Accident Analysis and Prevention 132 (2019) 105267

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Relationship between hazard-perception-test scores and proportion of hard-braking events during on-road driving – An investigation using a range of thresholds for hard-braking

Assaf Botzer[a,*], Oren Musicant[a], Yaniv Mama[b]

[a] Department of Industrial Engineering & management, Ariel University, Ariel 40700, Israel
[b] Department of Psychology, Ariel University, Ariel 40700, Israel

ARTICLE INFO

Keywords:
Hazard perception
Braking events
In-vehicle data recorder
Vehicle kinematics
Road safety

ABSTRACT

Drivers with higher proportion of hard braking events have greater potential to be involved in an accident. In this study, we tested if hard braking events might be accounted for by drivers' hazard perception (HP) ability. Our investigation was based on an original approach. Usually, researchers define hard braking according to a single deceleration threshold (e.g., $g < -0.5$). In this study, we chose different thresholds for hard braking (-0.25 to -0.6 g) and for each threshold, we examined the linkage between HP test (HPT) scores and the proportion of hard braking events. We hypothesized that this linkage would be stronger if the threshold that defines hard braking is higher. This is because the stronger the braking events, the higher the likelihood that they resulted from later detection of hazards and the lower the likelihood that they resulted from other causes (e.g., road humps). Thirty-three drivers completed an HPT and used a smartphone app that recorded their vehicle kinematics. We estimated the coefficient of HPT score in a series of binomial regression models on the proportion of hard braking events. In accordance with our hypothesis, we found that the coefficient of HPT score changed as a function of the threshold for hard braking. This finding was based on a significant negative Spearman correlation between the coefficients and the threshold and on linear functions that we derived from two binomial models that allowed the coefficient of HPT to vary according to the threshold. Our findings show that hard braking events are related to HP ability and can inform safety interventions in response to excessive proportion of hard braking events. In addition, they demonstrate that using a range of thresholds for hard braking is a practical tool in the study of hard braking events. From a theoretical perspective, our findings provide strong support to hazard perception theory.

1. Introduction

Drivers with higher frequency of extreme driving events such as hard braking and sharp turning are at greater risk of being involved in a car accident compared to their peers (Gitelman et al., 2018; Simons-Morton et al., 2012; Toledo et al., 2008). Therefore, advanced insurance programs consider the rate of driving events in their premiums to incentivize drivers to drive safely (e.g., Desyllas and Sako, 2013; Litman, 2009). Similarly, companies install in-vehicle data recorders (IVDR) in their fleets to facilitate behavioral change (Jo and Kim, 2017; Lehmann, 1996). Instances that vehicle speed and acceleration exceed predefined thresholds are marked in the IVDR's log file as driving events and can be presented to drivers as feedback to help them avoid such events in the future. And indeed, previous research demonstrated that drivers might reduce the rate of their risky driving events if they are confronted with their IVDR reports (Farmer et al., 2010; Musicant et al., 2014; Toledo et al., 2008).

However, the way people drive is determined not only by how they choose to drive (e.g., aggressive driving), but rather, it is also determined by their driving skills (Martinussen et al., 2014; Taubman-Ben-Ari et al., 2004). Therefore, to reduce driving events, one must not only focus on motivational factors, they should also investigate what skills might be linked to driving events and seek to improve them (e.g., Isler et al., 2011). One such skill might be hazard perception, which can be viewed as awareness to dangerous traffic situations (Horswill and McKenna, 2004b), or alternatively, as the ability to identify hazardous situations while driving (Borowsky et al., 2010). This ability is widely acknowledged as key to avoiding road risk (e.g., Drummond, 2000; Elander et al., 1993; Horswill and Mckenna, 2004a, b); Horswill, 2016; McKnight and McKnight, 2003; Quimbly et al., 1986; Watts and

* Corresponding author.
E-mail address: assafbo@ariel.ac.il (A. Botzer).

https://doi.org/10.1016/j.aap.2019.105267
Received 3 April 2019; Received in revised form 7 August 2019; Accepted 11 August 2019
Available online 22 August 2019

Quimby, 1979) and authorities in Britain and in several Australian states have decided that a basic score on a hazard perception test (HPT) (e.g., Horswill and McKenna, 2004b) would be a prerequisite for obtaining a driver license (Horswill et al., 2011).

In addition to reducing road risk, we suggest that higher hazard perception ability might also reduce the occurrence of driving events. This is because drivers that can detect hazards can respond to them before they develop (Fuller, 2011) and hence, they might resort less often to hard braking and swerving that one might perform to avoid road accidents (Han and Yang, 2009).

According to Saunier et al. (2011) who studied drivers' responses during 295 road conflicts, braking was the most frequent response to avoid accidents in comparison to accelerating and swerving (˜48%, 10% and 2%, respectively). Therefore, in our investigation, we tested the relationship between HPT score and the proportion of hard braking events.

Until recently, no studies have tested the relationship between HPT score and driving events. However, very recently, Hill, Horswill, Whiting & Watson (2019) recorded hard braking events that were equal or higher than 0.5 g during on-road driving and reported that indeed, their rate was negatively related to computer-based hazard perception test scores. These findings added support to the validity of HPT scores as an index for driving safety among younger drivers. In our study, we have tested the relationship between HPT score and the occurrence of hard braking events as Hill et al. (2019) did. Yet, we undertook a different approach to this investigation.

In Hill et al's. (2019) investigation, the authors noted that researchers who would choose higher deceleration values (g) as the threshold for defining a hard-braking event would risk in excluding too many safety-relevant events from their dataset. On the other hand, researchers who would choose lower deceleration values (g) as the threshold for defining a hard-braking event would risk in including too many safety-irrelevant events in the dataset. In other words, this tradeoff is between excluding too many safety-relevant (e.g., hazard-related) events from the analysis and analyzing too many braking events that any driver would perform regardless of her/his hazard perception ability (e.g., when decelerating before a traffic light). To balance between the two undesirable outcomes Hill et al. (2019) set the threshold for defining a hard-braking event in their study on 0.5 g.

In our study, we will not choose a certain threshold for defining a hard-braking event in order to balance between the two undesirable outcomes that Hill et al. (2019) described. Instead, we will use a range of thresholds for hard braking and test the possible relationship between the threshold for defining a hard-braking event and the extent to which hazard perception ability would account for the proportion of drivers' hard braking events. We maintain that this possible relationship is derived from hazard perception theory, and that testing it provides an encompassing investigation of the hypothesized relationship between hazard perception ability and hard-braking events.

According to hazard perception theory, drivers with higher hazard perception ability would detect hazards sooner in comparison to drivers with lower hazard perception ability (Borowsky et al., 2010; Horswill and McKenna, 2004b). Therefore, they would have longer time to respond to them (Fuller, 2011). Note that if this is indeed the case, then on average, the deceleration rate to avoid crashes (DRAC) of drivers with higher hazard perception ability should be lower (braking can be weaker; see Cheng et al., 2011; Guido et al., 2012) than that of drivers with lower hazard perception ability. And, similarly, on average, the proportion of hard braking events of drivers with higher hazard perception ability should be lower than that of drivers with lower hazard perception ability.

