Trip characteristics impact on the frequency of harsh events recorded via smartphone sensors

Abstract

The Internet of Things (IoT) constantly offers new opportunities and features to monitor and analyze driver behavior through wide use of smartphones, effective data collection and Big Data analysis, resulting in assessment and improvement of driver behavior and safety. The objective of the present study is to investigate the impact of detailed trip characteristics on the frequency of harsh acceleration and harsh braking events through an innovative smartphone application developed within the framework of BeSmart project. A 200-driver naturalistic experiment spanning 12 months is carried out since July 2019. During the first two months, participants were asked to drive in the way they usually did, without receiving any feedback on their driving behavior from the application. Over the subsequent two months, participants were provided with personalized feedback, a trip list and a scorecard regarding their driving behavior, allowing them to identify their critical deficits or unsafe behaviors. Some of the most important risk factors, such as speed and driving above the speed limit, usage of mobile phone while driving and harsh events (acceleration and braking) are recorded through the application and subsequently analyzed. Generalized Linear Mixed-Effects Models were fitted to the trips of car drivers who made frequent trips for both experiment phases in order to model the frequencies of harsh events. Results indicate that maximum speed, the percentage of speeding duration and total trip duration are positively correlated with both harsh acceleration and harsh braking frequencies. On the other hand, the exposure metric of total trip distance was found to be negatively correlated with both harsh event types. A small positive correlation of the percentage of mobile use duration with harsh accelerations was also detected.

Keywords: road safety; driver monitoring; naturalistic experiment; smartphone application; harsh events; Generalized Linear Mixed-Effects Models

1. Introduction

Measuring driving efficiency has been the focus of many studies in driving behavior literature in the past [1, 2]. From a road safety perspective, it is extremely significant to identify the parameters that affect driving behavior and therefore crash risk. It is only when these parameters are quantified that proper road safety measures can be effectively taken.

There is a significant number of risk factors affecting crash probability identified in literature. The most important risk factors recognized in the literature [3, 4] are human factors (speeding, distracted driving, driving under the influence of alcohol and other psychoactive substances etc.), unsafe road infrastructure, unsafe vehicles and inadequate law enforcement of traffic laws. Among them, human factors are likely to be the most crucial cause of road traffic fatalities and injuries every year and therefore the importance of studying how these factors can affect crash risk is high [5]. The predominance of human factors as crash causes is indicated by the respective percentage, which has been estimated to amount to up to 94% [6].

1.1. Human risk factors

Distracted driving, as one of the most important human factors that influence crash risk, has been attracting the attention of researchers the past decades. Mobile phone use (handheld or hands-free) and complex conversation (at mobile phone or with passengers) appear to be the most critical in-vehicle distraction factors [7]. Given that mobile usage is an inevitable part of everyday driving process and is expected to increase over the years [8], its impact on driving behavior in traffic and road safety is particularly crucial and merits further investigation. The literature so far has showed that when drivers are using mobile phones while driving, several impacts manifest on their behavior expressed in terms of loss of control, response to incidents, or crash occurrence [9, 10].

Speeding has also been the subject of extensive research in the transportation field. Excess and inappropriate speed are responsible for a high proportion of the mortality and morbidity that result from road crashes [11, 12]. In high-income countries, speed contributes to about 30% of deaths on the road, while in some low-income and middle-income countries, speed is estimated to be the main contributory factor in about half of all road crashes [13]. Elvik et al. [14] engaged in relevant research by first questioning whether speed is still as important for road safety as it was in the past, taking into account the penetration of constantly evolving vehicle safety systems in the global automotive market. After reviewing recent research studies regarding the impact of speed on road safety, they conclude that speed remains an important risk factor both for crash occurrence and for injury severity.

Furthermore, apart from speeding and distracted driving, more recent studies highlight the importance of investigating the phenomenon of harsh events in greater detail, as they have been associated with driving risk assessment, risk level correlation and classification [15, 16, 17]. This is because harsh driving events, such as harsh accelerations and harsh brakings, indicate an overall aggressive and unsafe driving behavior; unsafe distance from adjacent vehicles, possible near misses, lack of concentration, increased reaction time, poor driving judgement or low level of experience. From a research scope, during recent years harsh or safety-critical events have been adopted as crash surrogate measures and the understanding of related opportunities and challenges in their interpretation is under increased examination (e.g. [18, 19]).

Although both harsh accelerations and harsh brakings are associated with unsafe driving behavior, they constitute two different types of event and should therefore examined as such. More specifically, harsh acceleration events may reveal high levels of anxiety and anger while driving [20, 21] leading to a risky driving behavior characterized by drivers' involvement in situations of high risk. Harsh brakings may indicate driver struggle to anticipate the occurrence of a critical situation, which most of the times would not have occurred at the first place if it were not for driver's inattention, high speed development, inadequate distances from adjacent vehicles and other unsafe behavior indicators. As a result, harsh breaking events harsh are often used to locate safety critical events in Naturalistic Driving (ND) data [22, 23].

Additionally, given the strong correlation between harsh events and driving risk, it is not surprising why harsh accelerations and harsh brakings have been investigated by insurance industry in the context of usage-based motor insurance (UBI) schemes [24, 25], allowing for more behavioral parameters being used in UBI models. Harsh events, in combination with other driving behavioral indicators such as speeding and distracted driving are being increasingly used by Pay How You Drive (PHUD) Usage Based Insurance schemes as the critical risk factor indicators in terms of driving behavior [26].

1.2. Smartphone data exploitation

The importance of driver monitoring is progressively established in the transportation field; despite that, researchers have been struggling with the difficulty of collecting accurate real-time driving data by adopting low-cost collection and processing methods. In that environment, the high penetration rate of smartphones and social networks nowadays provide new opportunities and features to monitor and analyze driver behavior. Apart from the wide smartphone application capabilities and the low cost and ease of use in data collection, experiments under naturalistic conditions with the use of smartphones allow for drivers to be recorded under normal driving conditions and without any influence from external parameters, resulting at being considered as one of the most appropriate methods for the assessment of driving behavior [27].

Smartphones are equipped with a variety of sensors, such as motion sensors (e.g. accelerometer and gyroscope), position sensors (e.g. magnetometer), global navigation satellite system (GNSS) receivers, environmental sensors (barometers, photometers, and thermometers), microphone, cameras, etc. As a result, the exploitation of the various sensors for the purpose of transportation and road safety research allows for continuous, inexpensive and fast data collection, with plenty of studies confirming and even improving the reliability of smartphone measurement data implementing state-of-the-art machine learning and big data algorithms [28, 29]. Vlahogianni and Barmpounakis [30] examined the use of smartphones as an alternative for driving behavior analysis and they concluded that the smartphone-based algorithms may accurately detect four distinct patterns (braking, acceleration, left cornering and right cornering) with an average accuracy comparable to other popular detection approaches based on data collected using a fixed position device.

