Examining the Impact of Feedback on Traffic and Safety Behavior of Car Drivers in a Naturalistic Driving Study

Armira Kontaxi
Ph.D. Candidate, Research Associate
Department of Transportation Planning and Engineering
National Technical University of Athens, 5 Heroon Polytechniou Str., Athens, Greece, GR-15773
Email: akontaxi@mail.ntua.gr

Apostolos Ziakopoulos
Postdoctoral Researcher
Department of Transportation Planning and Engineering
National Technical University of Athens, 5 Heroon Polytechniou Str., Athens, Greece, GR-15773
Email: apziak@central.ntua.gr

George Yannis
Professor
Department of Transportation Planning and Engineering
National Technical University of Athens, 5 Heroon Polytechniou Str., Athens, Greece, GR-15773
Email: geyannis@central.ntua.gr

Word Count: 5973 words + 3 tables (250 words per table) = 6723 words

Submitted [01 Aug 2023]

ABSTRACT

Rapid technological advances, especially in telematics and Big Data analytics, as well as the increasing penetration and use of information technology by drivers (e.g. smartphones), provide new capabilities for monitoring and analyzing driving behavior. This paper examines the impact of driver feedback delivered through a smartphone application on driving behavior risk indicators, within a 21-month multiphase naturalistic driving experiment involving a sample of 175 car drivers. First, a preliminary analysis utilizing summary statistics and Wilcoxon signed-rank test shed light upon the effects of upgraded feedback features on key risk indicators across experiment phases. Subsequently, Structural Equation Models (SEM) on a 73,869 trip dataset provided significant insights into how feedback mechanisms and exposure factors influence driving behaviors. Results indicate that the examined feedback mechanisms are effective in reducing the percentage of speeding time and harsh braking events, although there is a slight increase in harsh accelerations, which may require further refinement of the feedback system. The scorecard feature has the highest positive effect, indicating its crucial role in modifying driving habits, with gamification (competition and challenges) being the second most influential feedback mechanism. Regarding the exposure indicators, morning peak is associated with more aggressive driving, while afternoon peak tends to be less risky. Additionally, results showcase strong positive correlations between speeding, harsh braking, and harsh accelerations highlight the interconnected nature of aggressive driving behaviors. These findings may be beneficial for insurance companies, fleet management applications, and policymakers, enabling them to leverage results to improve traffic safety and driver behavior.

Keywords: Driver Feedback, Road Safety, Naturalistic Driving, Multiphase Experiment, Wilcoxon Signed-Rank Test, Structural Equation Models

INTRODUCTION

Despite considerable progress in road safety over the past decade, road traffic crashes remain a pervasive public health issue globally, resulting in around 1.19 million road traffic deaths in 2021 (*1*), corresponding to a rate of 15 road traffic deaths per 100 000 population. The identification of critical risk factors leading to road traffic crashes has been researched by numerous studies over the years. Among these factors, human elements are consistently recognized as the most significant, accounting for the vast majority of road crashes. In fact, human error is cited as the cause of 95% of all road crashes (2). This underscores the importance of understanding and addressing driver behavior as a key component of road safety initiatives. By analyzing driver behavior, targeted interventions can be developed, aiming to mitigate risky actions such as distracted driving, speeding, and impaired driving.

The significance of driver monitoring is becoming more widely acknowledged in the transportation sector (3). Nevertheless, researchers encounter difficulties in collecting accurate real-time driving data with affordable collection and processing techniques. In this context, the widespread use of smartphones and social networks presents new opportunities for monitoring and analyzing driver behavior (4). The capabilities of smartphone applications, combined with their low cost and ease of use, facilitate data collection. These advancements facilitate the provision of direct feedback and trip analysis to drivers, potentially reducing road crashes and casualties. Going one step further, the conduction of driving experiments under naturalistic conditions using smartphones allows for the recording of drivers in their normal driving environments without external influences, and thus for the effective assessment of driver behavior (5). Despite the growing interest from both manufacturing companies and transportation researchers in driver behavior, there is a notable gap in research quantifying the influence of driver feedback on road safety, particularly in terms of comparing data before and after feedback provision.

In this regard, the present study aims to leverage large-scale trip data from smartphone sensors to assess the impact of driver feedback on key performance indicators, such as speeding, harsh braking, and harsh acceleration events. For this purpose, a naturalistic driving experiment has been conducted, thousands of trips have been used first to examine the trend of the risk driving indicators and then, Structural Equations Models (SEM) are applied to identify feedback effects to risky driving indicators. The outputs of the two methods are combined to provide some critical insights on whether driver feedback influences driving behavior and in what extent.