In addition, we also note that the difference between the proportion of hard braking events between drivers with different hazard perception abilities should be smaller than the average difference if we choose relatively lower thresholds to define hard braking. This is because the proportion of lower-deceleration braking events would be composed of braking events that were not necessarily related to whether hazards were detected or not (e.g., road humps). Thus, we should expect that drivers with different hazard perception abilities would differ more in their proportion of hard braking events, if we increase the deceleration (g) value that defines hard braking. In that case, hazard perception ability would be more able to account for the proportion of hard braking events.

To test this hypothesis, we ran a naturalistic driving study in which drivers drove their daily trips with a data recorder for three weeks. Our data recorder logged all braking events that exceeded 0.15 g into the dataset with their corresponding values (e.g., 0.15…0.25…0.35, etc). Next, when we analyzed the data, we chose a range of thresholds to define hard braking (0.25-0.6 g in increments of 0.01), and for each threshold, we estimated the coefficient of HPT score in a model that accounts for the proportion of hard braking events. We hypothesize that the coefficient of HPT score in the model would be more negative for higher thresholds for hard braking.

2. Method

2.1. Participants

Participants were 33 students (mean age = 24.3, SD = 2.1, 26 Females and 7 males) from the departments of Psychology and Industrial Engineering at Ariel University who participated in the study in return for a ticked in a lottery for a new mobile phone. Psychology students also received experiment participation points that they collected in partial fulfillment of the requirements for a bachelor's degree in psychology. Mean driving experience was 6.42 years (SD = 2.1).

Participants’ own mobile phones were fitted with an IVDR smartphone application that identified the occurrence of driving events (braking, accelerating etc.). All Participants were asked to use the IVDR application for the following six weeks. In the last three weeks, a second application operated to block participants' in-coming phone messages. Therefore, we report here on the data that we collected during the first three weeks, before the blocking intervention started.

2.2. Tools

2.2.1. IVDR application

We used the IVDR smartphone application of GreenRoad technologies (GreenRoad, Austin, Texas, USA) to collect vehicle kinematic indices. This IVDR is a more advanced version of the in-vehicle GreenRoad IVDR that was used in previous studies (e.g., Albert et al., 2014; Musicant and Lampel, 2010; and see Lotan et al., 2010 for a detailed description of the technical principles of operation). The app's output is based on the fusion of data from the smartphone accelerometer and GPS. In cases that the fusion of data from the accelerometer and GPS, and sometimes together with a gyroscope, does not allow computation of the vehicle orientation (e.g., because the driver held the phone and tilted it) the app stops recording until orientation can be redetermined. These principles of operation resemble those of other smartphone data recorders that are being increasingly utilized for research and practical practices (Hong et al., 2014; Pfriem and Gauterin, 2014; Romera et al., 2016; Zheng and Hansen, 2016). In the latter respect, the kinematic data that the GreenRoad application provides is used by vehicle fleets in their driving safety and fuel consumption reduction programs (e.g., Stagecoach Group, Perth, Scotlandand) and by insurance intermediaries (e.g., Arthur J. Gallagher, Illinois, USA) as the basis for negotiations on cover terms.

The resulting database from the IVDR app included 560 trips during a total of roughly 212 driving hours. The application registered approximately 7412 driving maneuvers (turns, accelerations, braking). A sample of the data log is shown in Table 1, with each line indicating a driving event. The data for each event included date and time, participant ID, g-force, event’s name, and speed. In the current study, our

Table 1
Example of Event Log.

Subject ID	Event Time	Lateral Acceleration (g)	Longitudinal Acceleration (g)	Maneuver Name	Speed (km\h)
23	10/04/2017 19:31	−0.34	0.08	Turning	52
38	10/04/2017 19:31	−0.05	−0.38	Braking	81
38	10/04/2017 19:32	0.32	−0.15	Braking	24
47	10/04/2017 19:33	−0.33	0.06	Turning	29

Fig. 1. A screenshot from a hazard perception test clip. Participants were asked to click on the mouse when they identified a hazard. In this clip, it was the cyclist (marked with a circle in the figure).

analysis focused on the occurrence of the 2739 braking events between 0.15-0.75 g.

2.2.2. *Hazard perception test (HPT)*

We used the online HPT that was developed by the Driver and Vehicle Standards Agency (DVSA- U.K.). The test is an online simulation of the HPT for obtaining a driver license in Britain. All participants received an email with instructions and a link to the test. They were informed that they were about to watch video clips that showed driving scenarios from the viewpoint of the driver (Fig. 1) and that in each of the clips they would need to click on the mouse when they think that they identified a hazard (e.g., a pedestrian about to walk into the road; a car that changes lanes). They were further informed that the first clip would be a trial clip that would not be factored into statistics. This clip, and all other clips started to play when participants clicked on a line in the middle of the screen- "click to start".

Scores for individual clips were on a scale from 1 to 5 according to the following scoring system: A 3-second period from the time when the hazard was visible was divided into 5 equal intervals of 0.6 s each. Identifying the hazard during the first 0.6 interval granted 5 points, identifying the hazard during the next 0.6 s interval granted 4 points, and so forth. Failing to identify the hazard within the 3 s period resulted in a zero score. Participants were not informed of the scoring system, yet they had been informed that the later they would identify the hazards the lower their score would be. Clicking on the mouse on a pace of 1 click per second or faster, resulted in the stopping of the clip automatically after eight or fewer clicks (depending on the pace) and in a zero score. Thus, clicking repeatedly on the mouse to capture hazards per-adventure was not an effective strategy. Participants were instructed to click on the red "exit-test" icon on the right corner of the screen after they had completed six video clips (trial + 5 test clips). We used the average clip score as the overall HPT score.

Note that HPT score for obtaining a driver license in Britain is based on a greater number of clips than in our experiment. However, as participants in our study were not tested for a driver license, we believed that they would be more inclined to try harder if the test would be shorter. Therefore, and based on the procedure of Borowsky et al. (2010) (number of clips = 6) we decided to administer a shorter test.

2.3. *Procedure*

Participants learned about the experiment from an invitation that we posted on the students' web-board. We wrote in the invitation that we invite students who are using a smartphone, hold a valid driver license for at least three years and are using an in-vehicle phone mount to participate in a study on driving behaviour. In addition, we wrote how students would be compensated for their participation. Students that wanted to participate used the name and email address that we provided in the invitation to contact the experimenter. The experimenter scheduled a single-session meeting with each of the participants, after verifying that she or he indeed meet the inclusion criteria that were described in the invitation. In the meeting, the experimenter explained to the participants that they would install an IVDR application on their smartphone. She also explained to them that during the six weeks of the experiment, the IVDR will record their driving behaviour in terms of braking intensity, angular and longitudinal accelerations and speed. In addition, the experimenter explained that participants may opt, at any time, to end their participation by removing the application or by asking her to close their IVDR account, without incurring any negative consequences. After the experimenter confirmed that the participants understood the instructions, they signed an informed consent form and the experimenter installed the IVDR application on their smartphones. The experimenter thanked the participants and asked them to contact her in case of technical problems or if they wanted her to close their IVDR account.

