://doi.org/10.1037/a0016172.

Crant, J. M. (1995). The proactive personality scale and objective job performance among real estate agents. *Journal of Applied Psychology, 80*(4), 532–537. https://doi.org/10.1037/0021-9010.80.4.532.

Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology, 29*(1), 1–19. https://doi.org/10.1007/S10869-013-9308-7.

DeJoy, D. M., Schaffer, B. S., Wilson, M. G., Vandenberg, R. J., & Butts, M. M. (2004). Creating safer workplaces: Assessing the determinants and role of safety climate. *Journal of Safety Research, 35*(1), 81–90. https://doi.org/10.1016/j.jsr.2003.09.018.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., & Nikolic, D. (1999). *Situation awareness as a predictor of performance in en route air traffic controllers*. Springfield, VA: D. C. Office of Aviation Medicine Washington Ed..

Dysvik, A., & Kuvaas, B. (2008). The relationship between perceived training opportunities, work motivation and employee outcomes. *International Journal of Training and Development, 12*(3), 138–157. https://doi.org/10.1111/j.1468-2419.2008.00301.x.

Elnaga, A., & Imran, A. (2013). The effect of training on employee performance. *European Journal of Business and Management, 5*(4), 137–147.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society Annual Meeting, 32*(2), 97–101. https://doi.org/10.1177/154193128803200221.

Endsley, M. R. (1999). Situation awareness and human error: Designing to support human performance. *Paper presented at the Proceedings of the high consequence systems surety conference*.

Endsley, M. R. (2000). *Theoretical underpinnings of situation awareness: A critical review*. Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R., & Robertson, M. M. (2000). *Training for situation awareness in individuals and teams*. Mahwah, NJ: Lawrence Erlbaum Associates.

Evans, B., Glendon, A. I., & Creed, P. A. (2007). Development and initial validation of an aviation safety climate scale. *Journal of Safety Research, 38*(6), 675–682. https://doi.org/10.1016/j.jsr.2007.09.005.

Farh, C. I., Seo, M.-G., & Tesluk, P. E. (2012). Emotional intelligence, teamwork effectiveness, and job performance: The moderating role of job context. *Journal of Applied Psychology, 97*(4), 890–900. https://doi.org/10.1037/a0027377.

Fernandes, A., & Braarud, P. Ø. (2015). Exploring measures of workload, situation awareness, and task performance in the main control room. *Procedia Manufacturing, 3*, 1281–1288. https://doi.org/10.1016/j.promfg.2015.07.273.

Franco-Santos, M., Nalick, M., Rivera-Torres, P., & Gomez-Mejia, L. (2017). Governance and well-being in academia: Negative consequences of applying an agency theory logic in higher education. *British Journal of Management, 28*(4), 711–730. https://doi.org/10.1111/1467-8551.12249.

Givehchi, S., Hemmativaghef, E., & Hoveidi, H. (2017). Association between safety leading indicators and safety climate levels. *Journal of Safety Research, 62*, 23–32. https://doi.org/10.1016/j.jsr.2017.05.003.

Griffin, M. A., & Hu, X. (2013). How leaders differentially motivate safety compliance and safety participation: The role of monitoring, inspiring, and learning. *Safety Science, 60*, 196–202. https://doi.org/10.1016/j.ssci.2013.07.019.

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347–358. https://doi.org/10.1037//1076-8998.5.3.347.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: Process versus structural equation modeling. *Australasian Marketing Journal (AMJ), 25*(1), 76–81. https://doi.org/10.1016/j.ausmj.2017.02.001.

Hays, R. D., Hayashi, T., & Stewart, A. L. (1989). A five-item measure of socially desirable response set. *Educational and Psychological Measurement, 49*(3), 629–636. https://doi.org/10.1177/001316448904900315.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55. doi:10.1080/10705519909540118.

Huang, Y.-H., Chen, P. Y., Krauss, A. D., & Rogers, D. A. (2004). Quality of the execution of corporate safety policies and employee safety outcomes: Assessing the moderating role of supervisor safety support and the mediating role of employee safety control. *Journal of Business and Psychology, 18*(4), 483–506. https://doi.org/10.1023/B:JOBU.0000028448.01394.bf.

K.E. Huber S.E. Hill S.M. Merritt Minding the gap: Extending mindfulness to safety-critical occupations Industrial and Organizational Psychology 8 4 2015 699 705 101017/iop.2015.103

Ingram, A., Peake, W. O., Stewart, W., & Watson, W. (2019). Emotional intelligence and venture performance. *Journal of Small Business Management, 57*(3), 780–800. https://doi.org/10.1111/jsbm.12333.

Irwin, A., Caruso, L., & Tone, I. (2019). Thinking ahead of the tractor: Driver safety and situation awareness. *Journal of Agromedicine, 24*(3), 288–297. https://doi.org/10.1080/1059924x.2019.1604279.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386–408. https://doi.org/10.2307/20159208.

Judge, T. A., & Zapata, C. P. (2015). The person–situation debate revisited: Effect of situation strength and trait activation on the validity of the big five personality traits in predicting job performance. *Academy of Management Journal, 58*(4), 1149–1179. https://doi.org/10.5465/amj.2010.0837.

Karimi, L., Leggat, S. G., Bartram, T., & Rada, J. (2020). The effects of emotional intelligence training on the job performance of australian aged care workers. *Health Care Management Review, 45*(1), 41–51. https://doi.org/10.1097/HMR.0000000000000200.

Kell, H. J., Rittmayer, A. D., Crook, A. E., & Motowidlo, S. J. (2010). Situational content moderates the association between the big five personality traits and behavioral effectiveness. *Human Performance, 23*(3), 213–228. https://doi.org/10.1080/08959285.2010.488458.

Khalili, A. (2012). The role of emotional intelligence in the workplace: A literature review. *International Journal of Management, 29*(3), 355–370.

Kim, T. Y., Cable, D. M., Kim, S. P., & Wang, J. (2009). Emotional competence and work performance: The mediating effect of proactivity and the moderating effect of job autonomy. *Journal of Organizational Behavior, 30*(7), 983–1000. https://doi.org/10.1007/s10869-016-9469-2.

Law, K. S., Wong, C.-S., & Song, L. J. (2004). The construct and criterion validity of emotional intelligence and its potential utility for management studies. *Journal of Applied Psychology, 89*(3), 483–496. https://doi.org/10.1037/0021-9010.89.3.483.

Leung, M.-Y., Chan, Y.-S., & Yuen, K.-W. (2010). Impacts of stressors and stress on the injury incidents of construction workers in hong kong. *Journal of Construction Engineering and Management, 136*(10), 1093–1103. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000216.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*(2), 151–173. https://doi.org/10.1207/S15328007SEM0902_1.

Mattingly, V., & Kraiger, K. (2019). Can emotional intelligence be trained? A meta-analytical investigation. *Human Resource Management Review, 29*(2), 140–155. https://doi.org/10.1016/J.HRMR.2018.03.002.

Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence, 22*(2), 89–113. https://doi.org/10.1016/S0160-2896(96)90011-2.

Miao, C., Humphrey, R. H., & Qian, S. (2017). A meta-analysis of emotional intelligence and work attitudes. *Journal of Occupational and Organizational Psychology, 90*(2), 177–202. https://doi.org/10.1111/joop.12167.

Millner, U. C., Brandt, D., Chan, L., Jette, A., Marfeo, E., Ni, P., ... Rogers, E. S. (2020). Exploring counselor-client agreement on clients' work capacity in established and consultative dyads. *Journal of Employment Counseling, 57*(3), 98–114. https://doi.org/10.1002/joec.12148.

Miner, A. G., & Glomb, T. M. (2010). State mood, task performance, and behavior at work: A within-persons approach. *Organizational Behavior and Human Decision Processes, 112*(1), 43–57. https://doi.org/10.1016/j.obhdp.2009.11.009.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*(2–3), 145–165. https://doi.org/10.1080/08959285.1998.9668029.

Muchinsky, P. M. (2000). Emotions in the workplace: The neglect of organizational behavior. *Journal of Organizational Behavior, 21*(7), 801–805. https://doi.org/10.1002/1099-1379(200011)21:7<801::AID-JOB999>3.0.CO;2-A.

Nazir, S., Colombo, S., & Manca, D. (2012). The role of situation awareness for the operators of process industry. *Chemical Engineering Transactions, 26*, 303–308. https://doi.org/10.3303/CET1226051.

Neal, A., & Griffin, M. A. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946–953. https://doi.org/10.1037/0021-9010.91.4.946.

O'Boyle, E. H., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior, 32*(5), 788–818. https://doi.org/10.1002/job.714.

Oliver, T., Hausdorf, P., Lievens, F., & Conlon, P. (2016). Interpersonal dynamics in assessment center exercises: Effects of role player portrayed disposition. *Journal of Management, 42*(7), 1992–2017. https://doi.org/10.1177/0149206314525207.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452.

Rhee, S.-Y., Hur, W.-M., & Kim, M. (2017). The relationship of coworker incivility to job performance and the moderating role of self-efficacy and compassion at work: The job demands-resources (jd-r) approach. *Journal of Business and Psychology, 32*(6), 711–726. https://doi.org/10.1007/s10869-016-9469-2.
Rode, J. C., Mooney, C. H., Arthaud-Day, M. L., Near, J. P., Baldwin, T. T., Rubin, R. S., & Bommer, W. H. (2007). Emotional intelligence and individual performance: Evidence of direct and moderated effects. *Journal of Organizational Behavior, 28*(4), 399–421. https://doi.org/10.1002/job.429.
Rothbard, N. P., & Wilk, S. L. (2011). Waking up on the right or wrong side of the bed: Start-of-workday mood, work events, employee affect, and performance. *Academy of Management Journal, 54*(5), 959–980. https://doi.org/10.5465/amj.2007.0056.
Sætrevik, B. (2013). Developing a context-general self-report approach to measure three-level situation awareness. *International Maritime Health, 64*(2), 66–71.
Sánchez-Álvarez, N., Extremera, N., & Fernández-Berrocal, P. (2016). The relation between emotional intelligence and subjective well-being: A meta-analytic investigation. *The Journal of Positive Psychology, 11*(3), 276–285. https://doi.org/10.1080/17439760.2015.1058968.
Sanchez-Ruiz, M.-J., Mavroveli, S., & Poullis, J. (2013). Trait emotional intelligence and its links to university performance: An examination. *Personality and Individual Differences, 54*(5), 658–662. https://doi.org/10.1016/j.paid.2012.11.013.
Schriesheim, J. F. (1980). The social context of leader–subordinate relations: An investigation of the effects of group cohesiveness. *Journal of Applied Psychology, 65*(2), 183–194. https://doi.org/10.1037/0021-9010.65.2.183.
Siu, O.-L., Phillips, D. R., & Leung, T.-W. (2003). Age differences in safety attitudes and safety performance in Hong Kong construction workers. *Journal of Safety Research, 34*(2), 199–205.
Sneddon, A., Mearns, K., & Flin, R. (2013). Stress, fatigue, situation awareness and safety in offshore drilling crews. *Safety Science, 56*, 80–88. https://doi.org/10.1016/j.ssci.2012.05.027.
Sunindijo, R. Y., & Zou, P. X. (2013). The roles of emotional intelligence, interpersonal skill, and transformational leadership on improving construction safety performance. *Construction Economics and Building, 13*(3), 97–113.
Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88*(3), 500–517. https://doi.org/10.1037/0021-9010.88.3.500.
Tu, C., Guo, J., Hatcher, R. C., & Kaufman, J. C. (2018). The relationship between emotional intelligence and domain-specific and domain-general creativity. *The Journal of Creative Behavior, 54*(2), 337–349. https://doi.org/10.1002/jocb.369.
van Winsen, R., Henriqson, E., Schuler, B., & Dekker, S. W. (2015). Situation awareness: Some conditions of possibility. *Theoretical Issues in Ergonomics Science, 16*(1), 53–68. https://doi.org/10.1080/1463922X.2014.880529.
Wanberg, C. R., Zhu, J., & Van Hooft, E. A. (2010). The job search grind: Perceived progress, self-reactions, and self-regulation of search effort. *Academy of Management Journal, 53*(4), 788–807. https://doi.org/10.5465/amj.2010.52814599.
Wang, J., Zou, P. X., & Li, P. P. (2016). Critical factors and paths influencing construction workers' safety risk tolerances. *Accident Analysis & Prevention, 93*, 267–279. https://doi.org/10.1016/j.aap.2015.11.027.
Westaby, J. D., & Lowe, J. K. (2005). Risk-taking orientation and injury among youth workers: Examining the social influence of supervisors, coworkers, and parents. *Journal of Applied Psychology, 90*(5), 1027–1035. https://doi.org/10.1037/0021-9010.90.5.1027.
Wong, C.-S., & Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *The Leadership Quarterly, 13*(3), 243–274. https://doi.org/10.1016/S1048-9843(02)00099-1.
Xie, J. L., Roy, J.-P., & Chen, Z. (2006). Cultural and individual differences in self-rating behavior: An extension and refinement of the cultural relativity hypothesis. *Journal of Organizational Behavior, 27*(3), 341–364. https://doi.org/10.1002/job.375.
Zhang, J., Ding, W., Li, Y., & Wu, C. (2013). Task complexity matters: The influence of trait mindfulness on task and safety performance of nuclear power plant operators. *Personality and Individual Differences, 55*(4), 433–439. https://doi.org/10.1016/j.paid.2013.04.004.
Zhao, X., Hwang, B.-G., & Gao, Y. (2016). A fuzzy synthetic evaluation approach for risk assessment: A case of singapore's green projects. *Journal of Cleaner Production, 115*, 203–213. https://doi.org/10.1016/j.jclepro.2015.11.042.
Zhao, X., Wu, P., & Wang, X. (2018). Risk paths in bim adoption: Empirical study of china. *Engineering, Construction and Architectural Management, 25*(9), 1170–1187. https://doi.org/10.1108/ECAM-08-2017-0169.

Zhongmin Wang is a PhD candidate in College of Business, Law & Governance at James Cook University, Australia. His current research is focused on safety performance and behaviors in the aviation context. His research interests also include emotional intelligence, training, career development, and work and non-work wellbeing. His work has been published in *Journal of Organizational Behavior*, *Journal of Vocational Behavior*, *Personality and Individual Differences*, among others.

Zhou Jiang is an Associate Professor in College of Business, Government and Law, Flinders University, Australia. He received his PhD in management from Macquarie University, Sydney, Australia. His research interests include knowledge management, emotions at work, career development, organizational behavior, workplace safety, and work and subjective wellbeing. He currently serves as an Associate Editor for *Applied Psychology: An International Review*, and seats on editorial boards of many leading journals, such as *Journal of Organizational Behavior*, *Journal of Vocational Behavior*, and *Human Resource Management Journal*, and *Australian Journal of Management* where he also publishes often.

Anna Blackman is an Associate Professor College of Business, Law & Governance at James Cook University, Australia. She obtained her PhD in management from James Cook University. Her areas of expertise include business coaching, human resource management, career development, rural and regional leadership, and workplace safety. She has published extensively in these areas many peer-reviewed journal articles, book chapters, and case studies in areas such as business coaching, workplace learning, employee training. Her recent work has appeared in *International Journal of Training and Development* and *Human Resource Development Review*.