The remainder of the paper is organized as follows: first, the main findings of previous works are discussed, then, the methodology of the study is presented, including the experiment design, the data collected and the methodological tools that are used. Subsequently, the results of the analysis are discussed and finally, conclusions and suggestions for future research are drawn.

LITERATURE REVIEW

Numerous studies have focused on driving behavior and naturalistic observations, primarily examining behavior recording and subsequently analyzing and modeling driver profiles (6, 7). These studies also investigate unsafe behaviors such as speeding (8, 9), mobile phone use (10, 11), harsh driving events and driver aggressiveness (12, 13) and driver fatigue (14, 15). Additionally, researchers have developed technologies and machine learning algorithms to detect these behaviors (16–18) and technologies that provide feedback to drivers (*19*, *20*).

Feedback to drivers has been shown to be a highly effective method for enhancing road safety. Feedback itself has long been acknowledged as a powerful tool for shaping behavior in diverse areas, including education, healthcare, and human resource management (*21*–*23*). In traffic and road safety, the importance of feedback can be highlighted through several key points such as behavior modification, enhanced awareness, reduction in crash rates, stress and fatigue management, integration with advanced technologies and promotion of a road safety culture.

Many studies have examined the effect of feedback, however there is very little research that quantify the exact effect on driver behavior and safety, as in many cases the drivers were recorded after the feedback system/mechanism had been applied, without monitoring a baseline period. A naturalistic driving experiment (*24*) was conducted for 57 car drivers with a control and intervention group for 11 weeks through a smartphone application and the drivers received a text message after the completion of the trip with personalized feedback about the participant's risky driving behavior. Four separate Generalized Estimating Equations (GEE) linear regression models were developed for each driving indicators and the results showed that the treatment effects for feedback were consistently in the expected positive direction. Another recent study (*25*) conducted a 16-week ND experiment including 3 phases (i.e. baseline, different types of feedback, follow-up without feedback) and provided real-time and post-drive feedback to drivers. Results showed that real-time feedback alone and in conjunction with financial incentives were effective in raising speed limit compliance. It is also interesting to note that the effects did not sustain when feedback and incentives were removed. The post feedback effect is an aspect that should be further investigated as the few studies that have dealt with the matter have not come to conclusive results (*26*, *27*), while some showing both positive (*28*) and negative (29) effects.

The methods used in studies examining the effect of driver feedback vary widely. After establishing the context and research questions, methodologies employed in these studies, are also important to be discussed. Many studies (30–32) initially focus on basic correlation tests, presenting critical summary statistics that compare the feedback and non-feedback phases or groups. These basic statistical comparisons serve as a foundation for understanding the immediate effects of driver feedback.

However, relying solely on basic correlation tests can be limiting, as these methods do not account for the complexity and multifaceted nature of driving behavior. As a result, several studies employ these basic methods as a preliminary step before moving on to more advanced statistical modeling. For instance, (33, 34) use initial correlation analyses as a data exploration phase. This phase is crucial for identifying key patterns and relationships, which then inform the development of sophisticated statistical models that can better capture the nuances of driver behavior and the impact of feedback.

Advanced methodologies, such as multivariate regression models, machine learning algorithms, and time-series analyses, are increasingly being used to understand the effects of driver feedback in a more comprehensive manner (35, 36) . These methods allow researchers to control for various confounding factors, explore interactions between multiple variables, and predict outcomes based on complex data patterns. For example, the use of machine learning techniques like supervised learning algorithms (e.g., XGBoost) has proven effective in modeling the contributions of different driving features to the decision to engage in risky behaviors, such as mobile phone use while driving (11).

METHODOLOGY

Experiment Design

As part of a research project, a 21-month naturalistic driving experiment involving 230 participants was conducted from July 2019 to March 2021, counting for 106,776 trips. This study included different types of drivers, namely car drivers, professional van drivers, and motorcycle riders. This paper focuses on car drivers, who correspondent the majority of the study's participants.

The main objectives of this experiment were to identify critical risk factors through driver monitoring using an innovative smartphone application and to develop feedback features that inform, notify, and motivate drivers to improve their skills, reduce driving errors, and lower crash risk. The experiment was divided into six phases, each providing different types of feedback to drivers. Figure 1 displays the different feedback features provided at each phase. As it is shown, Phase 1 served as the baseline phase where drivers were recorded through the smartphone app and only a trip list was available with no other information about the driver behavior. In phases 2,3,4 and 5 different feedback features were added, as shown below, while drivers returned to no feedback in phase 6 for researchers to examine the post feedback effect.