Many studies have shown promising results using data collected through smartphone sensors under naturalistic driving conditions. By conducting naturalistic driving experiments by means of mobile phone, researchers aim either at examining the effect of various driving behavior indicators on driver performance and cash risk [31, 32, 33] or at identifying aggressive and dangerous driving profiles through a clustering approach [34, 35, 36]. Going one step further, smartphones have proven to be an extremely useful feedback tool, allowing drivers to get informed about their weak points in regards with safety, namely speeding and aggressive driving style [37] as well as eco-driving [38, 39]. The ultimate objective when providing feedback to drivers is to trigger their learning and self-assessment process and enable them to gradually improve their performance and monitor their evolution [40]. Toledo and Shiftan [41] found that feedback can lead to a reduction of 8% in safety incidents, and 3–10% in fuel consumption, with a higher reduction obtained for large vehicles.

In light of the aforementioned, the aim of the current research is the utilization of high-resolution smartphone data for the investigation of the impact of detailed trip characteristics on the frequency of harsh acceleration and harsh braking events. To that end, a naturalistic driving experiment was carried out in order to examine driving behavior as expressed by the frequencies of harsh accelerations and harsh brakings.

2. Methodology

2.1. Overview of the experiment

Within the framework of BeSmart Project, a 200-driver naturalistic experiment spanning 12 months is conducted starting from July 2019. The objectives of the experiment include primarily the identification of critical risk factors through driver monitoring via an innovative smartphone application (described in the following section), and subsequently, the development of driver feedback features allowing to inform, notify and motivate the drivers to improve their critical skills and reduce their driving errors and therefore their crash risk.

More specifically, the designed experiment consists of 6 different phases differing in the type of feedback provided to drivers. In the present paper, the authors focus exclusively on the two first phases of the experiment. The first phase lasted for 12 weeks; participants were asked to drive in the way they usually did, without receiving any feedback on their driving behavior from the smartphone application. At this phase, only the Trip List and the vehicle characterization were accessible to the application user (Figure 1a). The purpose of this phase, referred to as “Phase

1” in the following, was to identify the naturalistic driving characteristics of drivers, which are used to formulate a baseline for future comparison. An initial analysis of car driver trips during the first phase of BeSmart can be found in Ziakopoulos et al. [42]. The second phase, referred to as “Phase 2” in the following, lasted for 10 weeks; participants were provided with personalized feedback, a trip list and a scorecard regarding their driving behavior, enabling them to identify their critical deficits or unsafe behaviors (Figure 1b).

The respective feedback was provided to participant smartphones through the application, each time a trip was completed. More specifically, Figure 1b (left) demonstrates the Trip List which now enables the per trip score on a 0-100 rating scale. As shown on the right side of Figure 1b, the Scorecard is subsequently introduced, allowing for additional feedback score on the four driving behavior indicators (from the left to the right); speeding, mobile phone use, harsh breaking and harsh acceleration. As a note, the BeSmart Project is conducted in Greece, therefore the original application has an exclusively Greek interface. However, for the purposes of this paper, the application interface has been translated in English.

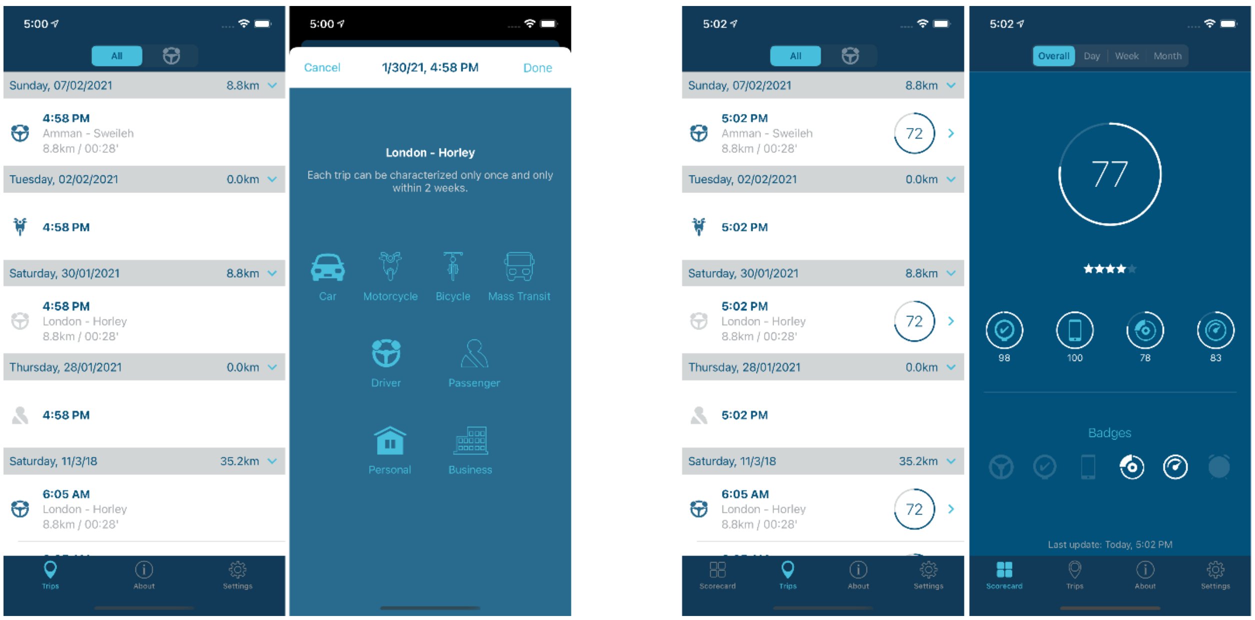

Figure 1 – Example screenshots from the application features in: (a) Phase 1 (Baseline) and (b) Phase 2 (Feedback)

2.2. The BeSmart Application

In order to achieve the research objective, an innovative smartphone application was developed aiming at the assessment and improvement of driver behavior and safety. The application is developed by OSeven (www.oseven.io), aiming to record driver behavior using the hardware sensors of the smartphone device. Furthermore, a variety of APIs is exploited to read sensor data and temporarily store them to the smartphone’s database before transmitting them to the central (back-end) database. Data collected from the application has been utilised in earlier research papers which also feature additional details regarding the application [33].

The standard procedure that is followed every time a new trip is recorded by the application is clearly presented in Figure 1. The data collected are highly disaggregated in space and time. Once stored in the backend cloud server, they are converted into meaningful driving behavior and safety indicators, using signal processing, Machine Learning (ML) algorithms, Data fusion and Big Data algorithms. This is achieved by using state-of-the-art technologies and procedures, which operate in compliance with standing Greek and European personal data protection legislation (GDPR).

Fig. 2. The OSeven data flow system.

A variety of different metadata are eventually calculated, including indicatively the following exposure indicators:

- Total distance (mileage)
- Driving duration
- Type(s) of the road network used (given by GPS position and integration with map providers e.g. Google, OSM)
- Time of the day driving (rush hours, risky hours)

The driving behavior indicators that are also calculated from the data include indicatively:

- Speeding (duration of speeding, speed limit exceedance etc.)
- Number and severity of harsh events
- Harsh braking (longitudinal acceleration)
- Harsh acceleration (longitudinal acceleration)
- Distraction from mobile phone use (mobile phone use is considered any type of phone use by the driver e.g. talking, texting etc.).