Journal of Safety Research 78 (2021) 303–313

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Making safety training stickier: A richer model of safety training engagement and transfer

Tristan Casey [a,*], Nick Turner [b], Xiaowen Hu [c], Kym Bancroft [d]

[a] Safety Science Innovation Lab, Griffith University, Brisbane, Qld, Australia
[b] Haskayne School of Business, University of Calgary, Calgary, AB, Canada
[c] QUT Business School, Queensland University of Technology, Brisbane, Qld, Australia
[d] Serco Asia Pacific, Brisbane, Qld, Australia

ARTICLE INFO

Article history:
Received 5 February 2021
Received in revised form 23 March 2021
Accepted 4 June 2021
Available online 25 June 2021

Keywords:
Learning and development
Safety training evaluation
Safety refresher training
Training relapse prevention

ABSTRACT

Introduction: Compared to other types of occupational training, safety training suffers from several unique challenges that potentially impair the engagement of learners and their subsequent application or "transfer" of knowledge and skills upon returning to the job. However, existing research on safety training tends to focus on specific factors in isolation, such as design features and social support. The aim of this research is to develop an overarching theoretical framework that integrates factors contributing to training engagement and transfer. *Method:* We conducted a comprehensive qualitative review of safety training research that was published between 2010 and 2020. We searched Web of Science, Scopus, and Google Scholar, yielding 147 articles, and 38 were included. We content analyzed article summaries to arrive at core themes and combined them with contemporary models of general occupational training to develop a rich model of safety training engagement and transfer. *Results:* We propose that training engagement is a combination of pre-training factors such as individual, organizational, and contextual factors, that interact with design and delivery factors. Safety training engagement is conceptualized as a three-component psychological state: affective, cognitive, and behavioral. Organizations should prioritize pre-training readiness modules to address existing attitudes and beliefs, optimize the safety training transfer climate, and critically reflect on their strategy to design and deliver safety training so that engagement is maximized. *Conclusions:* There are practical factors that organizations can use before training (e.g., tailoring training to employees' characteristics), during training (e.g., ensuring trainer credibility and use of adult learning principles), and after training (e.g., integrating learned concepts into systems). *Practical Applications:* For safety training to 'stick,' workers should be affectively, cognitively, and behaviorally engaged in learning, which will result in new knowledge and skills, improvements in attitudes, and new safety behaviors in the workplace. To enable engagement, practitioners must apply adult learning principles, make the training relevant, and tailor the training to the job and individual needs. After training, ensure concepts are embedded and aligned with existing systems and routines to promote transfer.

1. Introduction

Safety training is a core component of modern safety management. The goal of safety training is to provide workers with safety knowledge and motivation, encourage them to perform safety-relevant behaviors more often and more effectively, and ultimately contribute to a reduced risk of injury through safety behaviors (Burke et al., 2006; Griffin & Neal, 2000). From a contemporary safety science perspective, safety training also helps to improve organizational resilience (i.e., the ability of a system to succeed under varying conditions; Hollnagel, 2011) by equipping workers with improved capabilities to anticipate, respond, learn, and monitor. For instance, Malakis, Kontogiannis, and Kirwan (2010) investigated the role of safety training in air traffic control and showed that equipping operators with cognitive strategies contributes to overall system resilience. However, when safety training is poorly designed and executed, the consequences can go beyond the loss of financial and human resource investment; lives can be lost, errors made, and productivity reduced when safety training fails to transfer from the learning environment back to the workplace (Burke et al., 2006).

* Corresponding author.
E-mail address: tristan.casey@griffith.edu.au (T. Casey).

https://doi.org/10.1016/j.jsr.2021.06.004
0022-4375/

According to Krauss, Casey, and Chen (2014), safety training possesses unique challenges that are different from other types of occupational training. First, safety behaviors tend to be highly routinized and regulated, thus highly resistant to change. Safety training programs are often mandated by regulators and clients/-customers, meaning that there is a reduced sense of choice and self-determinism for the organization as well as the attendees. The motivation to engage with the safety training and its subsequent transfer might be further hindered by the bureaucratization of safety training (Dekker, 2019; Smith, 2018), where multiple and sometimes redundant or irrelevant training programs are mandated to employees. Finally, some of the knowledge and skills taught during safety training may only be used in emergency settings, increasing decay over time as there are limited opportunities to apply the knowledge. Transfer of emergency training is critical to safeguard life and minimize damage to infrastructure and assets. It has recently been studied in the context of immersive virtual reality technologies, which appear to increase learning of new skills and stimulate behavioral change and transfer (Feng et al., 2018; Wang et al., 2014).

Nevertheless, empirical evidence also indicates that when pre-training factors, features of training design and delivery, contextual factors, and post-training factors are considered, safety training can successfully reduce injuries and incidents at work, as well as promote better safety performance (Brahm & Singer, 2013; Robson et al., 2012). However, prior studies tend to focus on a narrow set of factors, such as social support (Freitas et al., 2017, 2019), personality, and motivation (Lingappa et al., 2020). Currently, there is no model that integrates multiple contributing factors and their impact on the safety training process, and that would help to develop a holistic and nuanced understanding of how training engagement and transfer could be facilitated (Krauss et al., 2014). Although the wider occupational training literature has generated integrative models (e.g., Ford, Baldwin, & Prasad, 2018; Sitzmann & Weinhardt, 2018), given the unique characteristics of safety training, the direct application of these models to safety might be inappropriate (Krauss, 2005; Krauss et al., 2014). For instance, in a safety training context, learners may hold attitudes toward safety or possess a safety "locus of control," there may be limited opportunity to apply learnings (e.g., for emergency scenarios), or safety training can be either mandated or voluntary, which might affect learner engagement.

In this paper, we synthesize these two separate lines of research to develop an integrative model for safety scientists and practitioners, present key theoretical insights, and identify practical opportunities to improve safety training engagement and transfer. We focus on onsite safety training that is delivered as standalone sessions, rather than a sub-section within broader occupational training. Our model draws on two key concepts that are relevant for training effectiveness—training engagement and training transfer—which describe the trainee's engagement with the training during its delivery and the subsequent application and generalization of skills and knowledge in their work setting, respectively. We draw on the latest training models from the occupational training literature, which highlight the importance of understanding training from a chronological and multilevel perspective and focus on both training design and delivery factors that create learner engagement, and, importantly, are within organizations' and trainers' reach to influence. With this theoretical backbone, we then enrich our model by incorporating empirical research in safety training from the past decade. In doing so, we respond to recent calls for more "consumer-centric" research that enables practitioners to design, deliver, and measure more effective training activities (Baldwin, Ford & Blume, 2017).

We start by providing a high-level overview of our proposed model and then discuss each of its components in detail. In the discussion, we focus on recommendations to improve safety training effectiveness through pre-training communication, using more engaging and impactful learning strategies, and integrating safety training into organizational systems and processes to provide insights to learners about when and how to apply their safety training.

2. Model development

To develop an integrative model (see Fig. 1), we first identified two key concepts that produce overall training effectiveness: training engagement and training transfer. Training engagement is a relatively understudied construct and either not explicitly defined by scholars (e.g., Sitzmann & Weinhardt, 2018) or defined implicitly through measurement proxies, such as number of levels or content accessed as part of a training program (Harvey, Balzer, & Kotwicki, 2020). Consequently, we conceptualize safety training engagement as the combination of optimal cognitive, emotional, and behavioral activity that drives motivation to learn and other training-approach behaviors. Drawing from the educational psychology literature, we concur that learner engagement has a multidimensional nature, with cognitive engagement considered as the mental effort invested in the training to think about and attend to the materials, behavioral engagement as actively participating in the training program, and affective or emotional engagement as a positive mental state in relation to the learning task at hand (Ben-Eliyahu et al., 2018). Using training engagement in a safety context fills a void in the research, which tends to model pre-training motivation, learning, and post-training motivation as the primary variables of interest. Safety training engagement as a within-training construct allows measurement and evaluation to venture into the learning process itself and enables diagnosis of in-situ effects of training design and delivery factors.

Training transfer is the focal outcome of training events and refers to the generalization and maintenance of learned knowledge and skills (Ford et al., 2018). Theoretically, training engagement is a proximal antecedent of learning that also affects training transfer. Without the experience of engaging with the training, training transfer cannot occur. Given that learning is affected by task engagement (Kanfer & Ackerman, 1989) and is malleable through the design and delivery features of training (Kanfer, 1990), engagement should be included within an enriched model of safety training transfer.

Adapting the seminal model outlined by Baldwin and Ford (1988) and recently consolidated by Ford et al. (2018), we propose three major categories that impact training transfer via training engagement: trainee factors that are specific to the individual (e.g., personality, beliefs, pre-training motivation), training factors that relate to how the training is designed and delivered (e.g., the level of engagement, use of adult learning principles), and contextual factors that arise from the team or organization (e.g., safety climate, safety training transfer climate). Outside safety training, meta-analytic studies have demonstrated that these categories of factors are most strongly related to training transfer (e.g., Blume et al., 2010). Within the context of safety training, design and delivery factors have been shown to be important in predicting transfer (Burke et al., 2006), and there is some evidence for contextual factors like safety training transfer climate (Krauss, 2005). Regarding personal characteristics, trainees' motivation (even after accounting for bias introduced by same-source and same-measurement contexts) has, among others, the highest correlation with learning and transfer (Blume et al., 2010). Consequently, pre- and post-training motivation is one of the most important variables that training designers and deliverers can target.

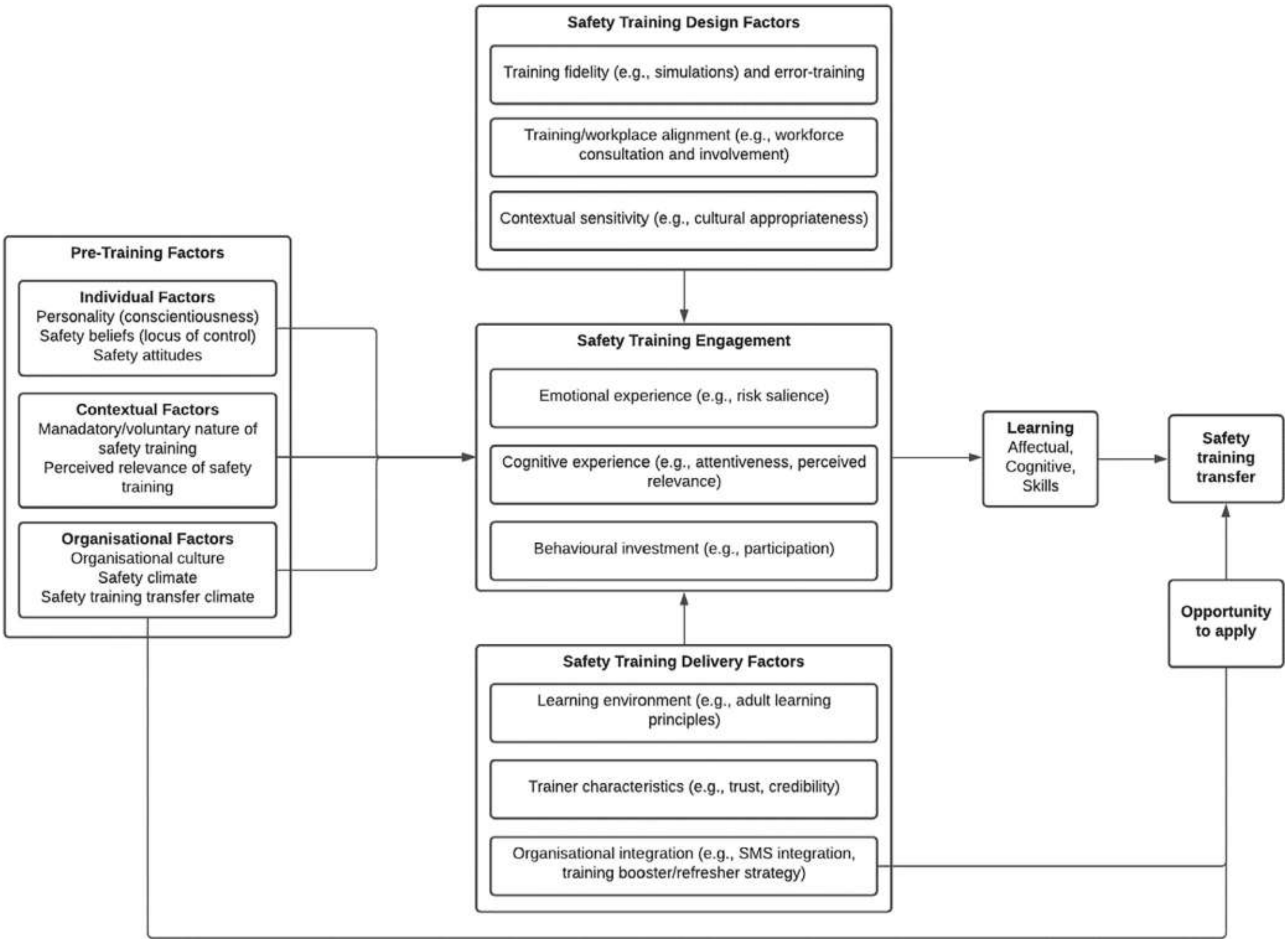

Fig. 1. Model of safety training engagement and transfer.

3. Review method

To develop a safety-contextualized understanding of each category, we scanned the safety training literature from 2010 onwards, focusing on specific and separate safety training delivered onsite or in occupational contexts rather than in classroom environments. Narrow search phrases including "safety training engagement," and "safety training transfer" were used across Scopus and Web of Science databases. An additional scan of Google Scholar search results (using the same search terms) was undertaken, and any unique articles not included in the first scan considered as part of the review. In total, 147 articles were sourced through these search activities. After removing duplicates and irrelevant articles by scanning the abstracts (i.e., published in a language other than English, non-peer reviewed), 57 articles were identified and reviewed in depth. A total of 38 articles had information that was relevant to the scope of this paper, as identified by the first author and a research assistant. Each of the 57 initial papers was evaluated as being relevant, irrelevant, or undecided by reviewing them in detail independently. Disagreements were then discussed and resolved. Articles were included if they (a) provided unique information not captured or repeated by other articles and (b) focused specifically on employee safety rather than other forms of safety (e.g., food safety). In the next section, we summarize this body of research across the categories formed by the safety training engagement and transfer model.

3.1. Safety training engagement

To transfer safety training successfully, trainees must learn and retain information. Safety training engagement is defined as a three-component psychological construct involving affective, cognitive, and behavioral elements in combination to optimize the energy and motivation directed towards the learning task. Such motivation to learn during safety training results in greater retention of knowledge, greater uptake of attitudinal change, and higher intention to engage in learned behaviors. Thereafter, training engagement promotes transfer through learning, and also via post-training motivation to apply/transfer (Naquin & Holton, 2002).

Some emerging findings from the safety training literature suggest that emotions play a central role in highlighting the salience of hazards and their corresponding risk levels. Cuing emotions such as fear or dread during safety training may deepen the learning process and encourage training transfer post-training (Burke et al., 2011). Cognitively, learners must be sufficiently motivated during the training process, as evidenced by attentional regulation (Kraiger et al., 1993). Behaviorally, learners experience engagement in safety training through active participation and involvement in learning activities (Casey et al., 2018). In safety training, engagement is likely to enhance learning because it can overcome suboptimal pre-training factors like negative safety attitudes or a weak safety climate.

3.2. Safety training transfer

Typically, safety training transfer is defined as the application of learned skills, generalization to work scenarios, and maintenance over time (Baldwin & Ford, 1988; Ford & Weissbein, 1997). Other aspects of training transfer may be important for safety training evaluation. Quantifying or qualifying the degree of transfer on the spectrum ranging from negative to positive would provide a useful success metric because it would highlight any intended interactions or variations in trained behavior that could actually decrease safety performance. Also, understanding whether the safety training only transfers to very similar contexts or also to very discrepant and diverse contexts may offer insights into why safety training can fail to produce the desired results, for example, during emergency events where conditions may be unique or short-lived.

As discussed previously, safety training transfer tends to be measured directly by attendance figures and evaluation forms and indirectly via injury and incident statistics. Safety training transfer should instead be measured using more nuanced metrics that could be staged to represent a growing embedding and exploration of additional skills and practices. For instance, initial transfer intention should be measured by utility reactions and motivation to apply learning. Short-term transfer can be measured by near-transfer evaluation—in other words, did trainees apply the learning in similar environments to what was trained? Mid-term evaluation should concentrate on far-transfer and long-term retention of key skills and concepts. Long-term evaluation can identify whether the training overall produced positive, zero, or neutral effects and also measure the factors that either supported or hindered the training being applied. Formative and summative evaluations can also be used together to identify improvements and lessons learned for future safety training designs.