2.4. *Statistical analyses*

We studied the linkage between HPT scores and the proportion of hard braking events, i.e., the count of hard braking events divided by all recorded braking events, as a function of the threshold for hard braking. The formulation of the binomial regression to estimate the proportion of hard braking events is as follows:

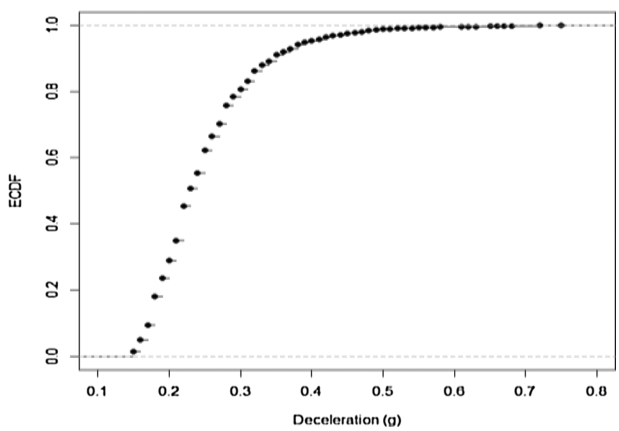

Fig. 2. Empirical cumulative distribution function of the deceleration (g) during braking events.

$$\mathrm{Ln}\left(\frac{\mathrm{E}(\mathrm{Y}_{g>T})}{N-\mathrm{E}(\mathrm{Y}_{g>T})}\right)=\beta_{1,g>T}\ \mathrm{HPT}+\ \beta_{2,g>T}\ \mathrm{G}+\ \beta_{3,g>T}\ \mathrm{S}+\ \beta_{4,g>T} \quad (1)$$

Where $\mathrm{E}(\mathrm{Y}_{g>T})$) is the expected count of braking events in which the deceleration exceeds a threshold T. N is the total count of braking events. The term $\mathrm{Ln}(\frac{\mathrm{E}(\mathrm{Y}_{g>T})}{N-\mathrm{E}(\mathrm{Y}_{g>T})})$ represents the logit link function used by the binomial regression. The explanatory variables are the hazard perception test (HPT) score, gender (G) and driving experience in years (S). We decided to include gender and driving experience in the model because numerous studies have shown that they affect driver behavior (e.g., Al-Balbissi, 2003; Maycock and Lockwood, 1993; Santamariña-Rubio et al., 2014; Williams, 1996). The terms $\beta_{1,g>T}$, $\beta_{2,g>T}$, $\beta_{3,g>T}$ in Eq.1 are the coefficients for the explanatory variables and $\beta_{4,g>T}$ is the intercept term (sometimes referred to as the "free" parameter). The term T is a vector of thresholds for hard braking.

Although our data included braking events with deceleration values as low as 0.15 g, we decided to set the minimal threshold for hard braking events on 0.25 g. This decision was based on Deligianni et al (2017) who analyzed maximal deceleration values during routine (safe) driving and found that they were usually around 0.25 g. Therefore, the values of the vector T ranged between 0.25 and 0.6 g. The increments along the vector were of 0.01 (overall 36 threshold values). We tested if the coefficients in the binomial regression model in Eq. 1changed as a function of the threshold for hard braking by estimating the coefficients thirty-six times, once for each value of the threshold.

Note that often, empirical investigations cannot test if the coefficients of the explanatory variables of hard braking events change as a function of the threshold for hard braking because the monitoring devices (e.g., off-the-shelf technologies) do not record all braking events. They only record the events that crossed a pre-set threshold for hard braking. Hence, it is impossible to know about braking events that would have been considered hard, had the threshold been different. The data in the current investigation were collected using a smartphone app that recorded all braking events and therefore, supported an analysis on the relationship between the coefficients of the explanatory variables and the threshold for hard braking.

In addition, note that in the binomial model that we present the dependent variable, i.e., the count of hard braking events divided by all recorded braking events, is a safety index that 'penalizes' less for hard braking if the overall number of braking events is larger. Hence, it accounts for two aspects of exposure; One, the expected increase in hard braking events if the driving duration/mileage is longer and two, the expected increase in hard braking events if the driving conditions require to brake more frequently (e.g. within a city or in heavy traffic in comparison to an open road).

An alternative way to control for exposure is to run a Poisson or a negative binomial model on the number of hard braking events (see Hill et al., 2019) and to introduce distance or duration to the model as the offset parameter (e.g., Musicant and Lampel, 2010; Van den Bossche et al., 2005). And indeed, a Poisson regression model that we ran on our data with duration of driving as the offset parameter yielded similar findings to those that we will report in the Results section. We decided to present the findings from the binomial model because we believe that analyzing the proportion of hard braking events provides better control for the driving conditions.

3. Results

3.1. Descriptive statistics

The mean HPT score of drivers in our study was 3.02 (SD = 0.845). The mean number of trips per driver was 16.97 (SD = 12.92) and the mean driving duration was 386.1 (SD = 315) minutes. Our data consisted of 2739 braking events across a range of longitudinal decelerations between 0.15 g and 0.75 g with a mean of 0.25 g (SD = 0.076). The empirical distribution function for the deceleration values is presented in Fig. 2. An inspection of the figure shows which events are more common, or conversely, less common in our sample. For example, the 50th percentile is at 0.23 g. The percentiles 0.85, 0.9, 0.95 and 0.99 correspond with decelerations of 0.32, 0.35, 0.39 and 0.52 g, respectively. These values represent less common braking intensities.

3.2. The coefficients of HPT score, gender, and experience in a binomial model on the proportion of hard braking events, as functions of the threshold for hard-braking

We studied the effects of HPT score, Gender, and Experience on the proportion of hard braking events, as functions of the threshold for hard braking, using a series of 36 binomial regression models, one model for each value of the threshold between 0.25 and 0.6 g. Fig. 3 presents the estimated coefficients (designated by open circles) of HPT score (upper left panel), of gender, (upper right panel), of experience (bottom left panel) and of the intercept (bottom right panel) as functions of the threshold for hard braking (x-axis). The grey area in the four panels of the figure represents the confidence intervals per each estimate.

A visual inspection of Fig. 3 shows that the coefficients in all panels change as a function of the threshold. These visual patterns are supported by inferential tests. In this section, we will use Spearman correlations as quantitative descriptions of the patterns, and in the next section, we will use more advanced statistical methods to describe the patterns in greater details.