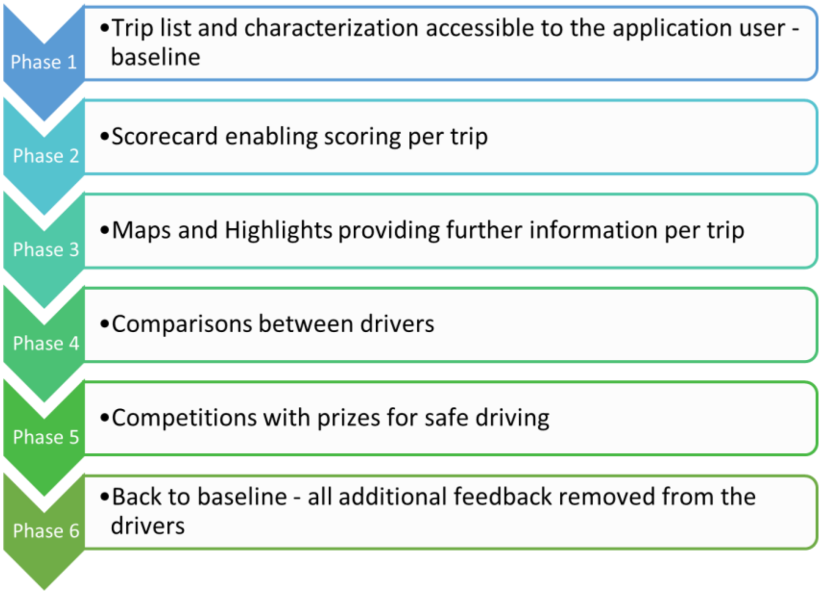

Figure 1 Experiment phases

Smartphone App and Data Collection

For that purpose, an innovative smartphone application, developed by OSeven (www.oseven.io), was utilized to assess and improve driver behavior and safety. This application records driver behavior using the smartphone's hardware sensors and various APIs to collect sensor data, which is temporarily stored on the device before being transmitted to a central database. The collected data is highly detailed in terms of time and location. Once uploaded to the backend cloud server, the data is processed into meaningful driving behavior and safety indicators using advanced signal processing, Machine Learning (ML) algorithms, data fusion, and Big Data techniques, all in compliance with Greek and European GDPR regulations. Figure 2 dipicts and summarizes the OSeven data flow system.

Figure 2 The OSeven data flow system

A variety of different metadata are eventually calculated, including both exposure indicators, namely trip distance, driving duration, road type, rush hours etc., and driving behavior indicators, namely speeding (duration of speeding, speed limit exceedance etc.), number and severity of harsh events, harsh braking, harsh acceleration, and distraction from mobile phone use (mobile phone use is considered any type of phone use by the driver e.g. talking, texting etc.). The reader can refer to previous studies in the framework of the same research project for further details in data processing and metadata calculation (33, 37). Figure 3 presents example screenshots from the application features in all experiment phases.

Figure 3 Example screenshots from the application features in all experiment phases

Theoretical background

The present analysis aims to examine the impact of feedback on driver behavior, i.e. in which ways driving risk factors are influenced by driver feedback.

Structural Equation Modeling (SEM) is a technique within the family of latent variable analysis. It is a multivariate method that supports both multiple-input and multiple-output modeling. In this study, SEM is used to formulate several unobserved constructs as latent variables from different types of variables collected through the naturalistic driving experiment. SEM is a widely recognized methodology with numerous applications. It has been employed in various studies to model complex interrelationships involving unobserved concepts expressed as latent variables. This includes applications in traffic engineering and road safety. For example, SEM has been used to model driving behavior and the probability of crash risk (38, 39) Additional examples include the use of SEM to connect task complexity and coping capacity with driving risk (40) or perception of risk and driving tasks on road safety attitudes of drivers (41, 42).

The underlying mathematical structure of SEMs can be defined as follows (43):

$$\eta = \beta\eta + \gamma\xi + \varepsilon \qquad (1)$$

where η is a vector expressing the dependent variables, ξ expressing the independent variables, ε expressing the regression error term, β expressing the regression coefficients for the dependent variables, and γ expressing the regression coefficients for the independent variables.