3.3. Pre-training factors

Pre-training factors include trainee factors, contextual factors, and organizational factors. Trainee factors include personality traits, attitudes, and beliefs. Contextual factors specific to safety training (Krauss et al., 2014) include the mandatory or voluntary nature of the training (with mandatory training potentially reducing pre-training motivation to learn) and the perceived relevance of the safety training (particularly in the case of safety training content that is mandated by regulators or other government bodies). Finally, an organization creates a context in which safety training occurs. For instance, safety climate conveys the value and importance of safety behavior, which could dramatically affect pre-training characteristics such as safety attitudes and training motivation, affecting the level of training engagement (Krauss et al., 2014). Such factors can also be relevant after training by directly impacting post-training motivation to learn and subsequent training transfer. After training, the safety training transfer climate sends a strong signal to employees about whether the organization values the application and/or practice of what was learned during safety training (Krauss, 2005), again impacting training transfer.

3.3.1. Individual factors

3.3.1.1. Safety beliefs. Beliefs are internal schema or models around meaning and knowledge or, more practically, "how the world works" (Fishbein & Raven, 1962). A safety-related belief can therefore be a conviction about any aspect of how safety should be managed in an organization, such as implicit accident causation models (i.e., what causes an accident), what constitutes a hazard, and core self-beliefs regarding safety-specific locus of control. For instance, Krauss (2005), in an unpublished doctoral thesis, explored the interactive effects of work locus of control on safety training transfer. Although the findings were non-significant, they trended in the predicted direction and it was proposed that a safety-specific locus of control construct may have been more appropriate.

Nevertheless, in a training context, deep-seated beliefs are virtually impossible to shift and can act as a handbrake on any learning or desired change, reducing training engagement (Murphy & Mason, 2006). Understanding beliefs prior to safety training could enable a more nuanced and tailored approach to safety training whereby different 'streams' of learning are identified to target or amplify certain pre-existing beliefs.

3.3.1.2. Safety attitudes. Attitudes are evaluations toward people and objects in their environment (Ajzen, 2005). Safety attitudes can be a positive or negative assessment of a safety-specific object, person, or action (Lingard & Rowlinson, 2005), such as engaging in a particular safety practice, using a safety tool or process, or a judgment of safety personnel. Safety attitudes are relevant to safety training because they may affect not only behavior intention to engage fully with and participate in the training but also individuals' intentions to apply what was learned (Krauss, 2005). Safety attitudes can also change over the course of a training event, meaning that an initially unfavorable or resistant attitude could evolve into a more conducive one by the conclusion of the event. This has implications for the design of training; an initial "safety training readiness" module or specific structuring of the training may have a positive relationship with subsequent engagement during training (Casey, Krauss, & Turner, 2018).

3.3.1.3. Personality. Safety training is delivered in strong regulatory/compliance framework with a moral overtone. As such, it is reasonable to expect that conscientiousness may be the most relevant predictor. In the general training transfer literature, Huang and Bramble (2016) found that trait, state, and task-contingent conscientiousness affected learning and training transfer, detail-oriented and duty-bound employees may feel an obligation to pay attention and engage with safety training exercises, and also feel a stronger need to apply safety training afterwards. Only limited work has been done in this space. One study by Lingappa, Kiran, and Mathew (2020) explores the role of Big 5 personality traits in predicting safety training motivation to learn, motivation to transfer, and self-reported training transfer among employees from an Indian chemicals company. The authors found that conscientiousness and locus of control positively affected safety training transfer, whereas risk-taking propensity was negatively associated with transfer. In addition, Hogan and Foster (2013) conceptualized a facet-level personality construct called 'Trainable,' which referred to a person's tendency to accurately estimate their safety competence, openness to feedback, and engagement in learning. An association was found between this construct and supervisor safety performance ratings, but the study did not explore the relationships with safety training engagement or outcomes.

3.3.2. Contextual factors

3.3.2.1. Mandatory/voluntary safety training. A lot of safety training is typically mandated by government bodies, such as health and safety regulators, or required by client/principal contractor organizations in industries like construction (Krauss et al., 2014). The general transfer training literature suggests mandatory/voluntary training conditions might interact with personality traits (i.e., learning goal orientation) to influence learning outcomes (Gegenfurtner et al., 2016). Following the same logic, we propose that existing safety attitudes and beliefs may interact with voluntary/mandatory status to influence training engagement and training transfer. When safety attitudes and beliefs are favorable, mandatory status may matter less than if they are negative. Given the prevalence of mandatory safety training, further research is needed to explore these relationships.

3.3.2.2. Perceived relevance. Many safety training programs suffer from a lack of domain specificity; for instance, many programs are designed as a one-size-fits-all solution where the audience needs and job characteristics are not taken into account (Casey, Krauss, & Turner, 2018). An example is when corporate areas from high-risk organizations participate in safety training that is primarily designed for their blue-collar operational colleagues. In these situations, the relevance of safety training is likely to result in audience disengagement and reduced uptake of learning. At the least, relevance will likely interact with training design factors to produce higher or lower attendee engagement in learning.

3.3.3. Organizational factors

3.3.3.1. Organizational culture. Fundamental beliefs and assumptions (Schein, 2010) relating to safety, such as what constitutes a hazard, what levels of risk are tolerable, and the nature of human relationships, are likely to affect training engagement, application and transfer. Studies in general training transfer have suggested that fundamental assumptions affect the attention and encoding of new information—contradictory or controversial ideas may even be ignored (Bunch, 2007). In relation to safety, training content that is misaligned with the dominant safety culture (i.e., assumptions surrounding such matters as hazards, risks, and interpersonal relationships, such as speaking up and stopping an unsafe act among coworkers) may be dismissed or ignored, leading to lower learning engagement. Given that the definition of safety is subjective (Dekker, 2019) and influenced by both individual- and group-level beliefs, the dominant culture operating within the trainee's local context could act as an information filter, highlighting or emphasizing some information and dismissing or downplaying other information that is misaligned or in conflict with the norms, values, and beliefs endorsed by the group. Culture is likely to also play a role in subsequent transfer because dominant norms may decrease intention to apply behaviors if they are not established or embedded in the team. Potentially, safety training should be conducted with intact teams to ensure social norms change.

3.3.3.2. Safety climate. Safety climate refers to the perceived value and importance of safety in an organization, as inferred through perceptions of policies, procedures, and practices (Zohar, 2010). Safety climate may affect the transfer of safety training by providing a broader social context that infers the priority of such training and the overall importance of safety. Indeed, Burke, Chan-Serafin, Salvador, Smith, and Sarpy's (2008) investigations across 68 organizations revealed that safety climate moderated the safety training–incidents relationships, with more positive safety climates amplifying reduction of incidents. Safety climate is a powerful contextual factor because it can influence pre-training factors like safety attitudes and motivation, within-training factors like motivation to learn (impacting engagement), and also post-training factors like motivation to apply (Griffin & Curcuruto, 2016).

3.3.3.3. Safety training transfer climate. In the general training literature, training transfer climate refers to the overarching priority of training, based on workers' perceptions of how valued and important it is to apply what is learned, with a focus on social support to apply (Baldwin & Ford, 1988; Rouillier & Goldstein, 1993; Tracey, Tannenbaum, & Kavanagh, 1995). It is distinct from safety climate because the former is a narrow perception of the importance and value of safety training, whereas safety climate refers to the higher order value of safety in general. Safety training may have a different priority compared to regular training. It may be considered less important or potentially wasting experienced employees' time. Thus, the social transfer environment in which safety training occurs is crucial. In an unpublished thesis, Krauss (2005) developed a measure of safety training transfer climate, finding that factors such as supervisor recognition for applying safety training, management encouragement of safety training, and opportunities to apply safety training indeed created a shared social context that influenced the transfer of safety knowledge.

3.4. Safety training design factors

Referring to the nature of how the training is created, training design factors are a prime target to improve safety training transfer. An interesting development in transfer research and practice is the use of technology in safety training to increase the level of fidelity, potentially triggering strong emotional and cognitive responses to embed learning (e.g., Bhandari & Hallowell, 2017). The alignment between "work as trained" and "work as done" is also likely to influence learner engagement and transfer. Involving workers in the design and development of safety training is likely to create greater ownership and enhance the learners' engagement. Finally, much safety training is delivered to diverse groups in terms of literacy and cultural background (e.g., Arcury et al., 2014). Ensuring the training design reflects these important differences will ensure that learners are suitably engaged.

3.4.1. Training fidelity and error training

Technology holds much promise when it comes to improving safety training engagement via fidelity. Much of safety is practical and hands-on, involving the development of skills such as hazard recognition, implementation of control measures, and safe work practices. Technology that simulates real-world conditions and gives a more nuanced and lifelike representation of safety scenarios is not only more engaging but also associated with better learning outcome in comparison to traditional classroom didactical training. In safety training, emerging evidence suggests that immersive injury simulations can induce strong emotions, which in turn increases interest in training content and contributes to risk-averse behavioral transfer Bhandari, Hallowell, & Correll (2019). Subsequent research with virtual reality in construction settings by Bhandari and colleagues (2020) found that emotions induced by simulated environments predicted higher risk perception and more effective safety decision-making.

Mixed-reality is an example of such technological advancement. In Hasanzadeh, de la Garza, and Geller's (2020) recent study, workers installed shingles onto a sloped roof (physical task) while at a simulated height of two meters (virtual environment). The experimenters manipulated various hazards and measured physiological markers to evaluate "presence" or mindful attention to the task. Others have investigated the use of building information modeling (BIM) in construction safety training contexts and found that compared to lectures, BIM-supported training resulted in greater knowledge transfer (Ahn, Kim, Park, & Kim, 2020). Finally, Chittaro, Corbett, McLean, and Zangrando (2018) explored the use of virtual reality to improve aircraft safety procedures. Virtual-reality-trained participants exhibited significantly faster life-vest donning and fewer errors than traditionally-briefed ones. Virtual reality (VR) technologies seem to have special relevance and utility in the context of emergency training, a form of safety training that concentrates on evacuation and other drills in response to incidents. In a systematic review published in the information sciences literature, Feng and colleagues (2018) put forward a model to inform the design and evaluation of virtual technologies in this setting. The study conceptualizes this use of VR equipment as a 'serious game' whereby deep learner engagement (as represented by our model through emotional, cognitive, and behavioral components) results in deeper learning and higher rates of post-training application. From a neuroscience perspective, inducing emotions through immersive virtual environments may facilitate stronger neuronal connections and associations between stimulus–response via the amygdala. Further, deeper learning of motor movements and physical skills required (e.g., operating a fire extinguisher) may be stimulated in the cerebellum through the immersive and feedback-rich environment of VR (as shown in healthcare and rehabilitation settings; e.g., Kim, Schweighofer, & Finley, 2019; Mao et al., 2014).

Gamification can also boost training engagement by enabling more detailed performance feedback to occur. Highlighting the use of technology to drive personalized training feedback, Jeelani, Han, and Albert (2018) evaluated the use of eye-tracking technology in construction, with hazard detection likelihood scores being used to improve safety performance among workers. Liang, Zhou, and Gao (2019) explored gamification in the mining industry and found that an immersive gaming environment using off-the-shelf equipment (i.e., HTC Vive and Unity 3D engine) improved miners' safety awareness and risk-aversion. Similarly, in one systematic review conducted in the construction industry, the authors found that gamified training as well as other computer aided technologies (simulations, augmented reality, virtual reality, mixed-reality) improved trainee engagement (Gao, Gonzalez, & Yiu, 2019).

Online safety training is gaining increasing traction, especially for mandatory site inductions in the high-risk domains such as construction and manufacturing, and represents a vehicle for greater individualization of safety training by allocating courses based on training needs analysis (Trout, 2016). Limited research has been done on online safety training, but one study has shown differences between older and younger workers (Wallen & Mulloy, 2006), which may be due to differing levels of computer self-efficacy or anxiety (Chen, 2017), innovativeness (Jokisch et al., 2020), and human–computer interaction factors like ease of use (Tsai et al., 2017)—all of which have been associated with age (older age adversely affects these variables). Early research has shown that less information-dense and visualized online safety training (e.g., videos, graphics, audio files) performed the best at stimulating learning across age groups (Wallen & Mulloy, 2006). Others have promoted the utility of blended learning for safety, which combines online modules with face-to-face activities. Specifically, Stuart (2014) found that for furniture manufacturing trainees, the anxiety and intimidation of the workplace setting, and practical exercises could be reduced through giving trainees access to prior online safety modules. Overall, more research is needed to understand how to design and deliver effective online safety training; however, general user experience principles founded on aesthetics, usability, and usefulness as per mainstream technology acceptance and computer anxiety models are good starting points for practitioners.

Indeed, blended learning is widely considered to be the most popular and effective mode of corporate training due to its flexibility, efficiency, and stimulation of continuous learning among students (Rasheed, Kamsin & Abdullah, 2020). Yet, and of relevance to safety training considering pre-training factors like existing safety beliefs and attitudes, blended learning has four key challenges: incorporating flexibility, stimulating student interaction, facilitating deep learning, and fostering an effective learning climate (Boelens, De Wever & Voet, 2017). As no systematic review of online or blended safety training research yet exists, research is required to clarify how the design of online safety training can be optimized. Cross-disciplinary collaboration among safety scientists, educators, and human–computer interaction specialists will be required.

Burke and colleagues (2011) conducted a meta-analysis on existing safety training studies and found that when the hazard (s) to be trained against were deemed higher risk, safety training was seen as more engaging. This effect occurs because the salience of hazards induces "dread" and elevates subjective perceptions of risk, resulting in higher learning engagement. These meta-analytic results are supported by several subsequent studies involving technology. For instance, Namian, Albert, Zuluaga, and Behm's (2016) study in construction found that engaging training combined with salient depictions of hazards resulted in more effective transfer. Gummesson (2016) discovered the utility of QR (quick response) codes in safety training for students, allowing them to view more detailed and vivid imagery of hazards. Loosemore and Malouf (2019) recommended that construction safety training should use more engaging and salient depictions of risk—ideally using technology.

Much has been written about the benefits of training through error exploration—it enables people to create deeper and more robust knowledge schema and develop more accurate mental models of how underlying processes or work systems operate (Keith & Frese, 2008). In such general error training, the protocol is to allow trainees to actively explore target activities or processes, encourage them to make errors and recover performance, and provide constructive and positive feedback. In high-risk settings, challenges such as automation create intractable systems that can escape employees' capacities to comprehend and develop an accurate situational awareness (Hollnagel, Wears, & Braithwaite, 2015). Using error-based learning may counter this issue and allow operators to learn how complex systems work and, importantly, how they fail.

Developing skills "at the edge" and even over the boundary of safety has been identified as a key strategy to make further improvements in safety performance in today's complex and dynamic work environment (Rasmussen, 1997). In health care, Browne et al. (2019) found that error training combined with bias-reduction strategies were effective at improving health-care providers' critical-thinking skills and subsequent error-management and safety performance. Finally, Choi, Ahn, and Seo (2020) used virtual reality to give forklift drivers the opportunity to make and learn from errors during driving, which boosted their situational awareness and safety performance. More work is needed to determine the impact of error-based learning on safety training engagement, particularly in the context of complex systems.

3.4.2. Training/workplace alignment

In many countries, employers are required by law to consult with workers about hazards and to inform the development of events like safety training (e.g., the Work, Health & Safety Act, 2011 in Australia). The purpose of this consultation is to ensure that the expertise of workers, who do the job on a daily basis, is incorporated into safety decisions made by management and ultimately results in more effective safety interventions (Safe Work Australia, 2018). When it comes to safety training, involving workers in its design and development may increase their engagement. One study in the agriculture industry found reduced training impact due to inadequate consideration of farmers' daily tasks, work context, and learning needs (Holte & Follo, 2018). The training was described as abstract, theoretical, and out of touch with farmers' needs and language. Others have found similar results, possibly due to a lack of consultation and involvement of the audience in the program design (Casey et al., 2018). In a recent study, Vigoros, Caffaro, and Cavallo (2020) tested a user-centered design model to develop visual safety tools for migrant farming workers. A significant difference in training satisfaction was found between the user-centered design group and the control group, highlighting the importance of involving workers in the development of safety training if learner engagement is to be maximized.