The upper left panel of Fig. 3 shows that in accordance with our expectation, the coefficient of HPT score is more negative if the threshold for hard braking is higher. This pattern indicates that the higher the threshold for hard braking, the stronger that HPT scores would be negatively associated with the likelihood for hard braking. The Spearman correlation between the coefficient of HPT score and the threshold for hard braking is ρ=-0.75, $p < 0.001$.

The upper right panel of Fig. 3 shows that the coefficient of Gender (females are the reference group) is higher for higher values of the threshold. This pattern indicates that the higher the threshold for hard braking, the higher the likelihood that men would be more involved than females in hard-braking events. The Spearman correlation between the coefficient of males (relative to the reference group) and the value of the threshold is $\rho = 0.88$, $p < 0.001$. The lower left panel of Fig. 3 shows that the coefficient of experience becomes lower for higher threshold values (ρ=-0.77, $p < 0.001$). This pattern indicates that the higher the threshold for hard braking, the stronger that experience would be negatively associated with the likelihood for hard braking.

Apart from the coefficients of the explanatory variables, the intercept also changed as a function of the threshold. The lower right panel of Fig. 3 shows that the intercept decreases as the threshold increases and hence, as expected, hard braking events would become less

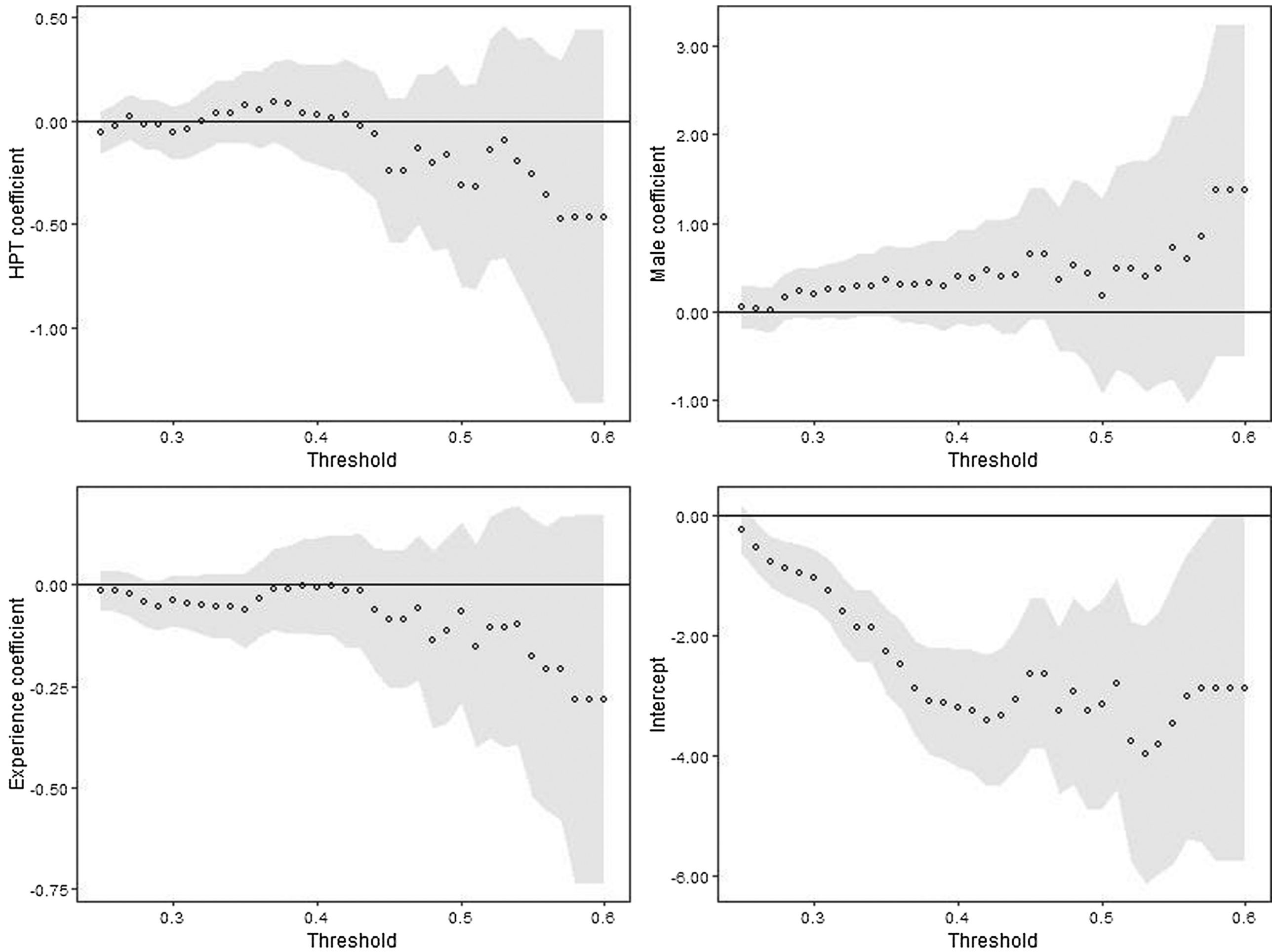

Fig. 3. Estimates of the coefficients in a binomial regression model on the proportion of hard braking events, as functions of the threshold for hard braking. The grey area represents the confidence intervals. The black horizontal line represents a zero coefficient.

probable as the threshold for hard braking is higher (ρ=-0.69, $p < 0.001$).

Note that according to Fig. 3, the coefficient of the intercept as a function of the threshold for hard braking is significantly different from zero across nearly the entire range of thresholds. In contrast, the coefficients of the explanatory variables- HPT score, gender, and experience are not significantly different from zero for any of the thresholds across that range. This can be inferred from that the grey area, which represents the confidence intervals, rarely covers the zero line (i.e., the black horizontal line) in the panel of the intercept (lower right panel) but entirely covers it in the panels for HPT score, gender and experience. Thus, while the Spearman correlations showed significant changes in the value of the coefficients as a function of the threshold, none of the single binomial regression models revealed that the coefficients were different from each other. Instead, according to the separate binomial models, none of the coefficients was significantly different than zero. We will refer to this finding in the Discussion. In the next section, we will present a finer grained description of the changes in the value of the coefficients as functions of the threshold for hard braking. We will base our description on more robust estimations of the coefficients.