Harsh event frequencies are the response variable of this research; harsh events are determined by the OSeven algorithms where the authors do not have access due to intellectual property protection of these algorithms. More specifically, OSeven uses data from all the axes of the accelerometer as well as GPS for the determination of harsh events. Harsh events are calculated via data fusion and machine learning algorithms and not a rule-based approach using as input the values of the accelerometer as well as values from additional sensors (e.g. orientation, magnetometer, GPS, gyroscope). These algorithms evaluate the processed time series from the smartphone sensors of the complete trip. Although the applied algorithms increase the accuracy of harsh events detection; by nature, they do not include specific threshold values or explicit explanations that could be presented in the existing paper for verification purposes. The reliability of the OSeven algorithms has been extensively evaluated against literature data, OBD data, on-road experiments by certified experts on the assessment of driving behavior, and experiments on driving simulators [43, 44]. It is noted that the OSeven product has already been adopted and used by major insurance companies in several countries, which is evidence regarding the acceptance of the OSeven algorithms.

3. Theoretical Background

This analysis aims to examine the impact of detailed trip characteristics on event frequencies across drivers (i.e. the number of harsh accelerations and harsh brakings per trip). Since harsh events can be considered as instances similar to road crashes, but more frequent, a statistical method suitable for dealing with frequency data (in other words, count data) can be implemented. Therefore, Generalized Linear Models (GLMs) which are used when dealing with event count data [45] were selected for the statistical analysis. Previous studies have also used similar analytical methods when dealing with harsh events [46, 47]

However, one may also consider that in the present dataset there are repeated measurements (trips) over the same units (drivers). Due to these repeated measurements, the observations are no longer independent, as required by the assumptions of linear models. Unless accounted for, this dependency may affect the accuracy of the modelling results. In fact, it is necessary to account for random heterogeneity due to differences between drivers, so as to make sure that the effects identified in the model are true effects of the independent variables on the dependent, and do not reflect unobserved differences between drivers.

Therefore, in order to capture personal driver traits, such as personality and experience, which affects their driving style, and thus the frequencies of events they exhibit that are unobserved, random effects are introduced to GLMs in order to extend them to Generalized Linear Mixed-Effects Models (GLMMs). In other words, GLMMs are formulated by using the fixed effects GLMs as basis and then adding random effects to the equation of the linear predictor. GLMMs were introduced by Breslow & Clayton [48].

The general form of the GLM models the log odds via the linear predictor. Following McCulloch [49], if y are the observed frequencies of harsh events per trip i (harsh brakings and harsh accelerations separately), and λ are the expected frequencies of harsh events to be predicted, then the model is specified as:

$$y_i \sim Poisson\ (\lambda_i) \tag{1}$$

And the linear predictor is:

$$log(\lambda_i) = \beta_0 + \beta_n x_n + \varepsilon \tag{2}$$

Where β are the fixed-effect parameters (constant and coefficients) for n independent variables, and ε is the error term. The GLM is then extended into a GLMM by adding random effects. Random effects in GLMMs are expressed as random variable coefficients (random slopes) or random intercepts. For a GLMM containing a random intercept and a random slope for a single independent variable j of the total n, Equation (2) would be formulated as:

$$log(\lambda_i) = \beta_{0i} + \beta_{ji} x_{ji} + \ \beta_{n-1} x_{n-1} + \varepsilon \tag{3}$$

Where β_{0i} and β_{ji} follow normal distributions centered at the value of their fixed counterparts:

$$\beta_{0i} \sim N\left(\beta_0, \sigma_{s,0}^2\right) \tag{4}$$
$$\beta_{ji} \sim N(\beta_j, \sigma_{s,j}^2) \tag{5}$$

Coefficient result interpretation is more intuitive when using relative risk ratios (sometimes called incidence rate ratios). Relative risk ratios are obtained by transforming the predictor to obtain the frequency. For an increase of one unit in one specific variable, k, with all other parameters remaining equal, the predicted original frequency λ_i is multiplied by: $\lambda_{ki} = \exp(\beta_{ki}) * \ \lambda_i$

As McCulloch [50] mentions, random effect models may use correlated independent variables as input, circumventing the limitations of traditional GLMs. Furthermore, it should be mentioned that for computational reasons during the GLMM fitting, the trip data underwent z-score scaling, a common standardization process which does not affect the obtained coefficients. Mathematically, for every parameter x with a mean $\bar{x}$ and a standard deviation S a scaled value is obtained:

$$x_{scaled} = (x - \bar{x})/S \tag{6}$$

The best-fitting model which contains the more informative variable combination and explains the highest degree of variance per given dataset is selected as the one with the minimum Akaike Information Criterion (AIC). It is critical to note that the added value of any random effects is assessed by conducting a custom ANOVA (log-likelihood test) between the fixed effects GLM and any formulated GLMMs.

4. Analysis and Discussion

4.1 Descriptive Statistics

Overall, during the first two phases of the experiment, 26,619 trips from a sample of 147 car drivers have been recorded. However, for the present analysis it was decided that the final sample should consist of drivers who have participated equally in both phases only. An additional criterion was set; all drivers chosen to be included in the analysis were required to have driven at least for 40 trips. This number approximately equals the typical monthly number of working trips for a driver assuming that each driver drives 2 trips per day for 5 working days per week. This number is reasonable to filter out drivers for which there are not enough observations, and it is also the 'industrial' criterion set by OSeven to start providing driver evaluation. As a result, from the 147 car drivers, 65 were ultimately selected creating a large dataset of 21,167 trips. Demographic information regarding the drivers' gender and age are shown in Table 1.

Table 1. Overview of the selected sample

	Age groups			
	<25	25-55	>55	Total%
Male	0	27	3	46%
Female	3	31	1	54%
Total %	5%	89%	6%	100%

The key indicator, and response variable for the purpose of this research is the frequency of harsh acceleration and harsh braking events during the two first experiment phases. Additional basic road safety indicators (i.e. possibly suggesting risky or reckless behavior) are the following: exceeding the speed limit (percentage of time driving over the speed limits per trip driving duration), and mobile usage (percentage of driving time using the mobile phone per trip driving duration). On the basis of the literature review results [15,16], it is assumed that the occurrence of harsh events is partly correlated with drivers' speeding behavior and mobile phone use, and thus some of its variance may be explained by changes in those variables.

Table 2 provides a description of the variables selected. Regarding the harsh event frequencies, it is noted that both during Phase 1 and Phase 2 drivers seem to incur more harsh brakings than harsh accelerations during their trip.