3.4.3. Contextual sensitivity

With an increasingly global workforce in safety–critical settings (Clarke, 2003), differences in national culture or ethnic background are likely to influence the effectiveness of safety training. National culture may affect learner engagement. Differences in language ability and interpretation of training materials are likely issues to explore. In a study investigating the design of multicultural safety training, Kovacic and Cunningham (2019) found that engaging delivery, combined with purposeful efforts to instill cultural respect into the training environment, and hands-on practical skill development and assessment activities tended to produce the best learning outcomes for multicultural workforces. Digging deeper into cultural beliefs, Yorio, Edwards, and Hoeneveld (2019) put forward several safety-specific propositions around Hofstede's (1980) cultural dimensions. Of relevance to safety training transfer, dimensions such as uncertainty avoidance (i.e., extent to which groups rely on norms and rules; Hofstede, 1980) may affect the transfer of certain types of training. For cultures with high uncertainty avoidance, safety training that focuses on legislation, standards, and rules may be more accepted and, hence, more likely to be transferred and applied. Burke et al. (2008) found that uncertainty avoidance was negatively related to safety training transfer in a meta-analytic study, but the type of training was not explored.

3.5. Training delivery factors

When corporate training uses principles such as adult learning (Knowles, 1996), transfer is improved (Burke & Hutchins, 2007). Adult learning includes strategies such as involving workers in the training program, scaffolding or building on existing knowledge, and encouraging adults to set their own learning tactics. In safety settings, particularly construction, safety training is often described as mundane, standardized, and infrequently incorporates adult learning principles (Bhandari et al., 2019). A meta-analysis conducted by Burke and colleagues (2006) found that more engaging and dynamic safety training results in better engagement and transfer. More recently, a study in the construction industry found that more engaging safety training methods resulted in attendees identifying more hazards and perceiving higher risk than those who attended less engaging training (Namian, Albert, Zuluaga, & Jaselskis, 2016). Again, in construction, Eggerth, Keller, Cunningham, and Flynn (2018) found that safety training that included narratives and discussion questions produced better learning than those without engaging methods.

Several individual studies in high-risk settings have replicated these results, showing that engaging in safety training results in higher risk salience, greater learning, and boosted application of learning on return to the workplace (e.g., Eggerth et al., 2018; Namian, Albert, Zuluaga, & Jaselskis, 2016). Trainers may carry different levels of credibility in the eyes of attendees, depending on whether they have an operational background or a safety science background. Indeed, operational safety professionals and trainers may more readily build trust and rapport with workers and deliver more enriched examples of targeted behaviors. Finally, integration of organizational processes and systems with safety training could impact not only the learning process but also the post-training motivation to learn through reinforcement and boosting of learned concepts and skills.

3.5.1. Learning environment

Active participation during safety training has long been established as a predictor of transfer. Through their seminal meta-analysis, Burke and colleagues (2006) found that safety training designed with adult learning principles and encouraged a high degree of involvement and participation tended to be more successful. Participating in safety training is likely to enhance its effectiveness because much of it is skill-based and so requires some behavioral investment (Krauss et al., 2014). By actively participating in safety training, attendees are more likely to develop behavioral routines and refine their performance, ideally in some sort of constructive feedback environment.

In a qualitative case study exploring corporate trainers' strategies to engage attendees, Arghode and Wang (2016) discovered that trainers use several different strategies. These strategies include being trainee-centered (e.g., providing interesting and relevant examples), using entertaining and interesting instruction techniques (e.g., humor), using a diverse range of instructional types (e.g., kinesthetic, didactic), encouraging trainees to participate in the session (e.g., role-play), and building rapport early in the session to maintain trust (e.g., an introductory ice-breaker activity). Further research is needed to explore the specific skills and strategies employed by safety trainers to create a positive learning environment that boosts engagement.

3.5.2. Trainer characteristics

Little research has been done on the characteristics of trainers themselves and their impact on safety training engagement. Burke and Hutchins' (2008) qualitative study in the general training literature found that trainers' subject matter knowledge, professional experience, and knowledge of learning principles were important factors. For safety training, trainer credibility may be particularly important. For many workers, safety training is seen as abstract or detached from the lived reality of their jobs (Holte & Follo, 2018). When a trainer is seen as an outsider or non-credible in the eyes of attendees, their willingness to engage in learning may be reduced. This may be related to the development of trust and rapport between trainer and trainee, similar to the concept of therapeutic alliance in counseling psychology (Elvins & Green, 2008). An element of trust is perceived competence or ability (Mayer, Davis & Schoorman 1995). Butler, Reed, and Le Grice (2007) found that vocational training in small business settings was important for knowledge transfer and improved performance.

3.5.3. Organizational integration

Integrating safety training concepts and practices within an existing safety management framework is likely to not only create additional opportunities to transfer, but also send signals regarding the priority and importance of such training, contributing positively to the safety training transfer climate. In the training transfer literature, the presence of an evaluation framework encourages transfer post-training (Hutchins, Burke, & Berthelsen, 2010). Measuring safety training using a combination of "lead and lag" indicators that go beyond training attendance and injuries/accidents is crucial to learn more about what promotes engagement and transfer. Training 'booster' interventions have thus far shown inconsistent effects on long-term transfer in the general training literature. In the general training literature, a strategy borrowed from clinical psychology called "relapse prevention" has been evaluated several times with inconsistent results (Hutchins & Burke, 2006). In safety settings, the results are also mixed. Casey and colleagues (2018) experimented with a training transfer relapse prevention module within the fishing industry. The module consisted of a structured checklist and the opportunity for attendees to brainstorm how they will overcome barriers to transfer. Because the overall training failed to show an effect on outcomes, the impact of the relapse prevention was not discernible. More work is needed to elucidate the impact of relapse prevention in safety training. Regarding booster training, the results are more positive. Kovacic and Cunningham (2019) and Ruttenberg and Rice (2019) explored the effectiveness of refresher training and found that participants apply concepts more often if refresher training is used. Boosters may prevent knowledge decline in safety training given reduced opportunities to apply learnings.

3.6. Opportunity to apply

Having the opportunity to apply training is one of the most important predictors of training transfer (Burke & Hutchins, 2007). Opportunities to apply can either be passive or active. Passive opportunities are when supervisors or managers create time for employees to practice learned skills, such as freeing up work commitments. Active opportunities to apply are when either transfer is directed/encouraged (e.g., practicing safety conversations during a Toolbox Talk) or is required on the job (e.g., an emergency event happens). The opportunity to apply is particularly important for safety training because some types of safety education cannot be directly practiced in the workplace, such as specific emergency events (Krauss et al., 2014). Implementing virtual reality technologies and simulations could be a promising way to provide opportunities to apply safety training in the future, and specifically for emergency training to provide simulated opportunities to use learned skills, maintain knowledge, and increase awareness and vigilance under times of stress (Feng et al., 2018). VR-based training could even be used to give employees opportunities to practice multiple roles during emergency events, providing cross-training and redundancy in the event a person in a critical role (e.g., fire wardens and first-aiders) becomes incapacitated or is unavailable.

4. Discussion

In this paper, we have outlined an enriched safety training transfer model. We reviewed the past decade of safety training literature to inform the development of this model, with a focus on contemporary topics like the use of technology in the safety learning space. This review takes stock of the current safety training landscape with a view to encouraging further research and more effective practice in the design and delivery of safety training, with a view to optimizing learner engagement and subsequent transfer. Several proposed factors specific to safety training have been proposed and warrant further research: trainer credibility, training fidelity, safety management system integration, and others.

4.1. Theoretical implications

This paper highlights a distinction between learner engagement and transfer. In the general training transfer and also safety-specific domain, there is a focus on transfer as an outcome of the 'black box' of intra-training factors. Separating learner engagement from this process and considering the roles of training design and delivery, as well as pre-training factors, may stimulate more nuanced and practical research. Although we know a lot about what predicts safety training transfer, less is known so far about how engagement can be increased, and specifically, which elements of engagement (emotional, cognitive, and behavioral) are most important within the context of different types of safety training.

Safety training warrants a particular focus when it comes to optimizing training transfer. Applying an individually-focused model like the theory of reasoned action or planned behavior (Fishbein & Azjen, 1975) suggests that safety attitudes, norms, and behavioral control will influence intention to use safety training. From an organizational perspective, group-level phenomena like safety culture and climate have been shown to affect training transfer, which points to the importance of thinking globally around training implementation.

Recent work done on the individualization of safety training suggests that there is a complex interaction between personal characteristics (e.g., personality and beliefs), training design and delivery factors, and contextual factors (e.g., safety training transfer climate). Just as the general safety climate literature has started to examine the complex interplay between individual, group, and organizational factors (e.g., Beus, Bergman, & Payne, 2016), safety training transfer research could also benefit from this approach. For example, Beus et al. (2010) found that organizational tenure attenuated safety climate strength in a non-linear fashion—employees with less tenure or who are less open to experience may be less affected by the social context, and so alternative strategies will need to be deployed.

In our model, safety training type is positioned as an important moderating or contextual variable. "Straightforward" safety training that concentrates on declarative knowledge for low-risk hazards is likely best done using didactic and traditional lecture-based methods (Burke et al., 2006). For problem-solving training, weaving in high-fidelity technologies and error-based learning will likely improve transfer outcomes. Further research is needed to examine the configurations of transfer factors required to optimize different types of safety training.

4.2. Practical implications

Several practical implications are apparent following this review. These implications have been arranged into actions that organizations can undertake either prior to, during, or following safety training implementation.

4.2.1. Before training

Before training, it would be advantageous to measure trainees' pre-existing safety characteristics (e.g., attitudes, beliefs) and use this information to stream attendees into different levels or types of training interventions. For instance, if safety attitudes are negative or neutral, an additional high impact and energy module might be effective at creating increased readiness to change and overall engagement in the learning.

We recommend that organizations measure and improve their safety training transfer climate before training is implemented to ensure the conditions for transfer are optimized. Measurement could be done by drawing on a published scale of training transfer, and adapting to the safety context, or using the safety-specific scale developed by Krauss (2005). Improvements to the safety training transfer climate could be achieved through the following interventions: (a) supervisor training that concentrates on pre- and post-training conversations around the value and importance of safety training, (b) targeted communications from senior management referring to the specific safety training and its benefits/importance, (c) developing post-training verification of competence and supportive conversations programs, (d) aligning safety training application with performance recognition programs, and (e) ensuring workers are given time at work to prepare for and practice safety training.

Preparatory communication should be delivered to ensure trainees are familiar with the reasons why the training is being delivered and consider using strategies to induce a learning goal orientation (e.g., framing the messages as a chance for self-betterment or improvement rather than achieving a grade or competency above others). Such work would lay a foundation for positive learner engagement.

Another strategy that organizations could adopt to improve safety training engagement is to develop a pre-training readiness module, which could include: messages of support and encouragement from senior leaders (e.g., video message), activities designed to measure training readiness, such as training anxiety, attitudes toward safety training, and include targeted "mini-interventions," a detailed overview of the training (including its purpose, benefits, and objectives), and, finally, targeted pre-work such as condensed readings of the topic or 1–2 key questions for trainees to consider prior to arriving at the training (e.g., asking attendees to take a photo of what makes them feel safe in their work environment).

Consultation is important to the design of effective and relevant safety training. Specifically, designers should involve workers in the design and development of training content and delivery methods. Resources such as health and safety representatives can be leveraged to keep the consultation manageable and targeted. Such consultation will help to increase the alignment between what is taught during safety training with what practices actually occur on the job (aligning work as imagined with work as done). Higher engagement in the learning is likely to result.

From a broader perspective, the organization should consider coupling safety climate improvement initiatives to major safety training events because the broader team and organizational social context will influence engagement and transfer. There are likely to be positive synergies between safety climate discovery and safety training, particularly if there is alignment between the opportunity areas identified by a climate survey and the areas targeted by training. Finally, organizations should look for ways to "declutter" safety training by removing redundant or irrelevant sections or parts that add little/no value to workers and reduce engagement—this can be achieved systematically by conducting formative training evaluations with a pilot group.

4.2.2. During training

The correct design of safety training can significantly enhance engagement. To optimize learning, designers should consider the type of safety training to be delivered and what combination of transfer factors are more important to optimize transfer performance and maintain cost/benefit efficiencies. For fundamental knowledge and skills, traditional techniques like lectures or group-based learning, the use of narratives or story-telling techniques, and discussion/application questions will be helpful. For hazard recognition, organizations should try to apply technology such as virtual reality and augmented reality. For problem-solving and decision-making, organizations could use error management and simulation training activities. And, finally, for empowerment, we recommend that organizations draw on engaging techniques like role-playing and expert demonstrations, providing examples of effective/ineffective performance, and providing detailed feedback on training performance. Research on this topic is in its infancy, but it may help to use immersive virtual, augmented, or mixed-reality technologies when the objective of the training is to improve hazard recognition, appraisal, and/or appreciation of risk where high-fidelity creates engagement and emotional arousal in response to target stimuli.

Biofeedback technologies could be used to improve learner engagement and transfer. For instance, eye-tracking can help inform what people need to learn about in the area of hazard recognition training, and heart rate variability monitoring can be used to provide ongoing feedback about the application of stress and distraction management techniques. Further, designers can use high-fidelity imagery and other media to induce strong emotional responses to high-risk hazards; however, attendees may need to be psychologically prepared if the imagery is graphic or potentially upsetting.

4.2.3. After training

Combining safety training concepts with existing routines and processes embedded within management systems is likely to cue learned content and boost transfer. In other words, organizations could identify ways that safety training can be integrated within existing SMS and safety processes; for example, embedding training language or concepts into risk assessment forms or modifying incident investigation processes to include appreciative inquiry skills learned during safety training may be beneficial.

An area that requires additional research but nonetheless seems important for transfer is refresher training. One recommendation is to time refresher training to ensure learning is retained and embedded. To date, only a few studies have been conducted on the design and timing of refresher training—one study by Kluge and Burkolter (2012) found that physical practice of a process control task resulted in better learning retention and transfer than a "symbolic rehearsal" or written refresher task. These refreshers took place 2–3 weeks after the initial training. In our view, ensuring a safety training refresher booster approximately one month after the initial training event is probably optimal for long-term retention. However, further work is needed to identify more specific guidance around refresher timing.

Importantly, organizations should identify how the training transfer will be measured and use a range of metrics that go beyond names and numbers of attendees, incident reductions, and evaluative "smile sheets." Organizations should consider whether it is possible to measure safety training transfer using behavioral observations, diarized feedback, pulse surveys, and/or competency evaluations. Additional training metrics could include motivation and confidence to transfer/apply learning, actual training application, type of transfer achieved (near/far), and impact of training on safety performance (positive, negative, neutral). Finally, organizations should monitor the transfer of training through surveys and/or observations of training application—identifying and ameliorating barriers or challenges to transfer.

Numerous training transfer studies have shown that social support is critical (e.g., Burke & Hutchins, 2007, 2008). For organizations delivering safety training, target direct leaders and supervisors to ensure there is a strategy to provide social support for safety training transfer; at a minimum, supervisors should be holding post-training conversations with workers about how to apply learning, what they learned, and how ongoing safety development can occur.

4.3. Future research directions

An interesting direction of future research concerns the dynamic modeling of safety transfer over time. Too often, training transfer is operationalized as a binary phenomenon (i.e., either it happens, or it does not) and as a product or outcome of the training delivery factors combined with attendee and contextual factors (Bell & Kozlowski, 2010). Instead, transfer should be thought of as a process in itself (Foxon, 1997), one that unfolds and fluctuates over time with the ebb and flow of various predictors such as supervisor support (Olenick, Blume, & Ford, 2020). With respect to safety training transfer, modeling the organizational or even team safety climate as a dynamic variable that affects training application could be a powerful way to advance the field. Given that safety climate is a dynamic variable that is both a leading indicator ahead of incidents and a lagging indicator in response to incidents (Payne, Bergman, Beus, Rodríguez, & Henning, 2009), taking multiple measurements in parallel to transfer behaviors could help to explain why some safety training fails to be applied in practice.

5. Conclusion

Although safety training is a mainstay of many organizations' safety management systems, not all safety training is effective. Drawing from the general training transfer literature and considering the application of these findings to safety training in light of its specific features and challenges, there is clearly no one-size-fits-all solution. Safety training transfer requires a multi-pronged approach that considers the trainee, training, and contextual factors, their interactions, and how individual characteristics should be used to inform the organization's training transfer strategy. As more research is done on this topic, organizations will learn about how safety training can be optimized to produce the best financial returns and the most effective improvement in safety performance.