Path analysis is frequently used to represent the structural model, showing how a group of "explanatory" variables can affect a "dependent" variable. This method visually illustrates the relationships, indicating whether the explanatory variables act as correlated causes, mediated causes, or independent causes of the dependent variable.

In the realm of model selection, Goodness-of-Fit measures play a crucial role in any statistical model assessment. The goodness-of-fit metrics utilized in the current analysis are listed below.

The Comparative Fit Index (CFI) compares the fit of a hypothesized model with an independence model. Values range from 0 to 1, with over 0.90 generally accepted as a good fit. The formula is represented as follows:

$$\text{CFI} = 1 - \frac{\max\left(x_H^2 - df_H, 0\right)}{\max\left(x_H^2 - df_H, x_I^2 - df_l\right)} \qquad (2)$$

where x_H^2 and $df_H$are the chi-square value and degrees of freedom of the hypothesized model, and x_I^2 and df_l are those of the independence model.

The Tucker–Lewis Index (TLI) evaluates model parsimony, with values above 0.95 indicating a good fit. The formula is represented as follows:

$$TLI = \frac{\frac{x_I^2}{df_I} - \frac{x_H^2}{df_H}}{\frac{x_I^2}{df_l} - 1} \qquad (3)$$

The Root Mean Square Error Approximation (RMSEA) measures the unstandardized discrepancy between the population and the fitted model, with values below 0.08 typically considered a good fit and values below 0.05 indicating an excellent fit. The formula is represented as follows:

$$RMSEA = \sqrt{\frac{x_H^2 - df_H}{df_H(n-1)}} \qquad (4)$$

where x_H^2 is the chi-square value, df_H is the degrees of freedom, and nnn is the sample size.

The Standardized Root Mean Square Residual (SRMR) is the square root of the difference between the residuals of the sample covariance matrix and the hypothesized covariance model, with values below 0.05 indicating an excellent fit. The formula is represented as follows:

$$SRMR = \sqrt{\frac{\sum_{i=1j}^{n} \sum_{j=1}^{n} \left(s_{ij} - \sigma_{ij}\right)^2}{n(n+1)/2}} \quad (5)$$

where s_{ij} is the observed covariance between variables i and j, σ_{ij} is the predicted covariance between variables i and j based on the model and n is the number of observed variables.

RESULTS AND DISCUSSION

Descriptive Statistics

Overall, during the 21-months experiment 73,869 trips were recorded from a sample of 175 car drivers (54% female, all ages) who had participated in all experiment phases. This subchapter presents the trend in the most crucial indicators of driving behavior during the experiment, noting the change in the phase of the experiment, in order to draw some initial conclusions about the effect of feedback on driver behavior. Table 1 shows the summary statistics of the selected variables, while Figure 4 illustrate the trend of the same variables across the different experiment phases. The trends identified are the following: (i) There was an overall improvement in driving behavior from Phase 1 to Phase 2; (ii) The Covid-19 pandemic and subsequent lockdown measures significantly reduced the travel; (iii) There is a fluctuating improvement in driver behavior in subsequent phases; (iv) There is an improvement in driver behavior during the Competition phase; (v) There is a relapse to worse driving behavior once the Competition and Challenge phase is completed.

Table 1. Descriptive statistics of the per trip values of the variables recorded during the experiment

Experiment Phases	**Percentage of mobile use**		**Percentage of speeding time**		**Harsh brakings per 100km**		**Harsh accelerations per 100km**		**Speed above the speed limits**	
	mean	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**
Phase1	4.69%	0.13	6.17%	0.10	18.90	31.23	9.20	21.34	4.81	6.37
Phase 2	3.95%	0.12	3.71%	0.07	19.35	32.72	9.64	22.37	4.22	6.30
Phase 3	4.55%	0.13	3.70%	0.08	19.75	31.87	11.34	25.11	3.71	5.76
Phase 4	4.44%	0.13	3.88%	0.08	17.70	31.62	10.53	23.98	3.33	5.28
Phase 5/ Competition	3.03%	0.10	2.69%	0.06	13.21	24.92	8.23	20.44	2.38	4.13
Phase 5/ Challenges	3.04%	0.11	3.37%	0.07	16.21	29.26	8.62	21.86	2.51	4.34
Phase 6	2.35%	0.09	3.72%	0.07	16.89	29.79	8.01	19.72	3.01	4.51

Figure 4 Trends of the driving behavior parameters over the different experimental phases

Before moving to the advanced statistical analysis, Wilcoxon signed-rank test was used to compare the examined variable of the different phases and assess whether their population mean ranks differ. Since none of the variables met the assumptions for normality and homogeneity of variances, the Wilcoxon signed-rank test (a non-parametric test) was used for all variables. Previous within subjects design studies in the field of naturalistic driving research have also utilized the test to compare means among different feedback phases (44, 45).