Table 2. Description of the variables used in the analysis

Variable	Description
Total Trip Duration [s]	Total trip duration [sec]
Total Trip Distance [km]	Total trip distance [km]
Average Speed [km/h]	Mean driving speed per trip [km/h]
Maximum Speed [km/h]	Maximum of driving speed per trip [km/h]
Percentage of Mobile Use Duration [%]	Share of mobile use per trip [%]
Percentage of Speeding Duration [%]	Share of time over the speed limit per trip [%]
Harsh accelerations [count]	Harsh acceleration events per trip [count]
Harsh brakings [count]	Harsh braking events per trip [count]

The descriptive statistics of the parameters that were recorded per trip for both experiment phases are shown in Table 3 for Phase 1 and Table 4 for Phase 2.

Table 3. Descriptive statistics of the per trip values of the variables recorded for Phase 1

Variable	Mean	Minimum	Median	Maximum	St. Dev.
Total Trip Duration [s]	992.55	61.00	696.00	17642	1051.86
Total Trip Distance [km]	10.17	0.50	4.8.00	387.20	20.14
Average Speed [km/h]	37.27	14.00	33.00	59.00	15.35
Maximum Speed [km/h]	68.64	19.00	64.00	206.00	26.68
Percentage of Mobile Use Duration [%]	4.00	0.00	0.00	93.80	10.00
Percentage of Speeding Duration [%]	5.00	0.00	1.00	70.00	9.00
Harsh accelerations [count]	0.76	0.00	0.00	19.00	1.52
Harsh brakings [count]	1.51	0.00	1.00	33.00	2.34

Table 4. Descriptive statistics of the per trip values of the variables recorded for Phase 2

Variable	Mean	Minimum	Median	Maximum	St. Dev.
Total Trip Duration [s]	950.00	61.00	687.00	11349.00	904.81
Total Trip Distance [km]	8.25	0.50	4.30	325.50	14.09
Average Speed [km/h]	34.55	14.00	30.00	56.35	14.07
Maximum Speed [km/h]	64.90	18.00	59.00	228.00	26.42
Percentage of Mobile Use Duration [%]	3.00	0.00	0.00	93.20	10.00
Percentage of Speeding Duration [%]	3.00	0.00	0.00	65.00	7.00
Harsh accelerations [count]	0.67	0.00	0.00	21.00	1.36
Harsh brakings [count]	1.36	0.00	1.00	27.00	2.20

For further investigation, the differences between the two Phases were also explored with t-tests based on the drivers' trips during Phase 1 and Phase 2, respectively. A t-test analysis, based on the 65 selected drivers' trips, rejected the null hypothesis that both distributions of harsh events have an equal mean, demonstrating that the decrease in harsh events to be significant; both harsh accelerations (t-stat = 4.90, $p < 0.001$) and harsh brakings (t-stat = 4.82, $p < 0.001$). Additionally, results reveal that both the average speed (t-stat = 13.417, $p < 0.001$) and the maximum mean speed (t-stat = 10.25, $p < 0.001$) seem to be statistically significantly reduced during Phase 2 of the experiment. This also applies to the percentage of time driving over the speed limits per trip driving duration, taking into consideration the significant reduction of the particular variable, namely from 5% to 3% (t-stat = 17.63, $p < 0.001$). Finally, although the percentage of time using the mobile phone per trip driving duration is also found to be reduced in Phase 2 of the experiment, from 4% to 3%, the reduction is not statistically significant (t-stat = 1.15, $p = 0.248$).

4.2 Generalized Linear Mixed-Effects Models

In order to model the expected frequency of events per trip for the participant drivers, models in a GLM framework were calibrated, as previously explained. Since the BeSmart application allows for a high resolution, big-data oriented collection scheme, it was attempted to include random effects in order to capture the unique driving behavior traits for each driver. This entails having a critical minimum sample of trips for each driver to achieve a meaningful outcome. Therefore, a screening was made among participant drivers, as described above, and drivers that had over 40 trips each were selected for the GLMM analysis.

GLMMs were fitted in R-studio (with the lme4 package) via maximum likelihood and using z-factor scaling, following Bates et al. [51]. A number of models were tested with different configurations in the collected parameters in both fixed effects and random effects. The Poisson function with the log-odds link function was implemented. After conducting log-likelihood test (ANOVA) comparisons, the most informative configuration of random effects was the inclusion of both random intercepts and random slopes in the GLMMs to capture unique driver traits (lowest LogLikelihood and highest χ^2). It should be noted that conceptually, for harsh event frequencies, random slope models without random intercepts make little sense and are thus avoided.

Results of mixed effect selection are shown on Table 5 for harsh accelerations and on Table 6 for harsh brakings:

Table 5. Log-likelihood comparison of mixed effect selection for harsh acceleration models for both phases

Experiment Phase	Model Family	Model Configuration	D.f.	LogLikelihood	χ^2	P(>χ^2)	Sig.
Phase 1	GLM	Fixed effects only [baseline]	6	-11078.6	–	–	–
	GLMM	Fixed effects & Random Intercepts	7	-9929.5	2298.34	<2e-16	***
	GLMM	Fixed effects, Random Intercepts & Random Slopes	9	-9860.8	137.35	<2e-16	***
Phase 2	GLM	Fixed effects only [baseline]	6	-10639.5	–	–	–
	GLMM	Fixed effects & Random Intercepts	7	-9787.4	1704.32	<2e-16	***
	GLMM	Fixed effects, Random Intercepts & Random Slopes	9	-9741.7	91.34	<2e-16	***

Significance codes: '***': 0.000 \| '**': 0.001 \| '*': 0.01 \| '.': 0.05 \| ' ': ≥ 0.1

Table 6. Log-likelihood comparison of mixed effect selection for harsh braking models

Experiment Phase	Model Family	Model Configuration	D.f.	LogLikelihood	χ^2	P(>χ^2)	Sig.
Phase 1	GLM	Fixed effects only [baseline]	5	-15634	–	–	–
	GLMM	Fixed effects & Random Intercepts	6	-14352	2565.04	< 2e-16	***
	GLMM	Fixed effects, Random Intercepts & Random Slopes	8	-14158	388.04	< 2e-16	***
Phase 2	GLM	Fixed effects only [baseline]	5	-15588	–	–	–
	GLMM	Fixed effects & Random Intercepts	6	-14045	3086.73	<2e-16	***
	GLMM	Fixed effects, Random Intercepts & Random Slopes	8	-13965	159.73	<2e-16	***

Significance codes: '***': 0.000 \| '**': 0.001 \| '*': 0.01 \| '.': 0.05 \| ' ': ≥ 0.1

The final models were selected as the ones with the lowest AIC values. Fixed effect results appear on Table 7 for harsh acceleration frequencies and on Table 8 for harsh braking frequencies. A dash sign ('–') on the tables indicates that the specific variable was not used in the particular model. Furthermore, taking into account the value range of the examined data, the models do not yield any negative predictions, thus leading to no concerns for the negative value of the intercept which describes residual unexplained variance.