Acknowledgment

This research was partially funded by Urban Utilities, Queensland, Australia.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2021.06.004.

References

Asterisked papers were included in our 10-year review of safety training research

*Ahn, S., Kim, T., Park, Y.-J., & Kim, J.-M. (2020). Improving effectiveness of safety training at construction worksite using 3D BIM simulation. *Advances in Civil Engineering, 2020*, 1–12. https://doi.org/10.1155/2020/2473138.

Ajzen, I. (2005). *Attitudes, personality, and behavior*. London: McGraw-Hill Education.

Arcury, T. A., Nguyen, H. T., Summers, P., Talton, J. W., Holbrook, L. C., Walker, F. O., ... Quandt, S. A. (2014). Lifetime and current pesticide exposure among Latino farmworkers in comparison to other Latino immigrants. *American Journal of Industrial Medicine, 57*(7), 776–787.

Arghode, V., & Wang, J. (2016). Exploring trainers' engaging instructional practices: A collective case study. *European Journal of Training and Development, 40*(2), 111–127.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*(1), 63–105.

Baldwin, T. T., Kevin Ford, J., & Blume, B. D. (2017). The state of transfer of training research: Moving toward more consumer-centric inquiry. *Human Resource Development Quarterly, 28*(1), 17–28.

Bell, B. S., & Kozlowski, S. W. J. (2010). Toward a theory of learner-centered training design: An integrative framework of active learning. In S. W. J. Kozlowski & E. Salas (Eds.), *SIOP organizational frontiers series. Learning, training, and development in organizations* (pp. 263–300). Routledge/Taylor & Francis Group.

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology, 53*, 87–105.

Beus, J. M., Bergman, M. E., & Payne, S. C. (2010). The influence of organizational tenure on safety climate strength: A first look. *Accident Analysis & Prevention, 42*(5), 1431–1437.

*Bhandari, S., & Hallowell, M. R. (2017). Emotional engagement in safety training: Impact of naturalistic injury simulations of the emotional state of construction workers. *Journal of Construction and Engineering Management, 143*(12), 04017090. 10.1061(ASCE)CO.1943-7862.0001405.

*Bhandari, S., Hallowell, M. R., & Correll, J. (2019). Making construction safety training interesting: A field-based quasi-experiment to test the relationship between the emotional arousal and situational interest among adult learners. *Safety Science, 117*, 58–70. https://doi.org/10.1016/j.ssci.2019.03.028.

*Bhandari, S., Hallowell, M. R., Boven, L. V., Welker, K. M., Golparver-Fard, M., & Gruber, J. (2020). Using augmented virtuality to examine how emotions influence construction-hazard identification, risk assessment, and safety decisions. *Journal of Construction and Engineering Management, 146*(2), 04019102. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001755/.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management, 36*(4), 1065–1105.

Boelens, R., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review, 22*(1), 1–18.

*Brahm, F., & Singer, M. (2013). Is more engaging safety training always better in reducing accidents? Evidence of self-selection from Chilean panel data. *Journal of Safety Research, 47*, 85–92. https://doi.org/10.1016/j.jsr.2013.09.003.

*Browne, A. M., Deutsch, E. S., Corwin, K., Davis, D. H., Teets, J. M., & Apkon, M. (2019). An IDEA: Safety training to improve critical thinking by individuals and teams. *American Journal of Medical Quality, 34*(6), 569–576. https://doi.org/10.1177/1062860618820687.

Bunch, K. J. (2007). Training failure as a consequence of organizational culture. *Human Resource Development Review, 6*(2), 142–163.

Burke, M. J., Chan-Serafin, S., Salvador, R., Smith, A., & Sarpy, S. A. (2008). The role of national culture and organizational climate in safety training effectiveness. *European Journal of Work and Organizational Psychology, 17*(1), 133–152.

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review, 6*(3), 263–296.

Burke, L. A., & Hutchins, H. M. (2008). A study of best practices in training transfer and proposed model of transfer. *Human Resource Development Quarterly, 19*(2), 107–128.

Burke, M. J., Salvador, R. O., Smith-Crowe, K., Chan-Serafin, S., Smith, A., & Sonesh, S. (2011). The dread factor: How hazards and safety training influence learning and performance. *Journal of Applied Psychology, 96*(1), 46–70.

Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American Journal of Public Health, 96*(2), 315–324.

Butler, A., Reed, M., & Le Grice, P. (2007). Vocational training: Trust, talk and knowledge transfer in small businesses. *Journal of Small Business and Enterprise Development, 14*(2), 280–293.

*Casey, T. W., Krauss, A. D., & Turner, N. (2018). The one that got away: Lessons learned from the evaluation of a safety training intervention in the Australian prawn fishing industry. *Safety Science, 108*(1), 218–224. https://doi.org/10.1016/j.ssci.2017.08.002.

*Chittaro, L., Corbett, C. L., McLean, G. A., & Zangrando, N. (2018). Safety knowledge transfer through mobile virtual reality: A study of aviation life preserver donning. *Safety Science, 102*, 159–168. https://doi.org/10.1016/j.ssci.2017.10.012.

*Choi, M., Ahn, S., & Seo, J. (2020). VR-based investigation of forklift operator situation awareness for preventing collision accidents. *Accident Analysis and Prevention, 136*, 105404. https://doi.org/10.1016/j.aap.2019.105404.

Clarke, S. (2003). The contemporary workforce: Implications for organisational culture. *Personnel Review, 32*(1), 40–57.

Dekker, S. (2019). *Foundations of safety science*. Miami: CRC Press.

*Eggerth, D. E., Keller, B. M., Cunningham, T. R., & Flynn, M. A. (2018). Evaluation of Toolbox safety training in construction: The impact of narratives. *American Journal of Industrial Medicine, 61*(12), 997–1004. https://doi.org/10.1002/ajim.22919.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*(7), 1167–1187.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fishbein, M., & Raven, B. H. (1962). The AB scales: An operational definition of belief and attitude. *Human Relations, 15*(1), 35–44.

Ford, J. K., Baldwin, T. T., & Prasad, J. (2018). Transfer of training: The known and the unknown. *Annual Review of Organizational Psychology and Organizational Behavior, 5*(1), 201–225.

Ford, J. K., & Weissbein, D. A. (1997). Transfer of training: An updated review and analysis. *Performance Improvement Quarterly, 10*(2), 22–41.

Foxon, M. (1997). The influence of motivation to transfer, action planning, and manager support on the transfer process. *Performance Improvement Quarterly, 10*, 42–63.

*Freitas, A. C., Silva, S. A., & Santos, C. M. (2017). Predictors of safety training transfer support as in-role behavior of occupational health and safety professionals. *European Journal of Training and Development, 41*(9), 776–799. https://doi.org/10.1108/EJTD.03.2017.0024.

Freitas, A. C., Silva, S. A., & Santos, C. M. (2019). Safety training transfer: The roles of coworkers, supervisors, safety professionals, and felt responsibility. *Journal of Occupational Health Psychology, 24*(1), 92–107. https://doi.org/10.1037/ocp0000125.

*Gao, Y., Gonzalez, V. A., & Yiu, T. W. (2019). The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Computers & Education, 138*, 101–115. https://doi.org/10.1016/j.compedu.2019.05.003.

Gegenfurtner, A., Könings, K. D., Kosmajac, N., & Gebhardt, M. (2016). Voluntary or mandatory training participation as a moderator in the relationship between goal orientations and transfer of training. *International Journal of Training and Development, 20*(4), 290–301.

Griffin, M. A., & Curcuruto, M. (2016). Safety climate in organizations. *Annual Review of Organizational Psychology and Organizational Behavior, 3*(1), 191–212.

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347–358.

*Gummesson, K. (2016). Effective measures to decrease air contaminants through risk and control visualization—A study of the effective use of QR codes to facilitate safety training. *Safety Science, 82*, 120–128. https://doi.org/10.1016/j.ssci.2015.09.011.

Harvey, P. D., Balzer, A. M., & Kotwicki, R. J. (2020). Training engagement, baseline cognitive functioning, and cognitive gains with computerized cognitive training: A cross-diagnostic study. *Schizophrenia Research: Cognition, 19* 100150.

Hasanzadeh, S., de la Garza, J. M., & Geller, E. S. (2020). How sensation-seeking propensity determines individuals' risk-taking behaviors: Implication of risk compensation in a simulated roofing task. *Journal of Management in Engineering, 36*(5). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000813 04020047.

Hofstede, G. (1980). Motivation, leadership, and organization: Do American theories apply abroad? *Organizational Dynamics, 9*(1), 42–63.

Hogan, J., & Foster, J. (2013). Multifaceted personality predictors of workplace safety performance: More than conscientiousness. *Human Performance, 26*(1), 20–43.

Hollnagel, E. (2011). Prologue: The scope of resilience engineering. In E. Hollnagel, J. Paries, D. D. Woods, & J. Wreathall (Eds.), *Resilience engineering in practice: A guidebook*. Surrey: Ashgate Publishing Limited.
Hollnagel, E., Wears, R. L., & Braithwaite, J. (2015). From safety-I to safety-II: A white paper. The resilient health care net: published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia.
*Holte, K. A., & Follo, G. (2018). Making occupational health and safety training relevant for farmers: Evaluation of an introductory course in occupational health and safety in Norway. *Safety Science, 109*, 368–376. https://doi.org/10.1016/j.ssci.2018.05.020.
Huang, J. L., & Bramble, R. J. (2016). Trait, state, and task-contingent conscientiousness: Influence on learning and transfer. *Personality and Individual Differences, 92*, 180–185.
Hutchins, H. M., & Burke, L. A. (2006). Has relapse prevention received a fair shake? A review and implications for future transfer research. *Human Resource Development Review, 5*(1), 8–24.
Hutchins, H. M., Burke, L. A., & Berthelsen, A. M. (2010). A missing link in the transfer problem? Examining how trainers learn about training transfer. *Human Resource Management, 49*(4), 599–618.
*Jeelani, I., Han, K., & Albert, A. (2018). Automating and scaling personalized safety training using eye-tracking data. *Automation in Construction, 93*, 63–77. https://doi.org/10.1016/j.autcon.2018.05.006.
Jokisch, M. R., Schmidt, L. I., Doh, M., Marquard, M., & Wahl, H.-W. (2020). The role of internet self-efficacy, innovativeness and technology avoidance in breadth of internet use: Comparing older technology experts and non-experts. *Computers in Human Behavior, 111*, 106408. https://doi.org/10.1016/j.chb.2020.106408.
Kanfer, R. (1990). Motivation and individual differences in learning: An integration of developmental, differential and cognitive perspectives. *Learning and Individual Differences, 2*(2), 221–239.
Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(4), 657–690.
Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology, 93*(1), 59–69.
Kim, A., Schweighofer, N., & Finley, J. (2019). Locomotor skill acquisition in virtual reality shows sustained transfer to the real world. *Journal of NeuroEngineering and Rehabilitation, 16*(113). https://doi.org/10.1186/s12984-019-0584-y.
Kluge, A., & Burkolter, D. (2012). "Being prepared for the infrequent": A comparative study of two refresher training approaches and their effects on temporal and adaptive transfer in a process control task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 2437–2441.
Knowles, M. (1996). Andragogy: An emerging technology for adult learning. London, UK.
*Kovacic, M., & Cunningham, K. (2019). Effective electrical safety program training in multi-lingual/cultural environments. *IEEE Transactions on Industry Applications, 55*(4), 4384–4388.
Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*(2), 311–328.
Krauss, A. D. (2005). *The transfer problem: Examining the direct and interactive effects of safety training transfer climate and work locus of control on the transfer of safety training* (unpublished thesis). Colorado State University.
Krauss, A., Casey, T. W., & Chen, P. Y. (2014). Making safety training stick. *Contemporary Occupational Health Psychology, 3*, 181–197.
*Liang, Z., Zhou, K., & Gao, K. (2019). Development of virtual reality serious game for underground rock-related hazards safety training. *IEEE Access, 7*, 118639–118649. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2019.2934990.
Lingappa, A. K., Kiran K K., Oommen Mathew, A., & Wright, L. T. (2020). Safety training transfer in chemical manufacturing: The role of personality traits and learning motivation. *Cogent Business & Management, 7*(1), 1835335. https://doi.org/10.1080/23311975.2020.1835335.
Lingard, H., & Rowlinson, S. (2005). *Occupational health in construction project management*. USA: Spoon Press.
*Loosemore, M., & Malouf, N. (2019). Safety training and positive safety attitude formation in the Australian construction industry. *Safety Science, 113*, 233–243. https://doi.org/10.1016/j.ssci.2018.11.029.
Malakis, S., Kontogiannis, T., & Kirwan, B. (2010). Managing emergencies and abnormal situations in air traffic control (part I): Taskwork strategies. *Applied Ergonomics, 41*(4), 620–627.
Mao, Y., Chen, P., Li, L., & Huang, D. (2014). Virtual reality training improves balance function. *Neutral Regeneration Research, 9*(17), 1628–1634.
Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734.
Murphy, P. K., & Mason, L. (2006). Changing knowledge and beliefs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 305–324). Lawrence Erlbaum Associates Publishers.
Namian, M., Albert, A., Zuluaga, C. M., & Behm, M. (2016). Role of safety training: Impact on hazard recognition and safety risk perception. *Journal of Construction and Engineering Management, 142*(12). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001198 04016073.
*Namian, M., Albert, A., Zuluaga, C. M., & Jaselskis, E. J. (2016). Improving hazard-recognition performance and safety training outcomes: Integrating strategies for training transfer. *Journal of Construction and Engineering Management, 142*(10), 04016048. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001160.
Naquin, S. S., & Holton, E. F. III, (2002). The effects of personality, affectivity, and work commitment on motivation to improve work through learning. *Human Resource Development Quarterly, 13*(4), 357–376.
Olenick, J., Blume, B. D., & Ford, J. K. (2020). Advancing training and transfer research through the application of nonlinear dynamics. *European Journal of Work and Organizational Psychology, 29*(4), 541–555.
Payne, S. C., Bergman, M. E., Beus, J. M., Rodríguez, J. M., & Henning, J. B. (2009). Safety climate: Leading or lagging indicator of safety outcomes? *Journal of Loss Prevention in the Process Industries, 22*(6), 735–739.
Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education, 144*, 103701. https://doi.org/10.1016/j.compedu.2019.103701.
Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*(2-3), 183–213.
*Robson, L. S., Stephenson, C. M., Schulte, P. A., Amick, B. C. III., Irvin, E. L., Eggerth, D. E., ... Grubb, P. L. (2012). A systematic review of the effectiveness of occupational health and safety training. *Scandinavian Journal of Work, Environment & Health, 38*(3), 193–208.
Rouillier, J. Z., & Goldstein, I. L. (1993). Determinants of the climate for transfer of training. *Human Resource Development Quarterly, 4*, 377–390.
*Ruttenberg, R., & Rice, C. (2019). Assessing the impact of health and safety training: Increased behavioral change and organizational performance. *American Journal of Industrial Medicine, 62*(11), 986–995. https://doi.org/10.1002/ajim.v62.1110.1002/ajim.23026.
Safe Work Australia. (2018). Model code of practice: Work health and safety consultation, cooperation, and coordination. https://www.safeworkaustralia.gov.au/doc/model-code-practice-work-health-and-safety-consultation-cooperation-and-coordination.
Schein, E. H. (2010). *Organizational culture and leadership*. John Wiley & Sons.
Sitzmann, T., & Weinhardt, J. M. (2018). Training engagement theory: A multilevel perspective on the effectiveness of work-related training. *Journal of Management, 44*(2), 732–756.
Smith, G. (2018). *Paper safe – The triumph of bureaucracy in safety management*. Perth, WA: Wayland Legal Pty Ltd.
Stuart, A. (2014). A blended learning approach to safety training: Student experiences of safe work practices and safety culture. *Safety Science, 62*, 409–417.
Tracey, J. B., Tannenbaum, S. I., & Kavanagh, M. J. (1995). Applying trained skills on the job: The importance of the work environment. *Journal of Applied Psychology, 80*(2), 239–252.
Trout, G. (2016). E-learning and online training. *Professional Safety, 61*(6), 34–36.
Tsai, T.-H., Chang, H.-T., Chen, Y.-J., Chang, Y.-S., & Xia, F. (2017). Determinants of user acceptance of a specific social platform for older adults: An empirical examination of user interface characteristics and behavioral intention. *PloS One, 12*(8), e0180102.
*Vigoroso, L., Caffaro, F., & Cavallo, E. (2020). Occupational safety and visual communication: User-centred design of safety training material for migrant farmworkers in Italy. *Safety Science, 121*, 562–572. https://doi.org/10.1016/j.ssci.2018.10.029.
Wallen, E. S., & Mulloy, K. B. (2006). Computer-based training for safety: Comparing methods with older and younger workers. *Journal of Safety Research, 37*(5), 461–467.
Wang, B., Li, H., Rezgui, Y., Bradley, A., & Ong, H. N. (2014). BIM based virtual environment for fire emergency evacuation. *The Scientific World Journal, 2014*, 1–22.
Yorio, P. L., Edwards, J., & Hoeneveld, D. (2019). Safety culture across cultures. *Safety Science, 120*, 402–410.
Zohar, D. (2010). Thirty years of safety climate research: Reflections and future directions. *Accident Analysis & Prevention, 42*(5), 1517–1522.