3.3. A varying coefficients model on the proportion of hard braking events

In the analyses that we presented in the previous section, we calibrated a series of separate binomial regression models and estimated their coefficients. These analyses suffer from two drawbacks: (1) Each model in the series of models was calibrated without consideration of the trends in the coefficients that are now visible in Fig. 3. A more accurate estimation procedure would use a likelihood function that considers that certain coefficient estimates are more likely if they are aligned with the trend. (2) The characteristics of the coefficient trends are not yet known, for example, whether a linear trend is suitable to describe the effect of the threshold on the HPT coefficient? Or perhaps a more complex formulation is needed. To address these issues, in this section, we calibrate two Varying Coefficients (VC) models (hereafter Model 2 and Model 3) to perform an exploratory analysis of the curve of the coefficient values (Model 2) and to parameterize it (Model 3). The two VC models follow the formulation in Eq. 2.

$$\mathrm{Ln}\left(\frac{\mathrm{E}(\mathrm{Y}_{g>T})}{N-\mathrm{E}(\mathrm{Y}_{g>T})}\right)=\beta_1(\mathrm{T})\ \mathrm{HPT}+\ \beta_2(\mathrm{T})\ \mathrm{G}+\ \beta_3(\mathrm{T})\mathrm{S}+\beta_4(\mathrm{T}) \quad (2)$$

The dependent and explanatory variables are like in Eq. 1. Different than in Eq.1, the coefficients in this model are functions of the threshold T. Model 2 is based on non-parametric smooth functions (P-splines, as in Eilers and Marx (1996)) of T to represent β_1(T), β_2(T), β_3(T), and β_4(T). The smooth functions enable to model complex patterns, without setting any prior constraints (e.g., a linear relationship) regarding their characteristics. This non-parametric modeling approach (Model 2) allowed us to perform an exploratory analysis of the patterns in the coefficients that assisted us in our decision regarding the parametric formulation that we applied to these patterns in Model 3. To calibrate model 2, we used the Generalized Additive Modelling (GAM) framework from the R software. GAMs (Wood, 2006) are extensions of generalized linear models (GLM), in which the linear predictor is expressed

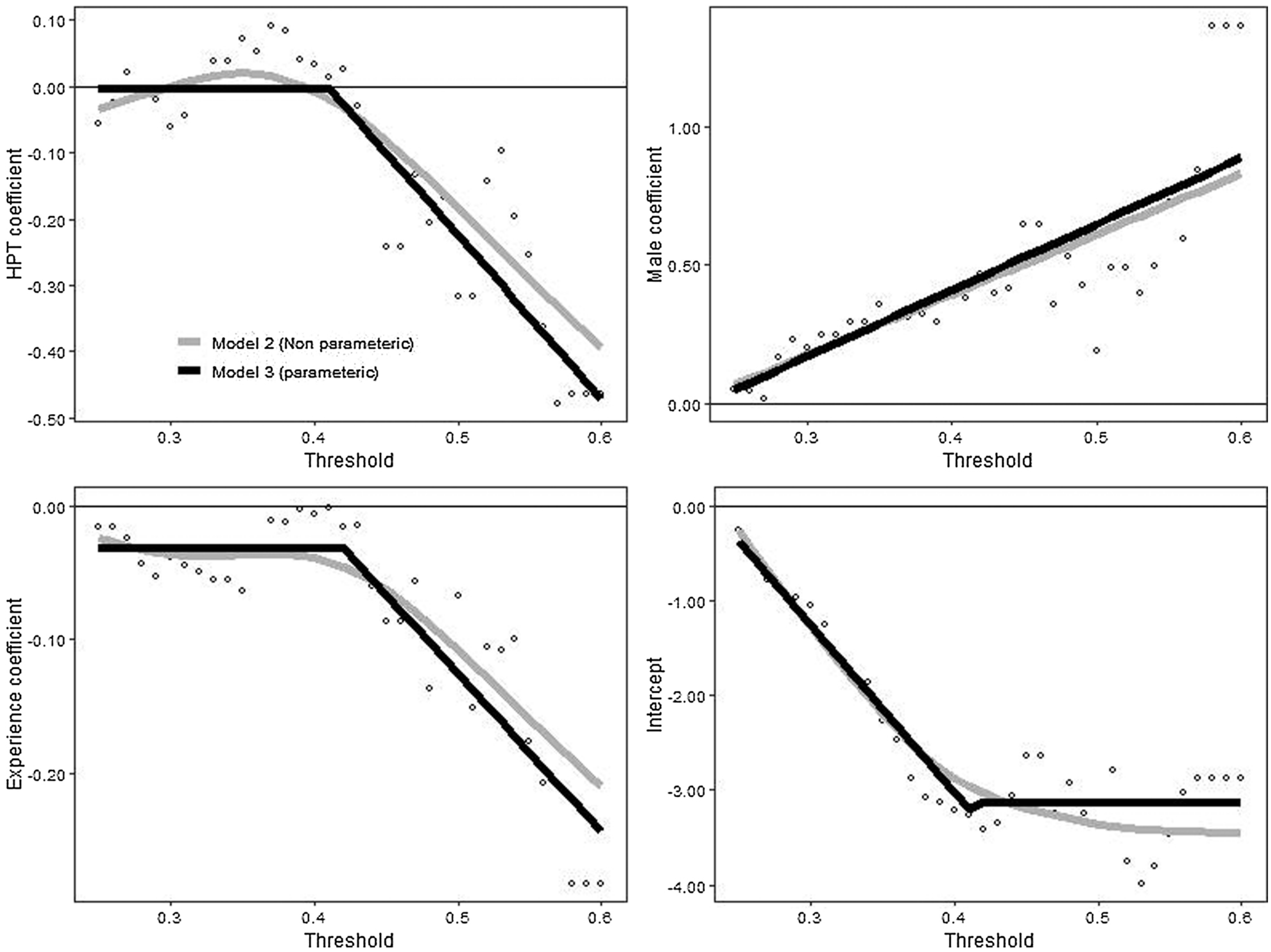

Fig. 4. Models 2 (non-parametric) and 3 (parametric) estimates of the coefficients in a binomial regression model on the proportion of hard braking events. The dots are the coefficient estimates according to Eq. 1.

as a sum of smooth functions of the covariates (The variable T in the current case). To calibrate Model 3, we used the GLM framework with a binomial response.

Fig. 4 shows the coefficients as functions of the thresholds according to Model 2 and Model 3 (solid grey and black lines, respectively). According to the upper right panel of the figure, the relationship between the coefficient of gender and the threshold for hard braking resembles a linear function. The smooth functions for HPT score, driving experience, and the intercept are more complex (upper and bottom left panels and bottom right panel, respectively). The visual analysis suggests that for lower values of the threshold (< 0.4 g) the curve for HPT score and Experience is flat, but for higher values of the threshold (> 0.4 g) the curve gradually departs for the zero line. The curve for the intercept follows a reverse trend, it is linear for the lower values of the threshold (around < 0.4 g) and flattens for the higher ones (> 0.4 g). Based on these observations, we calibrated Model 3 in which we replaced the non-parametric smooth functions (Eq.2) with the following:

$$\beta_1(\mathrm{T}) = \begin{cases} T < 0.42 & \beta_{1,0} \\ else & \beta_{1,0} + \beta_{1,1}(T - 0.42) \end{cases}$$

$$\beta_2(\mathrm{T}) = \beta_{2,0} + \beta_{2,1}(T - 0.25)$$

$$\beta_3(\mathrm{T}) = \begin{cases} T < 0.41 & \beta_{3,0} \\ else & \beta_{3,0} + \beta_{3,1}(T - 0.41) \end{cases}$$

$$\beta_4(\mathrm{T}) = \begin{cases} T < 0.41 & \beta_{4,0} + \beta_{4,1}(T - 0.25) \\ else & \beta_{4,2} \end{cases}$$

These functions describe a linear link with breakpoints at T = 0.42 for $\beta_1(\mathrm{T})$ and at 0.41 for $\beta_3(\mathrm{T})$ and for $\beta_4(\mathrm{T})$. The breakpoints were located after a search in the range [0.4, 0.45 g] for the point in which the AIC measure of fit (Akaike, 1987; Hox et al., 2017) is the lowest. Table 2 describes the estimated coefficient values for Model 3.