Table 7. GLMMs for harsh acceleration frequencies of 65 drivers (fixed effects)

Trip characteristic	GLMM for Phase 1					GLMM for Phase 2				
	Estimate	s.e.	p-value	Sig.	Relative Risk Ratio	Estimate	s.e.	p-value	Sig.	Relative Risk Ratio
Intercept	-0.927	0.091	0.000	***	0.395	-1.127	0.085	0.000	***	0.324
Maximum Speed	0.321	0.022	0.000	***	1.378	0.412	0.021	0.000	***	1.509
Percentage of Speeding Duration	0.074	0.013	0.000	***	1.076	0.035	0.012	0.003	**	1.035
Percentage of Mobile Use Duration	0.042	0.011	0.000	***	1.042	-	-	-	-	-
Log(Total Trip Duration)	0.848	0.051	0.000	***	2.334	0.729	0.050	0.000	***	2.073
Log(Total Trip Distance)	-0.231	0.050	0.000	***	0.793	-0.087	0.046	0.047	*	0.916

Significance codes: '***': 0.000 | '**': 0.001 | '*': 0.01 | '.': 0.05 | ' ': ≥ 0.1

Table 8. GLMMs for harsh braking frequencies of 65 drivers (fixed effects)

Trip characteristic	GLMM for Phase 1					GLMM for Phase 2				
	Estimate	s.e.	p-value	Sig.	Relative Risk Ratio	Estimate	s.e.	p-value	Sig.	Relative Risk Ratio
Intercept	-0.182	0.067	0.006	**	0.833	-0.313	0.075	0.000	***	0.731
Maximum Speed	0.327	0.016	0.000	***	1.387	0.331	0.015	0.000	***	1.395
Percentage of Speeding Duration	0.097	0.010	0.000	***	1.102	0.081	0.009	0.000	***	1.084
Log(Total Trip Duration)	0.885	0.045	0.000	***	2.423	0.723	0.038	0.000	***	2.061
Log(Total Trip Distance)	-0.298	0.036	0.000	*	0.742	-0.082	0.033	0.015	*	0.921

Significance codes: '***': 0.000 | '**': 0.001 | '*': 0.01 | '.': 0.05 | ' ': ≥ 0.1

Modelling results regarding the harsh acceleration frequencies reveal some interesting findings; the parameters of maximum speed, percentage of speeding duration and total trip duration have all been determined as statistically significant and positively correlated with harsh acceleration frequencies for both experiment phases. In the same context, total trip distance is statistically significant and negatively correlated with harsh acceleration frequencies for both experiment phases as well. Mobile use duration was found statistically significant only for Phase 1 with a small positive correlation.

More specifically, the aforementioned results could be further interpreted by calculating the relative risk ratio of every variable and thus measuring the increase in log-odds of the harsh acceleration frequencies. Maximum driver speed increases the frequencies of harsh accelerations; the effect appears to be higher in Phase 2 than Phase 1, as the respective estimates are 0.412 and 0.321, corresponding to risk ratios (in other words, incidence rate ratios for frequency frequencies) of 1.509 and 1.378. Exceeding the speed limit seems to increase the odds of harsh acceleration frequencies. A possible explanation is that certain drivers develop high driving speeds, reaching higher than the speed limit, they accelerate in a more abrupt way, allowing the application to detect a harsh acceleration event. In other words, overly aggressive drivers do not only exceed speed limits, they do so by accelerating harshly. When comparing the two phases, it is found that the effect of exceeding the speed limit is higher in Phase 1 than in Phase 2, with respective estimates of 0.074 and 0.035, corresponding to risk ratios of 1.076 and 1.035, respectively. Regarding the effect of mobile use while driving, it appears that 1% of mobile use duration per trip increases harsh acceleration event frequencies by 4% which makes it the least impactful variable among the examined ones, based on the respective estimate of 0.042. The fact that mobile phone use was not found significant for Phase 2 may be interpreted by the aforementioned reduction of the specific parameter during the two phases. Overall, it is noteworthy that all examined behavioral parameters (namely speed, percentage of speeding and mobile phone duration) are positively correlated with harsh acceleration frequencies, confirming the strong correlation between harsh events and unsafe driving behavior [16, 17].

As for the exposure parameters (trip distance and duration), there is a remarkable finding similar for both experiment phases. Trip duration seems to increase the odds of harsh acceleration frequencies; 1 sec of driving time increases acceleration frequencies by 2.3 and 2.1 times in Phase 1 and 2, respectively; risk ratios originating from parameter estimates of 0.848 and 0.729, respectively. However, the exposure parameter of total trip distance was found to be negatively associated with the odds of higher harsh acceleration counts, possibly because drivers may be prepared more effectively, physically and psychologically, when they are aware of the fact that they will cover a long distance. More precisely, a single unit of total trip distance travelled appears to reduce harsh acceleration counts by a lower degree in Phase 2 compared to Phase 1, as the respective estimates are -0.087 and -0.231, corresponding to risk ratios of 0.916 and 0.793.

With respect to harsh braking events (Table 8), apart from the mobile use variable, all the other variables, both driving behavioral and exposure ones, seem to have effects that are similar to the ones they have on the harsh acceleration events. The occurred finding is more obvious when it comes to both exposure parameters; trip distance and duration, where the effects appear to be identical. Specifically, trip duration appears to increase the odds of harsh braking frequencies; 1 sec of driving time increases braking frequencies by 2.4 and 2.1 times in Phase 1 and 2, respectively; risk ratios originating from parameter estimates of 0.885 and 0.723, respectively. On the other hand, similarly to the harsh acceleration models, the exposure parameter of total trip distance was found to be negatively associated with the odds of higher harsh braking counts; the effect seems higher in Phase 1 compare to Phase 2, as the respective estimates are -0.298 and -0.082, corresponding to risk ratios of 0.742 and 0.921.

The examined behavioral parameters, namely the parameters of maximum speed and percentage of speeding duration, have all been positively correlated with harsh braking frequencies for both experiment phases. Maximum driving speed increases the frequencies of harsh brakings; the effect appears to be similar for both Phases, as the respective estimates are 0.327 for Phase 1 and 0.331 for Phase 2, corresponding to risk ratios of 1.387 and 1.395, respectively. In the same context, exceeding the speed limit appears to increase the odds of harsh braking frequencies; the effect appears to be higher in Phase 1 than Phase 2, as the respective estimates are 0.097 and 0.081, corresponding to risk ratios of 1.102 and 1.084. As already mentioned in the harsh accelerations models, aggressive driving, expressed by higher driving speeds and higher percentage of speeding while driving, increases harsh braking frequencies, confirming once again the close relationship between harsh events and unsafe driving behavior. The absence of mobile phone use in the statistical model could be interpreted by the fact that drivers reduce speed while distracted, and therefore are less prone to harsh brakings. In other words, drivers decelerate in order to compensate for the distraction from mobile use, so they have the ability to brake normally and avoid harsh braking.

The visual representations of values of random intercepts and random slopes for the log of total trip duration per driver for both Phase 1 and Phase 2 are shown in the constructed caterpillar plots below for harsh acceleration frequencies (Figure 3) and for harsh braking frequencies (Figure 4), respectively. Personal differences per driver from the fixed effect intercept and slope are thus included in the linear predictor.