Tristan Casey is an expert in safety leadership and organisational culture, and a member of Griffith University's Safety Science Innovation Lab. With a career in workplace health and safety spanning 12 years, he has consulted nationally and internationally across a diverse range of industries. His passion is translating abstract/theoretical concepts into practical tools that have measurable impact.

Nick Turner, PhD (Sheffield) is Professor and Distinguished Research Chair in Advanced Leadership at the Haskayne School of Business, University of Calgary, Canada. His research focuses on work design, transformational leadership, and occupational health and safety.

Xiaowen Hu is an interdisciplinary, applied safety scientist with a passion for integrating diverse theoretical perspective to bring innovative solution to complex real-world problems. She is currently a Senior Lecturer at Queensland University of Technology and teaches a range of courses related to human resource management.

Kym Bancroft is an experienced HSE Executive with a Masters in Safety Leadership and Organisational Psychology. With close to 20 years' experience in safety, health, wellbeing, and injury management Kym has worked nationally and internationally assisting organisations to achieve safety transformation that is sustainable long term.

Journal of Safety Research 78 (2021) 28–35

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Modeling drivers' reaction when being tailgated: A Random Forests Method

Yueru Xu [c,d], Shan Bao [a,b,*], Anuj K. Pradhan [d]

[a] Industrial and Manufacturing Systems Engineering Department, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI 48128, USA
[b] University of Michigan Transportation Research Institute, 2901 Baxter Rd, Ann Arbor 48109, MI, USA
[c] Intelligent Transportation System Research Center, Southeast University, Nanjing 211189, China
[d] Mechanical and Industrial Engineering Department, University of Massachusetts Amherst, Amherst, MA 01003, USA

ARTICLE INFO

Article history:
Received 4 May 2020
Received in revised form 28 December 2020
Accepted 6 May 2021
Available online 25 May 2021

Keywords:
Road safety
Tailgating
Random Forests
Driving behavior
Naturalistic driving data

ABSTRACT

Background: Tailgating is a common aggressive driving behavior that has been identified as one of the leading causes of rear-end crashes. Previous studies have explored the behavior of tailgating drivers and have reported effective solutions to decrease the amount or prevalence of tailgating. This paper tries to fill the research gap by focusing on understanding highway tailgating scenarios and examining the leading vehicles' reaction using existing naturalistic driving data. *Method:* A total of 1,255 tailgating events were identified by using the one-second time headway threshold criterion. Four types of reactions from the leading vehicles were identified, including changing lanes, slowing down, speeding up, and making no response. A Random Forests algorithm was employed in this study to predict the leading vehicle's reaction based on corresponding factors including driver, vehicle, and environmental variables. *Results:* The analysis of the tailgating scenarios and associated factors showed that male drivers were more frequently involved in tailgating events than female drivers and that tailgating was more prevalent under sunny weather and in daytime conditions. Changing lanes was the most prevalent reaction from the leading vehicle during tailgating, which accounted for more than half of the total events. The results of Random Forests showed that mean time headway, duration of tailgating, and minimum time headway were three main factors, which had the greatest impact on the leading vehicle drivers' reaction. It was found that in 95% of the events, leading vehicles would change lanes when being tailgated for two minutes or longer. *Practical Applications:* Results of this study can help to better understand the behavior and decision making of drivers. This understanding can be used in designing countermeasures or assistance systems to reduce tailgating behavior and related negative safety consequences.

1. Introduction

Tailgating is defined as driving dangerously close behind another vehicle leaving insufficient distance to respond to potential emergency situations (Wang & Song, 2011). It is a common aggressive driving behavior and one of the most frequent forms of road rage. Road rage can refer to any display of aggression by a driver and has proven to have a consistently positive association with increased risk of serious motor vehicle accidents (Galovski et al., 2006). One study reported that about 62% of the investigated drivers claimed they often tailgate in an aggressive way as a result of road rage (Joint, 1995). In addition, tailgating has been identified as one of the leading causes of rear-end crashes (Lee et al., 2002). According to the National Highway Traffic Safety Administration (2015), rear-end collisions were the most common crash type in 2015, which accounted for about one third of the total crashes resulting in 2,203 fatalities and about 556,000 injuries. Two human-related factors were reported to be primarily associated with rear-end crashes, inattention and tailgating (Dingus et al., 1997), with tailgating indicated as a major cause of severe consequences (Carter et al., 1995). Studies also showed that tailgating during highway driving can be more dangerous than during non-highway driving due to the high driving speed (Carter et al., 1995; Evans & Wasielewski, 1982).

Time headway is the term used to describe the time needed for the following vehicle to cover the between-vehicle gap and reach the leading vehicle. Different time headway thresholds have been used in previous studies to define tailgating behavior (Evans &

* Corresponding author at: Industrial and Manufacturing Systems Engineering Department, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI 48128, USA.

E-mail address: shanbao@umich.edu (S. Bao).

https://doi.org/10.1016/j.jsr.2021.05.004
0022-4375/

Wasielewski, 1982; Michael et al., 2000; Monteiro et al., 2015). For example, Evans and Wasielewski (1982) used the one-second time headway threshold to define tailgating behavior and found that drivers with frequent tailgating behavior were more likely to have a higher number of traffic violations and crashes than those who tended to follow with longer time headways. Some studies also recommended longer time headway thresholds in defining tailgating, including the "Two-Second Rule" proposed by Michael and his colleagues (2000). In their study, it was suggested to drivers to keep a time headway of at least two seconds between their vehicle and the leading vehicle. In their study, following with time headways of less than two seconds was considered as tailgating. Another study recommended that drivers should keep the distance of one car length for each 10 miles per hour of speed, similar to a three-second rule (Monteiro et al., 2015). A general conclusion from those studies is that following a leading vehicle with time headway of one second or less shall be considered as tailgating.

Contributing factors associated with tailgating behavior have been previously investigated, including driver individual differences, risk perception and situational factors. Rajalin et al. (1997) researched driving behavior on two-lane highways for 157 close-following drivers and 178 conservative drivers and found that young males were more likely to be aggressive drivers and tailgated more often than other drivers. Through a questionnaire survey study, Wang and Song (2011) concluded that the top three self-reported causes for tailgating were "heavy traffic," "slow car ahead of my vehicle," and "I am in a hurry." Tlhabano et al. (2013) conducted a study in Botswana and stated tailgating behavior was attributable to a lack of driver awareness or an inability to determine safe following distances. Duan et al. (2013) proposed a model of risk perception when following another vehicle and observed drivers' tendencies to keep longer time headways behind trucks and SUVs than behind sedans.

The literature includes several studies of tailgating mitigation strategies. Michael et al. (2000) explored two tailgate intervention methods: presenting drivers with one hand-held sign stating "Please Don't Tailgate," or with the other sign stating "Help Prevent Crashes – Please Don't Tailgate." The latter message sign had a significant effect on reducing tailgating by increasing the average time headway from 2.11 seconds to 2.29 seconds, an average increase of 9%. Rama and Kulmala (2000) investigated whether providing time headway information would reduce tailgating and found that providing recommended minimum time headway through roadside traffic signs can effectively reduce tailgating behavior (defined as following with time headway of 1.5 seconds or shorter). Lertworawanich (2006, 2009)) conducted multiple studies to estimate safe car-following distances for different speed limits and developed the "dot" treatment pavement markings that were distance markings drawn on the road to remind drivers of the appropriate distance from leading vehicles. It was found that the average time headway increased after the implementation of marking at the study site. State-led tailgating mitigation programs have been conducted in Pennsylvania, Minnesota, and Maryland (Roadway Safety Foundation, 2001; Minnesota DOT, 2006; Song & Wang, 2010). Among all of these programs, the Pennsylvania Department of Transportation (PennDOT) program was considered to be very successful (Roadway Safety Foundation, 2001); their reflective dots and markings on the roadway led to a 60% drop of observed tailgating behavior. Their results suggest that pavement markings and signs can help drivers gauge their following distance. An in-depth investigation by Hutchinson (2008) also concluded that countermeasures such as advisory signs, pavement markings, and enforcement by the police could decrease tailgating and reduce rear-end crashes.

Many previous studies of tailgating behavior and corresponding countermeasures focused on the behavior of the following vehicle and infrastructure-based solutions. There is still a need to examine the reactions and decision-making processes of the leading vehicle drivers and explore additional countermeasures, as different reactions will cause different safety consequences (e.g., road rage). This study aims to fill the research gap by examining the leading vehicle drivers' reactions when being tailgated using naturalistic driving data. The objectives of this work are twofold: (1) to evaluate typical tailgating scenarios and associated factors, and (2) to model how vehicles respond to the situations when being tailgated. The types of reactions and associated characteristics from leading vehicles were identified and examined. A Random Forests based model was then applied to predict responses from the leading vehicle.

2. Method

2.1. Data extraction

This study used the light-vehicle fleet data from the Integrated Vehicle Based Safety System (IVBSS) program (Sayer et al., 2010). The IVBSS program deployed and tested an integrated in-vehicle crash warning system for both heavy trucks and light vehicles (LeBlanc et al., 2013). Each instrumented vehicle captured information regarding the driving environment, driver activity, system behavior, and vehicle kinematics. The data collection frequency was from 10 to 50 Hz (Wang et al., 2017). Driving data from a gender-balanced sample of 108 participants from three age groups (younger, between 20 and 30 years old; middle-aged, between 40 and 50 years old; and older, between 60 and 70 years old) were collected. The average age of the three groups were 25.1, 46.1, and 64.4 years old, respectively. All participants were recruited through several channels, such as using the University of Michigan's existing study volunteer pool and posting advertisements on Craigslist and local newspaper. All drivers participated in the study for a six-week period and received compensation for their participation in the study. Each participant used the instrumented research vehicle for their personal trip purposes during this six-week period and they were instructed to drive as naturally as they would normally do with their personal vehicle. More study details were described in the previous studies (Bao et al., 2020; Sayer et al., 2010; Wang et al., 2017).

To examine leading drivers' reactions while being tailgated, tailgating episodes were identified and corresponding data were then extracted from the IVBSS database. In this study, a tailgating episode was defined using the one-second time headway threshold, as illustrated in Fig. 1 and defined previously by Evans and Wasielewski (1982). Time headway (in seconds) was calculated by dividing the distance between the leading and following vehicles by the current speed of the following vehicle. The detailed criteria for tailgating event data extraction were:

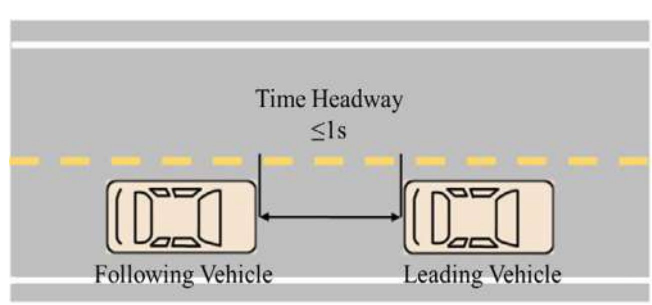

Fig. 1. Definition of tailgating between two vehicles.

- The time headway between the two vehicles was less than or equal to one second for at least a one-second duration (i.e., for ten consecutive data points).
- The tailgating event occurred during free-flowing traffic on highways (i.e., not during traffic jams) by using a speed threshold of 20 meters/second (i.e., 45 miles/hour) or higher.
- Driving data collected on ramps were excluded.
- Events where the following vehicle was attempting to change lanes were excluded.
- The time interval between two consecutive tailgating episodes was one second or longer.

Corresponding time-series and other related data of each episode were extracted directly from in-vehicle sensor data for traffic density, following distance and speed. Additionally, video coding provided data on number of traffic lanes and weather conditions. The calculation of vehicle kinematic features including following distances and event durations were all achieved based on in-vehicle sensor (e.g., radar) or vehicle controller area network (CAN) bus data.

2.2. Variables

More than half of the observed tailgating events ended with lane change maneuvers by the leading vehicle. To understand why drivers choose to make a lane change or not when being tailgated, the output variable of the analysis was set to be binary, *whether the leading vehicle made a lane change or not* (Yes/No). Nine input variables were used in this study including: minimum time headway (MinTH), mean time headway (MeanTH), duration time (DT), minimum speed (MinS), number of lanes (NL), lane type (LT), road line type (RLT), traffic density, day, and weather.

MinTH and MeanTH were the minimum and mean value of time headway between the two vehicles during each tailgating event. Duration time is the total recorded time of each tailgating event (i.e., a continuous following duration with time headway of one second or less). The type of vehicle traveling lane was video coded and divided into two types, fast lane or slow lane. For four-lane roads, the left two lanes were coded as fast lanes while the right two lanes were classified as slow lanes. For three-lane roads, the left and middle lanes were fast lanes while the right one was slow lane. The lane on the left was the fast lane and the one on the right was the slow lane for two-lane roads. Roadway type was classified as curvy or straight. Traffic density was divided into three groups: scarce traffic with one surrounding vehicle in the front of the leading vehicle, moderate traffic with two or three surrounding vehicles in the front, and dense traffic with four surrounding vehicles or more in the front. Weather condition was also video coded as adverse or normal condition.

2.3. Driver reaction prediction: Random Forests algorithm

Machine learning is a popular technique for pattern recognition and behavior modeling due to its effectiveness and efficiency with big data sets. Past applications include analysis of transportation safety and human driving behavior, such as, crash data analysis (Mohamed Radzi et al., 2017), driver distraction (Yao et al., 2018), and drowsy driving (Ngxande et al., 2017). Compared to the traditional generalized linear models, machine learning methods do not require presuppositions or predefined underlying relationships between dependent and independent variables. Such presuppositions and predefined relationships, if inappropriate within the context of the model, may affect the accuracy and reliability of the results. This study used the Random Forests (RF) algorithm to model leading vehicles' reactions when being tailgated. This machine learning method is an algorithm developed by Breiman (2001) based on a combination of a set of decision trees. It consists of non-parametric statistical approaches for conducting regression and classification analyses by recursive partitioning. RF is highly efficient in selecting large numbers of variables and handling overfitting problems in cases where a single classification tree yields inadequate results. Compared to other machine learning methods such as Artificial Neural Network (ANN) and Support Vector Machines (SVM), RF has the advantage of evaluating variable importance based on an out-of-bag (OOB) test of the decision trees in a forest (Breiman, 2001; Ma & Cheng, 2016).