4. Discussion

We investigated a prediction from hazard perception theory that the

Table 2
Parameter Estimates for Model 3.

Coefficient function	Parameter Estimates (S.E)
HPT: $\beta_1(\mathrm{T})$	$\beta_{1,0}$ = -0.003 (0.016)
	$\beta_{1,1}$ = -2.462 (0.540)***
Gender: $\beta_2(\mathrm{T})$	$\beta_{2,0}$ = 0.050 (0.051)
	$\beta_{2,1}$ = 2.396 (0.445)***
Driving Experience: $\beta_3(\mathrm{T})$	$\beta_{3,0}$ = -0.032 (0.008)***
	$\beta_{3,1}$ = -1.172 (0.279)***
Intercept: $\beta_4(\mathrm{T})$	$\beta_{4,0}$ = -0.368 (0.069)***
	$\beta_{4,1}$ = -17.732 (0.326)***
	$\beta_{4,2}$ = -3.122 (0.113)***

*** P value < 0.001.

proportion of hard braking events could be accounted for by HPT scores. According to hazard perception theory, higher hazard perception ability leads to earlier detection of hazards (e.g., Hill et al., 2019; Horswill and Mckenna, 2004a,b; Kinnear et al., 2013). Earlier detection of hazards, in turn, enables drivers to take actions before hazards develop (Fuller, 2011) and the sooner that drivers respond, the less likely that they would resort to harder braking to avoid an accident (Cheng et al., 2011). Based on this formulation, we hypothesized that HPT score would have a higher negative coefficient in a model on the proportion of hard braking events, the higher that we would set the threshold for hard braking.

In accordance with our hypothesis, we found a significantly negative Spearman correlation between the threshold for hard braking and the coefficients of HPT score (see Fig. 3, upper left panel). This finding corresponds with the theoretical framework of hazard perception theory and points to a relationship between drivers' HPT score and their proportion of hard braking events. A finer grained analysis of the coefficient estimates as functions of the threshold corroborated and refined the results of the Spearman test.

A varying coefficient non-parametric model suggested that indeed, the coefficients of HPT score changed as a function of the threshold for hard braking (Fig. 4 upper left panel – Model 2). Based on the shape of the smooth function between the threshold and the HPT coefficient, we calibrated a linear model with a breakpoint to our data (Fig. 4 upper left panel - Model 3). According to Model 3, in the range between 0.25-0.42 g the coefficient is not significantly different than zero (see the value of the regression coefficients in Table 2). But, if the threshold that defines hard braking is higher than 0.42 g, the coefficient of HPT score is expected to increase (in absolute terms) as a function of the threshold. Table 2 shows that the linear increase in the value of the coefficient as a function of the threshold for hard braking is statistically significant.

Note, that despite the significant change in the value of the coefficients as a function of the threshold, none of the single binomial regression models revealed that the coefficients were different from each other. Instead, according to the single binomial models, none of the coefficients was significantly different than zero. Thus, the single binomial models might have not been powerful enough to identify that for a range of thresholds, there is a significant linkage between HPT score and the proportion of hard braking events. In addition, we demonstrated in the Results section that none of the single binomial models identified a significant linkage between experience and gender and hard braking events. The lower power of the single binomial models was the result of the relatively small sample in our study, especially with respect to the less frequent braking intensities (higher than ˜0.4 g according to Fig. 2) that were associated with relatively greater confidence intervals (see Fig. 3).

We included gender and driving experience as explanatory variables in the binomial regression models on the proportion of hard braking events because of their established effects on driver behaviour (e.g., Al-Balbissi, 2003; Maycock and Lockwood, 1993; Santamariña-Rubio et al., 2014; Williams, 1996). Like our findings on the coefficient of HPT score, we found a significant Spearman correlation between the threshold for hard braking and the coefficients of gender and experience. A positive correlation with the coefficient of gender and a negative correlation with the coefficient of experience (upper right and lower left panels of Fig. 3, respectively). These findings are in line with the existing literature on driver behaviour.

Previous studies suggest that males take higher risks than females during driving (Rhodes and Pivik, 2011; Turner and McClure, 2003). Higher risks, by their nature, would translate into stronger braking to avoid accidents (see formulations by Zhao and Lee, 2018) and thus, it is more likely that males, rather than females would perform stronger braking events. Additional findings suggest that experienced drivers are more effective in their visual search and have greater ability to match the vehicle speed to the road conditions in comparison to inexperienced drivers (McKnight and McKnight, 2003). Therefore, in accordance with our findings, it is less likely that more experienced drivers would perform stronger braking events. The findings in Fig. 3 were supported by the finer grained analyses with the varying coefficients models.

Based on the results from the non-parametric model (Fig. 4 upper right and lower left panel – Model 2) we calibrated a linear model and a linear model with a breakpoint to the coefficients of gender and experience (Fig. 4 upper right and lower left panel – Model 3, respectively). Table 2 shows that the linear increase in the value of the coefficient of gender as a function of the threshold for hard braking is statistically significant. In addition, Table 2 shows that the coefficient of experience is constant and significantly different than zero from 0.25 g until a breakpoint at 0.41 g. Beyond this breakpoint, the coefficient of driving experience is a significant negative linear function of the threshold for hard braking. This pattern is very similar to the function that we described between HPT score and the threshold.

Possibly, the limited range of HPT scores and driving experience in our study did not allow us to find an existing effect of the threshold for hard braking on the coefficients of HPT score and experience across lower braking intensities. Alternatively, it is possible that mean braking intensity across the lower range of breaking intensities is less affected from the ability to detect hazards (HPT score) or from inappropriate speed (Experience). Future studies should investigate if the patterns that we identified are recurrent and if so, what might be their determinants. Interestingly, the breakpoints that we identified (0.41 g and 0.42 g) were in the range of thresholds that researchers usually choose when they use a single threshold for hard braking (e.g., 0.4 g in Cheng et al., 2011; 0.5 g in McGehee et al., 2007; and in Hill et al. (2019).

There are several practical implications to our findings. The first pertains to the employment of the correct means in response to alerts on excessive proportion of hard braking events. We noted in the Introduction that it might be beneficial to confront drivers with their driving reports in order to modify their driving behaviour, but that this might not be enough if their behaviour is skill related. Our findings suggest that higher proportion of hard braking events might result from lower hazard perception ability. Therefore, agents that are interested in changing drivers' behaviour (e.g., fleet managers, insurance companies) might try to test the HPT score of drivers with higher proportion of driving events. Those of them with lower HPT scores might benefit from training of their hazard perception ability using methods like the ones that were reported in Isler, Starkey& Sheppard (2011). Further, when testing whether such interventions were successful one should consider analyzing hard braking events across relevant thresholds (> 0.41 g according to the current study).