To visualise these results, driver-level deviations are shown relative to the mean (0.00; vertical gray line) with 95% confidence intervals around each intercept (blue circle). With respect to harsh acceleration frequencies, shown on Figure 3, the visual comparison of the intercept estimates illustrates that overall prediction averages in Phase 1 are similar to the ones of Phase 2. This is an indication that the unexplained variance for harsh acceleration occurrence is somewhat constant for each driver between the two Phases. Estimates for random slope effects of the log of total trip duration show a more erratic effect during Phase 1 comparted to Phase 2. In Phase 1, the impact of total duration on harsh acceleration counts appears to be mathematically higher for certain drivers. In practice, this indicates that there are certain individuals for whom total trip duration plays very different roles for harsh acceleration occurrence when they are not provided with feedback, a trend that greatly diminishes when feedback is provided.

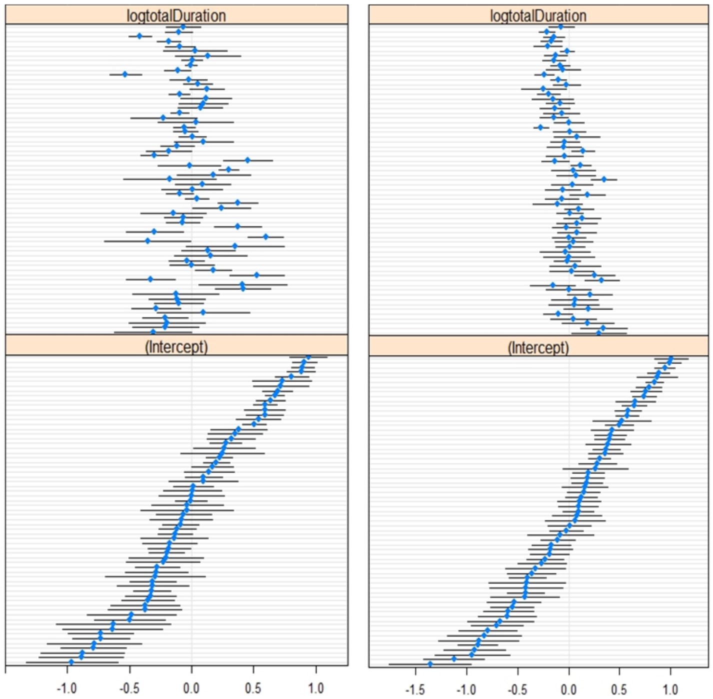

Figure 3: Random Intercepts and Random Slopes for log(total trip duration)
[Both in the GLMM for harsh acceleration frequencies during Phase 1 (left) and Phase 2 (right)]

However, different observations are drawn from the caterpillar plots when examining harsh braking frequencies, shown on Figure 4. The visual comparison of intercept estimates illustrates that prediction averages in Phase 2 overlap by a larger degree than in Phase 1. Furthermore, the gradient when connecting the blue data points of intercepts is steeper for Phase 2 if one takes the reduced range of the x-axis into account. Estimates for random slope effects of the log of total trip duration show little difference across the two Phases. It seems that the provision of feedback in Phase 2 led to overall reductions of harsh braking events per trip duration unit if the driver sample is examined on an aggregate level. Nonetheless, these reductions seem to have a certain amount of divergence across individual drivers.

As previously stated, the personal differences are captured in the random effects of the GLMMs. These differences may be parameters unobserved in the present models such as driver age, experience, aggressiveness, alertness and performance levels and other similar human factors. Since the dependent variables are harsh event frequencies, which also depend on the road environment, additional parameters which may affect harsh event occurrence can be thought to be integrated therein as well. Examples include temporal and spatial headways and more unforeseen events such as traffic conflicts or the presence of obstacles.

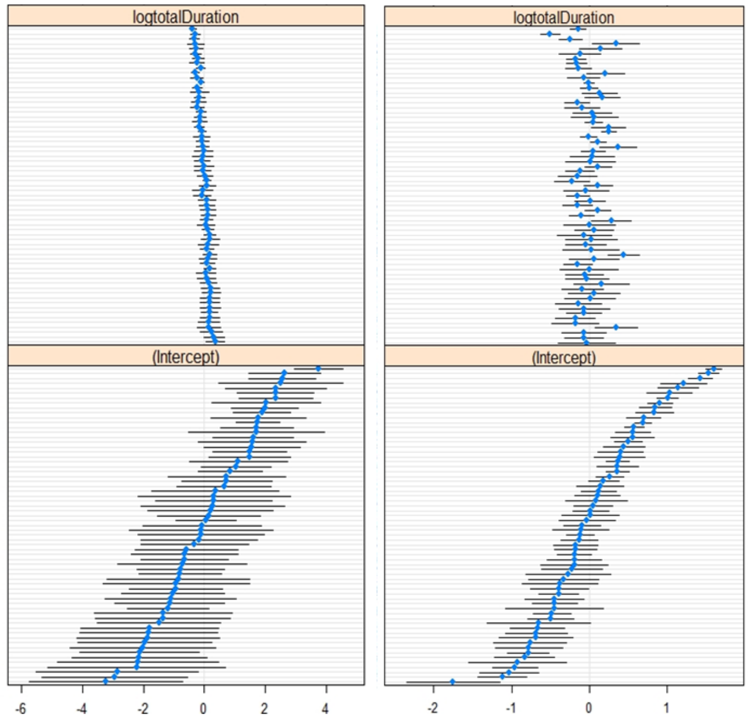

Figure 4: Random Intercepts and Random Slopes for log(total trip duration)
[Both in the GLMM for harsh braking frequencies during Phase 1 (left) and Phase 2 (right)]

5. Conclusions, Limitations and Future Research

This paper aimed to investigate the impact of detailed trip characteristics on the frequency of harsh acceleration and harsh braking events recorded by smartphone sensors. In order to achieve that objective, an ongoing naturalistic driving experiment was carried out within the framework of the BeSmart project in order to examine driving behavior as expressed by the frequencies of harsh accelerations and harsh brakings. In the present research, the first two phases were considered; Phase 1, where participants were asked to drive in the way they usually did, without receiving any feedback on their driving behavior and Phase 2, where participants were provided with personalized feedback, a trip list and a scorecard regarding their driving behavior, allowing them to identify their critical deficits or unsafe behaviors.

Generalized Linear Mixed-Effects Models were fitted to the trips of 65 car drivers who made frequent trips during both experiment phases in order to model the frequencies of harsh acceleration and harsh braking events. Results reveal correlations of harsh event frequencies with specific driving behavior and exposure metrics, at a more detailed level than existing studies (e.g. exceeding speed limit, trip duration). More specifically, results indicate that maximum speed, the percentage of speeding duration and total trip duration are positively correlated with both harsh acceleration and harsh braking frequencies. On the other hand, the exposure metric of total trip distance was found to be negatively correlated with both harsh event types.