The Wilcoxon signed-rank test results, showed in Table 2, reveal significant changes in driving behaviors across different phases. Notable improvements were observed between Phase 1 and Phase 2, with substantial reductions in mobile use (26.20%), speeding time (41.40%), and harsh braking (12.90%), indicating the effectiveness of feedback interventions. Between Phase 2 and Phase 3, although mobile use and speeding time continued to decrease, harsh braking slightly increased, suggesting mixed outcomes. The comparison between Phase 3 and Phase 4 showed an increase in mobile use (9.60%) but a decrease in harsh accelerations (11.20%), highlighting persistent distractions despite some improvements. Lastly, the shift from Competition and Challenges to the last phase resulted in significant increases in all risky behaviors, emphasizing that drivers may relapse to unsafe behaviors once not receiving anymore feedback.

Table 2. Wilcoxon signed-rank test results comparing means of risky driving behavior parameters across each phase

Phases compared	**Driving behavior parameters**	**mean diff.**	**S statistic**	**p-value**
Phase 1 - Phase 2	Percentage of mobile use	-26.20%	1.50591E+11	<**0.01**
Phase 1 - Phase 2	Percentage of speeding time	-41.40%	4.63629E+11	<**0.01**
Phase 1 - Phase 2	Harsh brakings per 100km	-12.90%	4.05868E+11	<**0.01**
Phase 1 - Phase 2	Harsh accelerations per 100km	-2.50%	2.07426E+11	<**0.01**
Phase 1 - Phase 2	Speed above the speed limits (km/h)	-18.00%	4.14412E+11	<**0.01**
Phase 2 - Phase 3	Percentage of mobile use	-26.80%	2.74877E+11	<**0.01**
Phase 2 - Phase 3	Percentage of speeding time	-16.70%	5.50967E+11	<**0.01**
Phase 2 - Phase 3	Harsh brakings per 100km	1.50%	7.48152E+11	<**0.01**
Phase 2 - Phase 3	Harsh accelerations per 100km	0.00%	4.86203E+11	<**0.01**
Phase 2 - Phase 3	Speed above the speed limits (km/h)	-27.60%	5.1229E+11	<**0.01**
Phase 3 - Phase 4	Percentage of mobile use	9.60%	1.08069E+12	<**0.01**
Phase 3 - Phase 4	Percentage of speeding time	-5.80%	2.46395E+12	<**0.01**
Phase 3 - Phase 4	Harsh brakings per 100km	-10.00%	2.68961E+12	<**0.01**
Phase 3 - Phase 4	Harsh accelerations per 100km	11.20%	1.98769E+12	<**0.01**
Phase 3 - Phase 4	Speed above the speed limits (km/h)	1.30%	2.27401E+12	<**0.01**
Phase 4 - Competition	Percentage of mobile use	-3.90%	27327559116	0.657
Phase 4 - Competition	Percentage of speeding time	-13.10%	72000240177	<**0.01**
Phase 4 - Competition	Harsh brakings per 100km	-3.20%	82995973298	<**0.01**
Phase 4 - Competition	Harsh accelerations per 100km	-10.30%	47511875657	<**0.01**
Phase 4 - Competition	Speed above the speed limits (km/h)	-20.90%	72401356187	<**0.01**
Competition - Challenges	Percentage of mobile use	10.00%	1502836826	<**0.01**
Competition - Challenges	Percentage of speeding time	50.70%	4321598228	<**0.01**
Competition - Challenges	Harsh brakings per 100km	41.50%	5775250817	<**0.01**
Competition - Challenges	Harsh accelerations per 100km	30.00%	2623882294	<**0.01**
Competition - Challenges	Speed above the speed limits (km/h)	24.30%	4081454761	<**0.01**
Challenges - Phase 6	Percentage of mobile use	2.90%	4712410638	<**0.01**
Challenges - Phase 6	Percentage of speeding time	4.00%	18068246934	<**0.01**
Challenges - Phase 6	Harsh brakings per 100km	-4.90%	21576313967	<**0.01**
Challenges - Phase 6	Harsh accelerations per 100km	1.80%	9765516199	<**0.01**
Challenges - Phase 6	Speed above the speed limits (km/h)	13.00%	18002330803	<**0.01**