Classification trees are one of the typical grouping algorithms for RF. A classification tree is built through a binary recursive partitioning approach, which requires an iterative process of splitting the data into partitions before further dividing on each of the branches (Harb et al., 2009). As shown in Fig. 2, the initial dataset in this study is created using all the input variables. Then, the algorithm systematically assigns each record to one of two subsets by using certain logic. For example, the algorithm can identify if the event occurs during daytime or nighttime and separate the entire dataset into two subsets. The objective of this step is to attain a

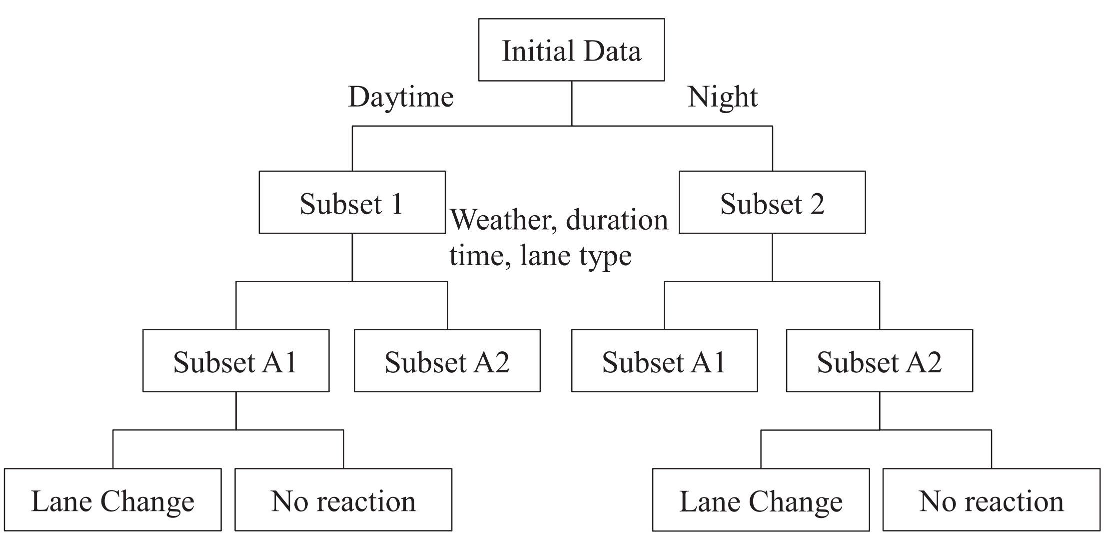

Fig. 2. Flow chart of classification trees in RF algorithm (adapted from Loh, 2011).

homogeneous set of labels in each partition. This process is then applied to each of the factors (weather, lane type, duration time, and so on) and continues until no more useful splits can be found. To choose the best splitting factor at a node, the algorithm tries every possible split to find the split associated with the largest decrease in diversity of the reaction type within each partition. The process continues at subsequent nodes until a full tree is generated. In the RF algorithm, a large number of classification trees are generated and each tree is built by selecting a random set of observations from the training dataset. The selected observations are called bootstrap samples and the left-out observations are called OOB samples. On average, each tree is grown using about 1 = e-1≈2/3 of the training database as bootstrap samples, leaving e-1≈1/3 as OOB samples (Lunetta et al., 2004).

At each node of the tree, rather than choosing the best split among all variables, the algorithm randomly selects several variables and chooses the best split among these variables. Different variables are used at each split in different trees. The result of the RF algorithm is obtained by averaging all the results of the individual classification trees. There is limited generalization error since a large number of trees are produced, to minimize overfitting concerns. In addition, The OOB samples are used to evaluate the accuracy and variable importance of RF, as well as to eliminate the need for a test set or cross-validation. This study used equation (1), where Xi are the inputs (e.g., weather, day, minTH, duration time) of *i*th OOB data and Yi is the actual reaction type of *i*th OOB data. The ensemble prediction $Y^{OOB}(X_i)$ is the reaction type calculated by RF algorithm (Bureau et al., 2005). The error rate (ER) of the prediction is calculated by:

$$ER = n^{-1}\sum_{i=1}^{n} I(Y^{OOB}(X_i) \neq Y_i) \quad (1)$$

where $I(*)$ is the indicator function; n means the number of trees in the RF.

To evaluate the importance of a specific variable, the values of each input variable are randomly permuted for the OOB samples. Next, modified OOB samples are applied to the tree to determine new reaction types. The difference between the misclassification rate for the modified and original OOB data is a measure of the importance of the variable. In this analysis, R software (R Core Team, 2014) was used to develop the RF algorithm and calculate the OBB error.

3. Results

A total of 1,255 valid tailgating episodes were identified from the IVBSS database by using the defined tailgating criteria. The tailgating data of the following vehicles were collected from 50 drivers (31 male, 19 female) from all three age groups (18 younger, 22 middle-aged, and 10 older drivers). The characteristics of all the tailgating events are shown in Fig. 3. Of the 1,255 tailgating events, 84.7% were from male drivers. Middle-aged drivers tailgated more frequently than the other cohorts (59.6% vs. 31.9%, 8.5% for younger and older drivers, respectively). In the IVBSS database, approximately 78% of all driving took place during daytime, in terms of both traveling time and traveling distance, while more than 97% of the tailgating events occurred during daytime (Fig. 3). Drivers were also found to have more tailgating events during normal weather driving, when compared to under adverse weather driving conditions (92.1% and 7.9%, respectively). In the database, normal weather represented about 93% of all the driving (Sayer et al., 2010). In addition, tailgating events occurred more often when drivers were using the fast lane (83.8%) than the slow lane (16.2%). Tailgating was more prevalent on two-lane roads (57.8%) than on roads with multiple available lanes (42.2%). More than half of the tailgating events (62.8%) occurred when there was moderate traffic density on the roads.

In this study, four different reaction types were identified from the leading vehicles during tailgating: changing lanes, slowing down, speeding up, or no response (Fig. 4). Drivers chose to change lanes while being tailgated during more than half of the cases (55.2%). Lead vehicle drivers increased their speed during about 13.8% of the cases. Leading drivers made no response 30.7% of the time. Only five events (0.004%) involved leading vehicles reducing their speed.

Since, the main interest of this study is to investigate leading vehicles' lane-changing response, the reaction types were further grouped into changing lanes or not and used in the RF analysis. The RF analysis used all the input variables described in the methods section.

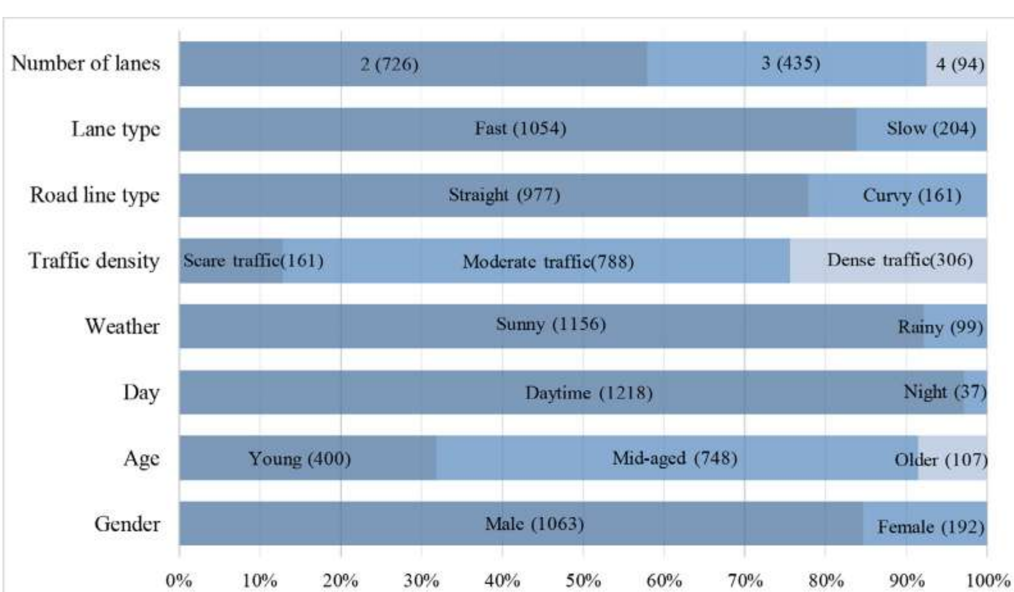

Fig. 3. Characteristics of tailgating events.

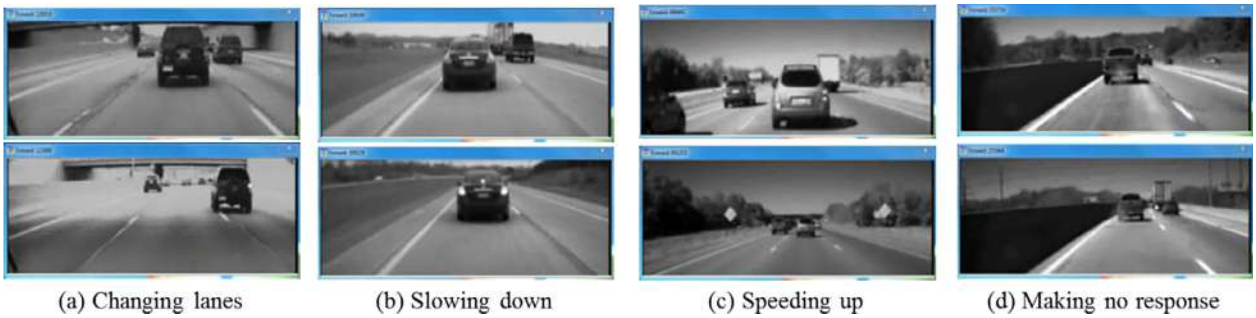

Fig. 4. Four reaction types of leading vehicles when being tailgated.

Table 1
Confusion Matrix of Training and Testing Data.

Training Data				Testing Data			
Predict Actual	No Lane Change	Lane Change	OOB Error	Predict Actual	No Lane Change	Lane Change	OOB Error
No Lane Change	284	119	29.5%	No Lane Change	111	45	28.8%
Lane Change	109	375	22.5%	Lane Change	48	164	22.6%

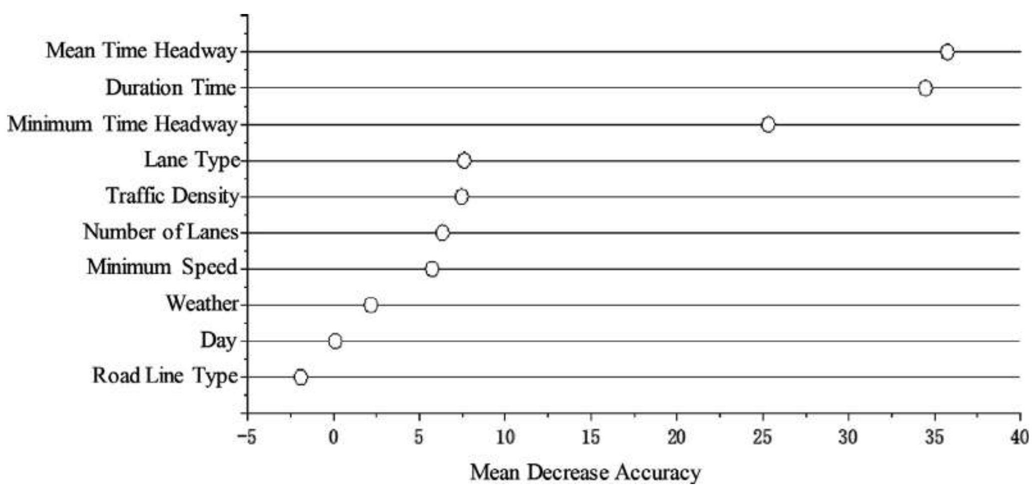

Fig. 5. Mean decrease accuracy of random forests for each input variable.

3.1. Random forest analysis

Two parameters in the RF analysis: the number of variables in each tree, *mtry*, and the number of trees, *ntree*, were tuned to obtain optimal performance. The parameter *ntree* usually provides stable OOB errors and the parameter *mtry* is used to pick the value with best performance between 2^{-1}*N and 2^{1}*N, where N represents one third of the total number of variables in regression problems. In this analysis, 10 variables were identified as inputs and four sets of *mtry* (2, 3, 4, and 5) were applied separately to build the RF.

After applying these parameters, the forest reached the lowest and most stable OOB error when *ntree* = 300 and *mtry* = 4. Therefore, these two values were used to establish the RF trees. The dataset was split randomly into two parts, with 70% of the dataset designated as training data while the remaining 30% was used as testing data. The confusion matrix of training and testing data was calculated and listed in Table 1.

The OOB errors for the two response types (i.e., Lane Change or No Lane Change) were 29.5% and 22.5% in the training data analysis, while in the testing data, the accuracy of the prediction was calculated as the percentage of correct prediction (total number of correct predictions divided by the total number of events). The accuracies of the prediction were similar, 74.3% and 74.7% for training data and testing data separately, suggesting that the model is valid with no over-fitting. In general, the model accuracy is reasonably high and thus can be used to describe the influence of various variables on the reaction of leading vehicle.

The importance of each variable in the final model was determined using the index of mean decrease accuracy (MDA) as an indication of its contribution in predicting the responses. MDA is the decrease of the model accuracy when an input variable is removed from the prediction model. For example, the RF model prediction accuracy, as shown in in Fig. 5, will be decreased by 37% when the variable of MeanTH is taken out. MeanTH, DT and MinTH are the three most important variables in the final model as these three variables resulted in the largest decreases in model accuracy values if they were omitted. The following sections will address the characteristics of these three key variables and how they relate to the leading vehicle's reaction when being tailgated.

3.2. Mean time headway (MeanTH)

The MeanTH distribution for tailgating events differs with the lane change response (yes or no). Situations where the leading vehicle changed lanes were associated with shorter Mean TH, while situations where the driver did not change lanes were associated with longer Mean TH, as shown in Fig. 6. The most frequent MeanTH for the no lane change responses was between 0.7 s and 1.0 s. When the lane change response occurred, the MeanTH was between 0.5 s and 0.9 s. In general, drivers are more likely to change lanes when the MeanTH is comparatively smaller. When the following vehicles were tailgating with a MeanTH of less than 0.8 s, the leading vehicles chose to change lanes and yield the right of way to the tailgating vehicle. Lead vehicles were more likely to stay in the same lane when there was an average following gap of 0.8 s or larger. The results indicate that leading vehicle drivers may feel more intimidated with a close following vehicle and are more likely to change lanes to remove themselves from the situation.

3.3. Minimum time headway (MinTH)

The MinTH values from all the tailgating records were divided equally into 10 categories from 0 s to 1 s. The numbers of tailgating records of different responses were counted and shown in the Heatmap of Fig. 7. Similar results were observed as the results of MeanTH analysis. The majority of tailgating events with no lane change responses had larger values of MinTH (typically greater than 0.6 s) than tailgating events with lane change responses. The MinTH of most tailgating events with lane change responses was between 0.3 and 0.7 s.

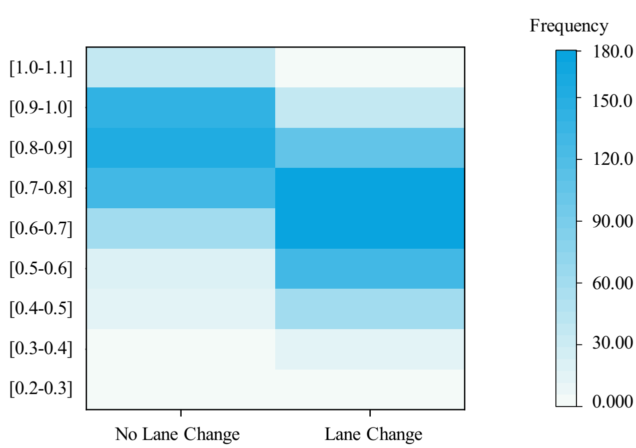

Fig. 6. Heatmap of MeanTH distribution of the two response types.

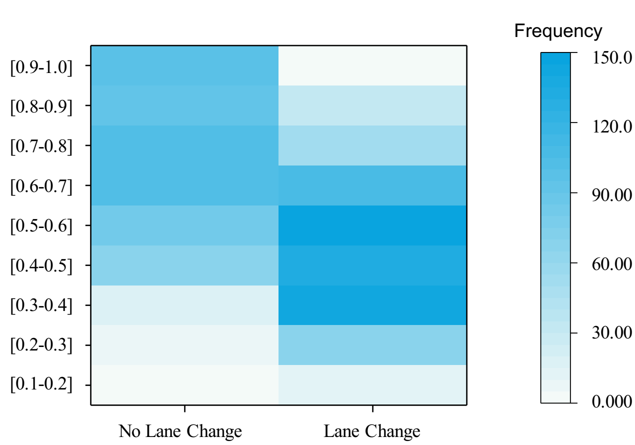

Fig. 7. Heatmap of MinTH.