Results for Phase 1 and Phase 2 indicate that both types of harsh events are influenced by the same explanatory variables, with the exception of mobile use while driving, which is found statistically significant only for harsh acceleration models for Phase 1 of the experiment. Additionally, for the majority of the variables, coefficient values seem to change between the two experiment phases in a similar direction for both harsh acceleration and harsh braking events. Specifically, the effect of speeding was found to be higher for the frequency of harsh events in Phase 1 than Phase 2. On the contrary, the effect of maximum drive speed is higher in Phase 2 than Phase 1 for both acceleration and braking events. A similar pattern is noticed for the exposure parameters, as it seems that trip duration effect is higher in Phase 1 than Phase 2, while trip distance effect was found higher in Phase 2 than in Phase 1, for both types of events.

Although initial findings from descriptive statistics suggest that drivers improved their driving behavior with regards to all the recorded driving behavior metrics, namely maximum driver speed, percentage of speeding duration and using the mobile phone duration, the application of suitable statistical methodologies for before-after evaluation leads to more reliable and accurate results. As the experiment progresses, different types of personalised feedback will be communicated to all drivers allowing them to identify their critical deficits or unsafe behaviors, while incentives within a social gamification scheme, with personalised target setting, benchmarking and comparison with peers will also be developed and provided through the smartphone application.

Alternative approaches to frequency modelling, such as the modelling of harsh event rates (harsh events per km), were considered. However, due to the nature of the data, they were ultimately discarded. Linear modelling led to negative predictions and very poor independent variable and overall model fits. Additionally, trips with zero harsh events in the dataset did not allow for log-normal transformations, and their subsequent removal would bias results. This paper does not explicitly focus on the evaluation of the effectiveness of driver feedback on improving driving behavior and increasing road safety levels from the smartphone application; this will be the dedicated focus of future research. It should be mentioned that this analysis is macroscopic overall, and should be treated as a high-level behavioral investigation. Within the present approach, there is no option to statistically examine if mobile phone use was exactly simultaneous with harsh events, but only to verify that they both occurred within the same trip. In other words, the temporal coincidence of data was not considered. Addressing these limitations will require the development of additional dedicated methodologies in the future, and the examination of in-depth datasets analyzing each trip per trip-second.

Another aspect that is worth investigating is the combined impact of traffic characteristics, road infrastructure and trip characteristics on the frequency of harsh events. Recent research suggest that there are strong correlations between road and traffic characteristics and harsh events [52], therefore the investigation of harsh events using multi-source data is a promising future research direction. Unfortunately, due to GDPR restrictions, the precise location of the trip data in the present research was not allowed, resulting in the lack of information regarding the road network. Nevertheless, the proposed methodology reaches noteworthy findings, even though the driving exposure and behavior metrics are solely based on smartphone sensors and no other methods, e.g. traffic video data. In this context, it is important to further explore the new possibilities for data collection and processing using such a portable and low-cost device in order to identify new ways of tackling and mitigating the prominent risk factors and improve road safety.

Finally, future research will also focus on the analysis of different driving behavior parameters identified by the road safety literature as risk factors (e.g. exceeding speed limit, mobile phone distraction) and their effect on driving performance and road safety. Furthermore, analyses per gender, age, crash history, self–assessment, driving experience and more demographic characteristics could be undertaken in order to capture any particular trends found in the categories of these parameters, possibly improving feedback processes, on the condition that this information can be provided while observing data protection laws.

6. Ethics approval

This study was conducted with the approval of the National Technical University of Athens – NTUA Ethics Committee

7. Acknowledgements

This research is co-financed by the European Union - European Regional Development Fund (ERDF) and Greek national funds through the Operational Program "Competitiveness, Entrepreneurship and Innovation" (EPAnEK) of the National Strategic Reference Framework (NSRF) - Research Funding Program: BeSmart - Multi-modal driver behavior and safety support system on the basis of smartphone applications.

8. References

[1] Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., & Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Human Factors*, *40*(1), 136-149.
[2] Young, M. S., Birrell, S. A., & Stanton, N. A. (2011). Safe driving in a green world: A review of driver performance benchmarks and technologies to support 'smart'driving. *Applied ergonomics*, *42*(4), 533-539.
[3] Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, *113*(10), 2636-2641.
[4] World Health Organization. Global Satus Report on Road Safety. 2018.
[5] Salmon, P., Young, K., Lenné, M., Williamson, A., Tomasevic, N., 2011. The Nature of Errors made by Drivers. Austroads Publication No. AP–R378/11. Austroads Ltd., Aus-tralia.
[6] Singh, S. (2018). Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. (Traffic Safety Facts Crash Stats. Report No. DOT HS 812 506). Washington, DC: National Highway Traffic Safety Administration.
[7] Papantoniou, P., Papadimitriou, E., Yannis, G. (2015). Assessment of driving simulator studies on driver distraction. Advances in Transportation Studies, 35, 129-144.
[8] Stutts, J.C., Reinfurt D.W., Staplin L., Rodgman E.A. (2001). The role of driver distrac-tion in traffic crashes. Report Prepared for AAA Foundation for Traffic Safety.
[9] Bellinger, D.B., Budde, B.M., Machida, M., Richardson, G.B., Berg, W.P. (2009). The effect of cellular telephone conversation and music listening on response time in braking. Transportation Research Part F 12 (6), pp. 441-451.
[10] Fitch, G. M., Bartholomew, P. R., Hanowski, R. J., & Perez, M. A. (2015). Drivers' visual behavior when using handheld and hands-free cell phones. Journal of safety research, 54, 105-e29.
[11] Elvik, R., Christensen, P., & Amundsen, A. (2004). Speed and road accidents. An evaluation of the Power Model. TØI report, 740, 2004
[12] Nilsson, G. (2004). Traffic safety dimensions and the power model to describe the effect of speed on safety (Doctoral dissertation, Univ.).
[13] ITF, 2018. Speed and Crash Risk. OECD/ITF, Paris.
[14] Elvik, R., Vadeby, A., Hels, T., & Van Schagen, I. (2019). Updated estimates of the relationship between speed and road safety at the aggregate and individual levels. Accident Analysis & Prevention, 123, 114-122.
[15] Bonsall, P., Liu, R., & Young, W. (2005). Modelling safety-related driving behavior—impact of parameter values. *Transportation Research Part A: Policy and Practice*, *39*(5), 425-444.
[16] Vaiana, R., T. Iuele, V. Astarita, M. V. Caruso, A. Tassitani, C. Zaffino, and V. P. Giofrè. Driving Behavior and Traffic Safety: An Acceleration-Based Safety Evaluation Procedure for Smartphones. Modern Applied Science. 2014 Feb 1;8(1):88.
[17] Gündüz, G., Yaman, Ç., Peker, A. U., & Acarman, T. (2018). Prediction of Risk Generated by Different Driving Patterns and Their Conflict Redistribution. IEEE Transactions on Intelligent Vehicles, 3(1), 71-80.
[18] Tarko, A. P. (2018). Surrogate measures of safety. In Safe mobility: challenges, methodology and solutions. Emerald Publishing Limited.
[19] Johnsson, C., Laureshyn, A., & De Ceunynck, T. (2018). In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. Transport Reviews, 38(6), 765-785.
[20] Stephens, A. N., & Groeger, J. A. (2009). Situational specificity of trait influences on drivers' evaluations and driving behavior. Transportation research part F: traffic psychology and behavior, 12(1), 29-39.
[21] Roidl, E., Frehse, B., & Höger, R. (2014). Emotional states of drivers and the impact on speed, acceleration and traffic violations A simulator study. Accident Analysis & Prevention, 70, 282-292.
[22] Hanowski, R. J., Perez, M. A., & Dingus, T. A. (2005). Driver distraction in long-haul truck drivers. Transportation Research Part F: Traffic Psychology and Behavior, 8(6), 441-458.
[23] Jansen, R. J., & Wesseling, (2018). SWOV. Harsh Braking by Truck Drivers: A Comparison of Thresholds and Driving Contexts Using Naturalistic Driving Data. In Proceedings of the 6th Humanist Conference.
[24] Boquete, L., Rodríguez-Ascariz, J. M., Barea, R., Cantos, J., Miguel-Jiménez, J. M., & Ortega, S. (2010). Data acquisition, analysis and transmission platform for a pay-as-you-drive system. Sensors, 10(6), 5395-5408.
[25] Paefgen, J., Staake, T., & Thiesse, F. (2013). Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach. Decision Support Systems, 56, 192-201.
[26] Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2017). Innovative motor insurance schemes: A review of current practices and emerging challenges. *Accident Analysis & Prevention*, *98*, 139-148.
[27] Ziakopoulos, A., Tselentis, D., Kontaxi, A., & Yannis, G. (2020a). A critical overview of driver recording tools. *Journal of Safety Research*.
[28] Ghose, A., Chowdhury, A., Chandel, V., Banerjee, T., & Chakravarty, T. (2016, January). An enhanced automated system for evaluating harsh driving using smartphone sensors. In *Proceedings of the 17th International Conference on Distributed Computing and Networking* (pp. 1-6).
[29] Tselentis, D. I., Vlahogianni, E. I., & Yannis, G. (2018). Comparative Evaluation of Driving Efficiency Using Smartphone Data. *METHODOLOGY*, *32*, 33.
[30] Vlahogianni, E. I., & Barmpounakis, E. N. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges. *Transportation Research Part C: Emerging Technologies*, *79*, 196-206.