A second practical implication pertains to the identification of drivers with excessive proportion of hard braking events. Our findings suggest that gender and experience affect the proportion of hard braking events. Therefore, as a first step, agents of change might target drivers who are outliers in their reference group (e.g., compare male drivers to male drivers).

Our third practical implication is a recommendation. We recommend practitioners and researchers to use recorders that record a wide range of braking-event intensities as opposed to only recording events beyond a preset threshold. Our findings suggest that such recorders are a practical tool to study the determinants of hard braking events.

From a theoretical perspective, our report extends previous findings regrading a relationship between HPT score and hard braking events (Hill et al., 2019). These previous findings were obtained with a single threshold for hard braking. We show that in accordance with hazard perception theory, the relationship between HPT score and the proportion of hard braking events is a function of the threshold for hard braking. Therefore, our findings provide further support to hazard perception theory.

We noted several limitations in the Discussion and in the Method sections of this paper. These limitations did not conceal the significant relationship between the value of the threshold for hard braking and the coefficients of HPT score, gender and experience in the binomial regression model on the proportion of hard braking events. However, certain methodological changes in future studies might lead to greater accuracy in estimating the values of the coefficients in models on the proportion of hard braking events.

First, we intend to use an HPT with a larger number of items. In the current study, we followed Borowsky et al. (2010) with respect to using only a few items in the HPT, with the aim of increasing participants' cooperation. However, it is possible that a test with a larger number of items would lead to more accurate estimations of drivers' hazard perception ability. This possibility should be tested in future studies. Findings from such studies might also contribute to the discussion on the gains from increasing the number of items in measurements of different types of constructs (see Bergkvist and Rossiter, 2007; Yeo and Neal, 2008).

Second, we intend to recruit a larger number of participants and to diversify our sample by recruiting a larger number of non-student and male drivers. We demonstrated that in accordance with previous literature (e.g., Rhodes and Pivik, 2011; Turner and McClure, 2003), the harder the braking, the greater the probability that a male driver executed it. Based on previous literature, we do not expect that this finding might be refined by an interaction between hazard perception and gender (see Drummond, 2000; Elander et al., 1993; Horswill and McKenna, 2004b; Quimbly et al., 1986). Nevertheless, a study with a larger number of male drivers would enable us to provide better estimations of the differences between male and female drivers in the proportion of hard braking events with respect to different decelerations. The additional suggested change in the sampling procedure, namely, to recruit a larger number of non-student drivers would assist in increasing the mundane realism of future studies on the determinants of driver performance (see Druckman and Kam, 2011). Finally, a larger number of participants, overall, would assist in gaining better parameter estimates for two reasons: It would reduce the variance in the data, and would enable us to include additional explanatory variables in our models like age and perception of driving skill (Elander et al., 1993).

In addition to increasing the number of test scenarios and participants we intend to use video data to boost our classification of driving events. Gitelman et al. (2018) were able to study the link between braking events, road crashes and road infrastructure using IVDR data alone. Nevertheless, they noted that more accurate identification of driving events can be obtained from crossing IVDR data with video footage.

5. Conclusions

We showed that the coefficient of HPT score in a binomial regression model on the proportion of hard braking events was more negative as the threshold for hard braking was set to higher deceleration values. This finding corresponds with the theoretical framework of hazard perception theory and points to a relationship between drivers' HPT score and their proportion of hard braking events. Based on our findings, we suggest that assessing drivers' hazard perception ability might help in employing the correct means in response to alerts on excessive proportion of hard braking events. In future studies we intend to investigate the relationship between hazard perception ability and braking events with a larger sample of drivers, with different tests of hazard perception ability and with various measurement tools.

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

This study was funded by Ariel University. We thank GreenRoad technologies (GreenRoad, Austin, Texas, USA) for letting us use their IVDR application and for their support in storing and extracting the IVDR data. We thank the two research assistants Liad Nakash and Ariel Veig for their contribution to this study.