Fig. 8. Percentage of duration time for different leading vehicle reactions.

3.4. Duration time (DT)

A histogram of all tailgating episodes' duration for each of the two reaction types is shown in Fig. 8. The results showed that about 80% of the tailgating events were less than 15 s long, with about 8.2% of the events having much longer durations (i.e., 2 minutes or longer). Long tailgating events can result in higher crash risks than short ones. Further, tailgating events with durations of 2 minutes or longer were more likely to result in lane change responses.

4. Conclusions and discussions

The investigation of tailgating behavior is an important human factors research topic in driving safety. Previous studies have identified the contributing factors of tailgating behavior and potential methods to reduce tailgating frequency from the perspective of a following vehicle, while the behavior and responses from the leading vehicle have been ignored. This study is designed to examine the tailgating problem from the leading vehicle drivers' perspective by investigating conditions associated with different reactions and developing algorithms to model drivers' responses when being tailgated.

Naturalistic driving data from the IVBSS program were used in this study. A total of 1,255 valid tailgating records were identified for analysis. Of all the valid records, 84.7% of the tailgating events were from male drivers, suggesting that male drivers are more likely to tailgate than female drivers. The data also show that middle-aged drivers were overrepresented in the tailgating events, accounting for 59.6% of all the tailgating records.

Four leading vehicle reaction types were identified in this study: changing lanes, slowing down, speeding up, and making no response. The results showed that most leading vehicle drivers responded to the tailgating events in a relative "polite" way by changing lanes during more than half of the events (55.22%) or speeding up during about 13.8% of the cases. The RF algorithm was then applied to the data to model leading vehicles' responses (change lanes or maintain lane) when being tailgated and to examine related contributing factors. All of the tailgating data were divided into two parts (training part and testing part) for modeling purposes. In the final model, the accuracy of training and testing datasets were reasonably high, 74.3% and 74.7%, respectively.

Results showed that MeanTH, DT, and MinTH are the three most important variables contributing to the model accuracy. Further analysis on the characteristics of these variables showed that drivers tend to make lane changes if followed by another vehicle very closely (i.e., short time headway) or if for long durations. Specifically, the most frequent MeanTH range for tailgating events with a lane change response was less than 0.8 s, while for the events with no-lane change reaction, the MeanTHvalue was greater than 0.7 ss. These results can help with setting the parameters in countermeasure designs.

In this study, 81.8% of the tailgating events lasted less than one minute, suggesting that most drivers (either the driver of the leading vehicle or the driver of the following vehicle) will have some kind of reaction within the first minute of the tailgating event. The comparison of the duration of tailgating events between the two types of reactions (lane change or without lane change) indicated that both reaction types were frequently observed during a short period of tailgating events. However, if the duration time lasts for more than two minutes, most leading vehicles will make a lane change to end the tailgating.

The conclusions of this study were analyzed and compared with some previous studies conducted by other researchers. The conclusion of this study is consistent with the majority of the existing studies on gender differences in driving behaviors that show male drivers tend to be more aggressive (Tavris et al., 2001; Turner & McClure, 2003). However, the result of this study differs from the conclusions made by some researchers that young men have the highest risk of aggressive driving (Begg & Langley, 2001; Mast et al., 2008). The potential reasons could be that middle-aged drivers was over represented in the highway driving data in the IVBSS data set, that about 42.8% of the highway driving data were contributed by middle-aged drivers, while the percentage of highway driving data for younger and older groups were 30.3% and 26.9%, respectively (Sayer et al., 2010). Moreover, middle-aged drivers tend to have increased prevalence on highways during peak hours for commuting, than younger and older drivers. This hypothesis was shared by another study, which reported that that tailgating was more common during commuting peak hours (Tlhabano et al., 2013). In the present study, drivers were more likely to tailgate during sunny weather, in daytime conditions, and while driving in the fast lane.

The accuracy of random forests used in this article were reasonably high compared to other studies. Similar model accuracy was identified in other relevant studies (McDonald et al., 2014; Li et al., 2017). These two studies adopted the RF approach to detect drowsiness-related lane departures by using wheel angles and explore the relationship between driving styles and relevant variables in highway traffic, separately. Both models had accuracy around 75%. The results of the current study have provided evidence that the RF algorithm is an efficient and promising method to explain the influence of various variables in predicting driver behavior, such as leading vehicle reactions while being tailgated.

Three key factors including MeanTH, MinTH, and DT have the most influence on leading vehicles' reaction. Small time headways generate dangerous situations (Vogel, 2003). Time headway and the variation of time headway have a major impact on crash potential (Smith et al., 2002). Therefore, leading vehicles' drivers may feel unsafe when being tailgated with a relatively small time headway and change lanes to bring themselves into a safer situation. In addition, the leading vehicle drivers were more likely to identify the tailgating behavior with the increase of duration time, since drivers are taught to check the rearview mirrors every five to eight seconds in defensive driving (Lund & Williams, 1985). Therefore, these factors can help with setting parameters in countermeasure designs.

There are some limitations of this study. The characteristics of the leading vehicle drivers could not be obtained, as all the data were collected through the sensors located in the following vehicles. We were unable to explore the influence of the leading vehicle driver characteristics on the reaction types (e.g., age, gender, number of years of driving experience). Another limitation is that the impact of risky driving characteristics (e.g., previous driving violation records) were not included in this analysis, as this study was designed to focus on examining the types of tailgating behavior and corresponding responses. The study of how different driving styles (i.e., risky or conservative) can impact on drivers' responses to tailgating behavior should be addressed in future studies, as this is critical in the design of future advanced driving assistant systems. In addition, we only focused on tailgating behavior under free flowing traffic with relatively high traveling speeds. The influence of traffic flow in this study is not significant, which may be a results of the free traffic flow selection criterion. The effect of traffic flow can be further evaluated in future studies.

In conclusion, this study provides a unique dataset and method to study tailgating behavior from the leading vehicle perspective by utilizing multiple sensor datasets. The results of this study can improve understanding of the behavior of both the leading and following vehicles involved in the tailgating behavior. This may help to design corresponding in-vehicle based countermeasures or assisting systems to mitigate crash risks that are initiated by following vehicles. In addition, it can also inform the research and design of automated vehicles, which will undoubtedly be tailgated by some other vehicles.

5. Declarations of interest

None.

Acknowledgements

This study was sponsored by Honda R&D Americas, Inc. This article solely reflects the opinions and conclusions of its authors and not Honda R&D or any other Honda entity. The authors want to thank Mary Lynn Buonarosa, Nicholas Sweet and Miriam Manary from UMTRI for their help with proofreading the paper.

References

Bao, S., Wu, L., Yu, B., & Sayer, J. (2020). An examination of teen drivers' car-following behaviour when Compared to adult drivers (2020). *Accident Analysis Prevention, 147* 105762.

Begg, D., & Langley, J. (2001). Changes in risky driving behavior from age 21 to 26 years. *Journal of Safety Research, 32*(4), 491–499.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Bureau, A., Dupuis, J., Falls, K., Lunetta, K. L., Hayward, B., Keith, T. P., & Van Eerdewegh, P. (2005). Identifying SNPs predictive of phenotype using random forests. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society, 28*(2), 171–182.

Carter, C. J., May, A., Smith, F., & Fairclough, S. (1995). An evaluation of an in-vehicle headway feedback system. *Contemporary Ergonomics*, 287.

Dingus, T. A., McGehee, D. V., Manakkal, N., Jahns, S. K., Carney, C., & Hankey, J. M. (1997). Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human factors, 39*(2), 216–229.

Duan, J., Li, Z., & Salvendy, G. (2013). Risk illusions in car following: Is a smaller headway always perceived as more dangerous? *Safety Science, 53*, 25–33.

Evans, L., & Wasielewski, P. (1982). Do accident-involved drivers exhibit riskier everyday driving behavior? *Accident Analysis & Prevention, 14*(1), 57–64.

Galovski, T. E., Malta, L. S., & Blanchard, E. B. (2006). *Road rage: Assessment and treatment of the angry, aggressive driver*. American Psychological Association.

Harb, R., Yan, X., Radwan, E., & Su, X. (2009). Exploring precrash maneuvers using classification trees and random forests. *Accident Analysis & Prevention, 41*(1), 98–107.

Hutchinson, P. T. (2008). Tailgating. Safety Centre for Automotive Safety Research, University of Adelaide. Australia, Research Report# CASR046.

Joint, M. (1995). *Road rage*. London: Automobile Association.

LeBlanc, D. J., Bao, S., Sayer, J. R., & Bogard, S. (2013). Longitudinal driving behavior with integrated crash-warning system: Evaluation from naturalistic driving data. *Transportation Research Record, 2365*(1), 17–21.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors, 44*(2), 314–334.
Lertworawanich, P. (2006). Safe-following distances based on the car-following model. *PIARC International seminar on Intelligent Transport System (ITS) in Road Network Operations*. Malaysia: Kuala Lumpur.
Lertworawanich, P. (2009). Evaluation of dot following-distance pavement markings in Thailand (No. 09-0709). Proceedings of the Transportation Research Board 88th Annual Meeting., Washington D.C., USA.
Li, G., Li, S. E., Cheng, B., & Green, P. (2017). Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part C: Emerging Technologies, 74*, 113–125.
Lund, A. K., & Williams, A. F. (1985). A review of the literature evaluating the defensive driving course. *Accident Analysis & Prevention, 17*(6), 449–460.
Lunetta, K. L., Hayward, L. B., Segal, J., & Van Eerdewegh, P. (2004). Screening large-scale association study data: Exploiting interactions using random forests. *BMC Genetics, 5*(1), 32.
Ma, J., & Cheng, J. C. P. (2016). Identifying the influential features on the regional energy use intensity of residential buildings based on Random Forests. *Applied Energy, 183*, 193–201.
Mast, M. S., Sieverding, M., Esslen, M., Graber, K., & Jäncke, L. (2008). Masculinity causes speeding in young men. *Accident Analysis & Prevention, 40*(2), 840–842.
McDonald, A. D., Lee, J. D., Schwarz, C., & Brown, T. L. (2014). Steering in a random forest: Ensemble learning for detecting drowsiness-related lane departures. *Human Factors, 56*(5), 986–998.
Michael, P. G., Leeming, F. C., & Dwyer, W. O. (2000). Headway on urban streets: Observational data and an intervention to decrease tailgating. *Transportation Research Part F: Traffic Psychology and Behaviour, 3*(2), 55–64.
Minnesota DOT. (2006). Minnesota Tailgating Pilot Project, Report and Summary, http://www.dot.state.mn.us/trafficeng/tailgating/Tailgating-finalreport.pdf.
Mohamed Radzi, N. H., Birgin, I. S., & Mustaffa, N. H. (2017). Road traffic crash severity classification using support vector machine. *International Journal of Innovative Computing, 7*(1).
Monteiro, N. M., Balogun, S. K., Kote, M., & Tlhabano, K. (2015). Stationary tailgating in Gaborone, Botswana: The influence of gender, time of day, type of vehicle and presence of traffic officer. *IATSS Research, 38*(2), 157–163.
National Highway Traffic Safety Administration. (2015). Traffic Safety Facts, 2012 Data: Pedestrians1. Annals of Emergency Medicine, 65(4), 452.
Ngxande, M., Tapamo, J. R., & Burke, M. (2017, November). Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques. In Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), 2017(pp. 156-161). IEEE.
R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rajalin, S., Hassel, S.-O., & Summala, H. (1997). Close-following drivers on two-lane highways. *Accident Analysis & Prevention, 29*(6), 723–729.
Rämä, P., & Kulmala, R. (2000). Effects of variable message signs for slippery road conditions on driving speed and headways. *Transportation Research Part F: Traffic Psychology and Behaviour, 3*(2), 85–94.
Roadway Safety Foundation. (2001). Saving Lives: RSF recognizes nine outstanding highway safety projects, Road Safety Reporter, December Issue.
Sayer, J.R., Buonarosa, M.L., Bao, S., Bogard, S.E., LeBlanc, D.J., Blankespoor, A.D., Funkhouser, D.S. & Winkler, C.B. (2010). Integrated vehicle-based safety systems light-vehicle field operational test, methodology and results report. UMTRI-2010-30.
Smith, D. L., Najm, W. G., & Glassco, R. A. (2002). Feasibility of driver judgment as basis for a crash avoidance database. *Transportation Research Record, 1784*(1), 9–16.
Song, M., & Wang, J. H. (2010, March). Studying the tailgating issue in Rhode Island and its treatment. In 51st Annual Transportation Research Forum, Arlington, Virginia, March 11-13, 2010 (No. 207265). Transportation Research Forum.
Tavris, D. R., Kuhn, E. M., & Layde, P. M. (2001). Age and gender patterns in motor vehicle crash injuries: Importance of type of crash and occupant role. *Accident Analysis & Prevention, 33*(2), 167–172.
Tlhabano, K. N., Kote, M., Pheko, M. M., Monteiro, N., & Balogun, S. K. (2013). Don't kiss my bumper: Investigating tailgating driving behaviour in Botswana. *IOSR Journal of Humanities and Social Science, 12*(1), 34–40.
Turner, C., & McClure, R. (2003). Age and gender differences in risk-taking behaviour as an explanation for high incidence of motor vehicle crashes as a driver in young males. *Injury Control and Safety Promotion, 10*(3), 123–130.
Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accident Analysis & Prevention, 35*(3), 427–433.
Wang, J. H., & Song, M. (2011). Assessing drivers' tailgating behavior and the effect of advisory signs in mitigating tailgating. Proceedings 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, University of Iowa, Iowa City, IA.
Wang, Y., Bao, S., Du, W., Ye, Z., & Sayer, J. R. (2017). Examining drivers' eye glance patterns during distracted driving: Insights from scanning randomness and glance transition matrix. *Journal of Safety Research, 63*, 149–155.
Yao, Y., Zhao, X., Du, H., Zhang, Y., & Rong, J. (2018). Classification of distracted driving based on visual features and behavior data using a Random Forest method. *Transportation Research Record*. 0361198118796963.

Dr. Xu is current a postdoctoral fellow in the Department of Architecture and Civil Engineering of Chalmers University of Technology after he achieved his PhD in the Southeast University, China. He did his visiting PhD study in the University of Michigan Transportation Research Institute, where he conducted this study. Dr. Xu has published multiple papers in different journals.

Dr. Bao is an Associate Professor in the Industrial and Manufacturing Systems Engineering Department, University of Michigan-Dearborn, with a joint appointment as Associate Research Scientist in the University of Michigan Transportation Research Institute's Human Factors Group. She is also an affiliated faculty member with UM Civil and Environmental Engineering department, UM Robotics Institute and Institute of Data Science. Dr. Bao received her Ph.D. in mechanical and industrial engineering from the University of Iowa in 2009. Her research over the last 17 years has been rooted in the human factors issues related to connected and automated vehicle technologies, ADAS system evaluation, driver distraction, driver behavior modeling, and big data analysis. She has served as the PI (with a total funding of 3.8 million dollars) or co-PI (with a total funding of 14 million dollars) of 49 research projects for the US DOT, US DOE, and the automotive industry. She has published 63 technical publications, including 33 referred journals articles. Dr. Bao is a member of the Human Factors and Ergonomics Society and has served as the chair of the Surface Transportation Technical Group of Human Factors and Ergonomics Society. She is a member of the TRB Vehicle User Characteristics committee and the TRB Human Factors in Road Vehicle Automation subcommittee. Dr. Bao is also the author of the recent IEEE e-learning course on "Human Factors Issues in Vehicle Automation".

Dr. Pradhan is an Assistant Professor in the Mechanical and Industrial Engineering Department at the University of Massachusetts Amherst. He is a co-director of the Human Performance Laboratory. Professor Pradhan received his MS & PhD in Industrial Engineering and Operations Research at UMass Amherst. He was a National Institutes of Health Intramural Research Training Award Fellow at the National Institutes of Child Health and Human Development in Bethesda, MD, until 2013. He was then an Assistant Research Scientist at the Young Driver Behavior and Injury Prevention Group and then the Human Factors Group at the University of Michigan Transportation Research Institute (UMTRI).