[31] Paefgen, J., Kehr, F., Zhai, Y., & Michahelles, F. (2012, December). Driving behavior analysis with smartphones: insights from a controlled field study. In *Proceedings of the 11th International Conference on mobile and ubiquitous multimedia* (pp. 1-8).
[32] Yannis, G., D. Tselentis, E. Vlahogianni, A. Argyropoulou. Monitoring distraction through smartphone naturalistic driving experiment. Proceedings of the 6th International Naturalistic Driving Research Symposium (NDRS), The Hague, 8-9 June 2017
[33] Papadimitriou, E., Argyropoulou, A., Tselentis, D. I., Yannis, G. (2019). Analysis of driver behavior through smartphone data: The case of mobile phone use while driving. Safety Science, 119, 91-97.
[34] Meseguer, J. E., Calafate, C. T., Cano, J. C., & Manzoni, P. (2013, July). Drivingstyles: A smartphone application to assess driver behavior. In *2013 IEEE Symposium on Computers and Communications (ISCC)* (pp. 000535-000540). IEEE.
[35] Ma, C., Dai, X., Zhu, J., Liu, N., Sun, H., & Liu, M. (2017). Drivingsense: Dangerous driving behavior identification based on smartphone autocalibration. *Mobile Information Systems*, *2017*.
[36] Mantouka, E. G., Barmpounakis, E. N., & Vlahogianni, E. I. (2018). Mobile sensing and machine learning for identifying driving safety profiles (No. 18-01416)
[37] Soriguera F., Miralles E. (2016). Driver feedback mobile app. Transportation Research Procedia, Volume 18, Pages 264-271.
[38] Dahlinger, A., Tiefenbeck, V., Ryder, B., Gahr, B., Fleisch, E., & Wortmann, F. (2018). The impact of numerical vs. symbolic eco-driving feedback on fuel consumption–A randomized control field trial. Transportation Research Part D: Transport and Environment, 65, 375-386.
[39] Michelaraki, E., Kontaxi, A., Papantoniou, P., Yannis, G. (2020) Correlation of driver behavior and fuel consumption using data from smartphones. Proceedings of 8th Transport Research Arena TRA (Conference cancelled)
[40] Tselentis, D., Kostoulas, G., Kontonasios, K.-N., Fortsakis, P., & Yannis, G. (2020) Driver feedback during naturalistic driving experiments: A review of types, methods and future challenges. Proceedings of 8th Transport Research Arena TRA (Conference cancelled)
[41] Toledo, G., & Shiftan, Y. (2016). Can feedback from in-vehicle data recorders improve driver behavior and reduce fuel consumption?. *Transportation Research Part A: Policy and Practice*, *94*, 194-204.
[42] Ziakopoulos, A., Kontaxi, A., Yannis, G., Fortsakis, P., Kontonasios, K.-N. & Kostoulas, G. (2020b) Advanced driver monitoring using smartphone applications: The BeSmart project. Proceedings of 8th Transport Research Arena TRA (Conference cancelled).
[43] Vlahogianni, E.I. & Barmpounakis, E.N. (2016). How accurate are smartphones for monitoring Driver's Behavior?, TRB 95th Annual Meeting January 10-14, Washington, D.C., US.
[44] Tselentis, D. I., Vlahogianni, E. I., & Yannis, G. (2019). Driving safety efficiency benchmarking using smartphone data. *Transportation research part C: emerging technologies*, *109*, 343-357.
[45] Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation research part A: policy and practice, 44(5), 291-305.
[46] Twisk, D., Stelling, A., Van Gent, P., De Groot, J., & Vlakveld, W. (2020). Speed characteristics of speed pedelecs, pedelecs and conventional bicycles in naturalistic urban and rural traffic conditions. Accident Analysis & Prevention, 150, 105940.
[47] Kamla, J., Parry, T., & Dawson, A. (2019). Analysing truck harsh braking incidents to study roundabout accident risk. Accident Analysis & Prevention, 122, 365-377.
[48] Winter, B. (2013). A very basic tutorial for performing linear mixed effects analyses. arXiv preprint arXiv:1308.5499.
[49] Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American statistical Association*, *88*(421), 9-25.
[50] McCulloch, C. E. (2003). Generalized linear mixed models. In NSF-CBMS regional conference series in probability and statistics (pp. i-84). Institute of Mathematical Statistics and the American Statistical Association.
[51] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(7), 1-23.
[52] Petraki V., Ziakopoulos A., Yannis G., (2020) "Combined impact of road and traffic characteristic on driver behavior using smartphone sensor data". Accident Analysis & Prevention, Vol. 144.