References

Akaike, H., 1987. Factor analysis and AIC. Selected Papers of Hirotugu Akaike. Springer, New York, NY, pp. 371–386.
Al-Balbissi, A.H., 2003. Role of gender in road accidents. Traffic Inj. Prev. 4 (1), 64–73.
Albert, G., Lotan, T., Toledo, T., Grimberg, E., Lasebnik, M., 2014. Are young drivers as careful as they deem? In vehicle data recorders and self-reports evaluations. Eur. Transp. Res. Rev. 6 (4), 469–476.
Bergkvist, L., Rossiter, J., 2007. The predictive validity of multipleitem versus single-item measures of the same constructs. J. Mark. Res. 44 (2), 175–184.
Borowsky, A., Shinar, D., Oron-Gilad, T., 2010. Age, skill, and hazard perception in driving. Accid. Anal. Prev. 42 (4), 1240–1249.
Cheng, B., Lin, Q., Song, T., Cui, Y., Wang, L., Kuzumaki, S., 2011. Analysis of driver brake operation in near-crash situation using naturalistic driving data. Int. J. Automot. Eng. Technol. 2 (4), 87–94.
Deligianni, S.P., Quddus, M., Morris, A., Anvuur, A., Reed, S., 2017. Analyzing and modeling drivers' deceleration behavior from normal driving. Transp. Res. Rec. 2663 (1), 134–141.
Desyllas, P., Sako, M., 2013. Profiting from business model innovation: evidence from pay-as-you-drive auto insurance. Res. Policy 42 (1), 101–116.
Druckman, J.N., Kam, C.D., 2011. Students as experimental participants. Cambridge Handbook Of Experimental Political Science 1, 41–57.
Drummond, A.E., 2000. Paradigm lost! Paradise gained? An Australian's perspective on novice driver safety. In Novice Drivers Conference, 2000, Bristol.
Elander, J., West, R., French, D., 1993. Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. Psychol. Bull. 113 (2), 279.
Farmer, C.M., Kirley, B.B., McCartt, A.T., 2010. Effects of in-vehicle monitoring on the driving behavior of teenagers. J. Safety Res. 41 (1), 39–45.
Fuller, R., 2011. Driver control theory: from task difficulty homeostasis to risk allostasis. In: Porter, B. (Ed.), Handbook of Traffic Psychology. Academic Press, London.
Eilers, P.H.C., Marx, B.D., 1996. Flexible smoothing with B-splines and penalties. Stat. Sci. 11 (2), 89–121.
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. Transp. Res. Part C Emerg. Technol. 91, 156–175.
Guido, G., Vitale, A., Astarita, V., Saccomanno, F., Giofré, V.P., Gallelli, V., 2012. Estimation of safety performance measures from smartphone sensors. Procedia-Social and Behavioral Sciences 54, 1095–1103.
Han, I., Yang, K.S., 2009. Characteristic analysis for cognition of dangerous driving using automobile black boxes. Int. J. Automot. Technol. 10 (5), 597–605. https://doi.org/10.1007/s12239-009-0070-9.
Hill, A., Horswill, M.S., Whiting, J., Watson, M.O., 2019. Computer-based hazard perception test scores are associated with the frequency of heavy braking in everyday driving. Accid. Anal. Prev. 122, 207–214.
Hong, J.H., Margines, B., Dey, A.K., 2014. A smartphone-based sensing platform to model aggressive driving behaviors. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 4047–4056 ACM.
Horswill, M.S., Anstey, K.J., Hatherly, C., Wood, J.M., Pachana, N.A., 2011. Older drivers' insight into their hazard perception ability. Accid. Anal. Prevent. 43, 2121–2127.
Horswill, M.S., Mckenna, F.P., 2004a. In: Banbury, S., Tremblay, S. (Eds.), A Cognitive Approach to Situation Awareness: Theory and Application. Aldershot. Ashgate Publishing, pp. 155–175.
Horswill, M.S., McKenna, F.P., 2004b. Drivers' hazard perception ability: situation awareness on the road. In: Banbury, S., Tremblay, S. (Eds.), A Cognitive Approach to Situation Awareness. Ashgate, Aldershot, UK, pp. 155–175.
Horswill, M.S., 2016. Hazard perception in driving. Curr. Dir. Psychol. Sci. 25 (6), 425–430. https://doi.org/10.1177/0963721416663186.
Hox, J.J., Moerbeek, M., Van de Schoot, R., 2017. Multilevel Analysis: Techniques and Applications. Routledge.
Isler, R.B., Starkey, N.J., Sheppard, P., 2011. Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. Accid. Anal. Prev. 43 (5), 1818–1827.
Jo, J., Kim, H., 2017. Development of a Safety Index to Identify Differences in Safety Performance by Postal Delivery Motorcyclists Based either in Different Regional Post Offices or within the Same Regional Office. ISPRS Int. J. Geoinf. 6 (11), 324.
Kinnear, N., Kelly, S.W., Stradling, S., Thomson, J., 2013. Understanding how drivers learn to anticipate risk on the road: a laboratory experiment of affective anticipation of road hazards. Accid. Anal. Prev. 50, 1025–1033.
Lehmann, G., 1996. The features of the accident data recorder and its contribution to road safety. Proceedings: International Technical Conference on the Enhanced Safety of Vehicles 1565–1568 Vol. 1996.
Litman, T., 2009. Pay-As-You-Drive Pricing For Insurance Affordability. Victoria Transport Policy Institute.

Lotan, T., Albert, G., Ben-Bassat, T., Ganor, D., 2010. Potential Benefits of In-vehicle Systems for Understanding Driver Behaviour. European Comission. Tech. Rep. Grant, Brussels, Belgium.
Martinussen, L.M., Møller, M., Prato, C.G., 2014. Assessing the relationship between the driver behavior questionnaire and the driver skill inventory: revealing sub-groups of drivers. Transp. Res. Part F: Traffic Psychol. Behav. 26, 82–91.
Maycock, G., Lockwood, C.R., 1993. The accident liability of British car drivers. Transp. Rev. 13 (3), 231–245.
McGehee, D.V., Raby, M., Carney, C., Lee, J.D., Reyes, M.L., 2007. Extending parental mentoring using an event-triggered video intervention in rural teen drivers. J. Safety Res. 38, 215–227.
McKnight, A.J., McKnight, A.S., 2003. Young novice drivers: careless or clueless? Accid. Anal. Prev. 35 (6), 921–925.
Musicant, O., Lampel, L., 2010. When technology tells novice drivers how to drive. Transp. Res. Rec. 2182 (1), 8–15.
Musicant, O., Bar-Gera, H., Schechtman, E., 2014. Temporal perspective on individual driver behavior using electronic records of undesirable events. Accid. Anal. Prev. 70, 55–64.
Pfriem, M., Gauterin, F., 2014. Employing smartphones as a low-cost multi sensor platform in a field operational test with electric vehicles. 2014 47th Hawaii International Conference on System Sciences 1143–1152.
Quimbly, A.R., Maycock, G., Carter, I.D., Dixon, R., Wall, J.G., 1986. 'Perceptual Abilities of Accident Involved Drivers', Research Report 27. Transport and Road Research Laboratory, Crowthorne, UK.
Rhodes, N., Pivik, K., 2011. Age and gender differences in risky driving: The roles of positive affect and risk perception. Accid. Anal. Prev. 43 (3), 923–931.
Romera, E., Bergasa, L.M., Arroyo, R., 2016. Need data for driver behaviour analysis? Presenting the public UAH-DriveSet. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) 387–392.
Santamariña-Rubio, E., Pérez, K., Olabarria, M., Novoa, A.M., 2014. Gender differences in road traffic injury rate using time travelled as a measure of exposure. Accid. Anal. Prev. 65, 1–7.
Saunier, N., Mourji, N., Agard, B., 2011. Using data mining techniques to understand collision processes. CIRRELT.
Simons-Morton, B.G., Zhang, Z., Jackson, J.C., Albert, P.S., 2012. Do elevated gravitational-force events while driving predict crashes and near crashes? Am. J. Epidemiol. 175 (10), 1075–1079.
Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O., 2004. The multidimensional driv- 583 ing style inventory - Scale construct and validation. Accid. Anal. Prev. 36, 323–332.
Toledo, T., Musicant, O., Lotan, T., 2008. In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transp. Res. Part C Emerg. Technol. 16 (3), 320–331.
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. Inj. Control Saf. Promot. 10 (3), 123–130.
Van den Bossche, F., Wets, G., Brijs, T., 2005. Role of exposure in analysis of road accidents: a Belgian case study. Transp. Res. Rec. 1908 (1), 96–103.
Watts, G.R., Quimby, A.R., 1979. Design and Validation of a Driving Simulator, Report LR 907. Transport and Road Research Laboratory, Crowthorne, UK.
Williams, A.F., 1996. Magnitude and characteristics of the young driver crash problem in the United States. New to the Road: Reducing the Risks for Young Motorists. Proceedings of the First Annual International Symposium of the Youth Enhancement Service.
Wood, S.N., 2006. Generalized Additive Models: an Introduction With R. CRC, BocaRaton, Florida.
Yeo, G., Neal, A., 2008. Subjective cognitive effort: A model of states, traits, and time. J. Appl. Psychol. 93 (3), 617.
Zhao, P., Lee, C., 2018. Assessing rear-end collision risk of cars and heavy vehicles on freeways using a surrogate safety measure. Accid. Anal. Prev. 113, 149–158.
Zheng, Y., Hansen, J.H., 2016. Unsupervised driving performance assessment using free-positioned smartphones in vehicles. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) 1598–1603.