SEM analysis

This section presents the results of the SEM analysis, focusing solely on the final models, presented in Table 3. In addition to the previously mentioned hard goodness-of-fit measures, the coefficient estimates produced were evaluated to ensure they provided logical and interpretable results. Efforts were made to avoid model misspecification by considering the appropriateness of the theoretical framework and the outcomes produced. During the modeling process, it was evident that certain model structures were significantly better suited to the experimental data based on specific criteria; only these best-fitting models are presented here. Variations within each latent variable structure were explored using the backwards elimination technique. All statistical analyses were conducted in R-studio (R Core Team, 2013) and SEM analysis in particular utilized the lavaan R package. Ultimately, the proposed SEM structure retained two latent unobserved variables:

- Feedback, expressing the influence of the different features of the smartphone app during the different phases of the experiment, namely Baseline, Scorecard feature, Maps feature, Compare feature, Competition and Challenges feature.
- Exposure, expressing the influence of the exposure metrics, namely Distance (for driving speed 30km/h – 50km/h), Morning peak and Afternoon peak.

Table 3: SEM model of Percentage of speeding time, Harsh Brakings per 100km & Harsh Accelerations per 100km

SEM Components		**Parameters**	**Estimate**	**S.E.**	**z-value**	**P(>|z|)**
Latent Variables	Feedback	Baseline	1.000	–	–	–
		Scorecard feature	2.076	0.014	148.640	**0.000**
		Maps feature	1.646	0.010	157.864	**0.000**
		Compare feature	1.215	0.029	41.754	**0.000**
		Competition & Challenges feature	2.053	0.038	54.447	**0.000**
	Exposure	Distance (for driving speed 30km/h – 50km/h)	1.000	–	–	–
		Morning peak	2.473	0.350	7.072	**0.000**
		Afternoon peak	-1.360	0.129	-10.579	**0.000**
Regressions	Percentage of speeding time	Intercept	0.409	0.003	138.941	**0.000**
		Exposure	0.326	0.043	7.627	**0.000**
		Feedback	-0.214	0.014	-15.655	**0.000**
	Harsh Accelerations per 100km	Intercept	0.099	0.001	95.037	**0.000**
		Exposure	0.028	0.010	2.769	0.006
		Feedback	0.026	0.004	6.493	**0.000**
		Competition & Challenges feature	-0.001	0.000	-2.748	**0.000**
		Afternoon peak	0.006	0.002	3.095	**0.002**
	Harsh Brakings per 100km	Intercept	0.184	0.001	158.258	**0.000**
		Exposure	0.077	0.014	5.542	**0.000**
		Feedback	-0.027	0.005	-4.976	**0.000**
Covariances	Percentage of speeding time	Harsh Brakings per 100km	0.007	0.001	7.686	**0.000**
	Harsh Accelerations per 100km	Percentage of speeding time	0.006	0.001	9.526	**0.000**
	Harsh Brakings per 100km	Harsh Accelerations per 100km	0.021	0.000	75.739	**0.000**
	Feedback	Exposure	-0.001	0.000	-5.558	**0.000**
Goodness-of-fit measures		CFI	0.940			
		TLI	0.944			
		RMSEA	0.049			0.845
		SRMR	0.025			

The model's goodness-of-fit measures further support the robustness of these findings, with a CFI of 0.940, TLI of 0.944, RMSEA of 0.049, and SRMR of 0.025. All four examined goodness-of-fit measures and the signs of the estimated coefficients indicate an excellent model fit. Additionally, the model's AIC was the lowest among the tested combinations, and no negative variances were observed, which would have suggested model misspecification (variance outputs are omitted for brevity). To enhance model fit, several variables were scaled linearly by factors of 10, which reduced variance discrepancies without affecting the interpretation of the coefficients. Moreover, an iterative process was employed to integrate covariances of the measured variables into the model. This involved comparing observed and fitted covariance correlations and addressing the largest discrepancies by including relevant covariance pairs, provided there were no significant theoretical objections. This approach significantly improved the overall model fit.

The path diagram of the present model is presented on Figure 5; green arrows denote positive correlations, while red arrows denote negative correlations. Several useful insights can be obtained from the produced SEM model results that are discussed in detail in the following subsections.

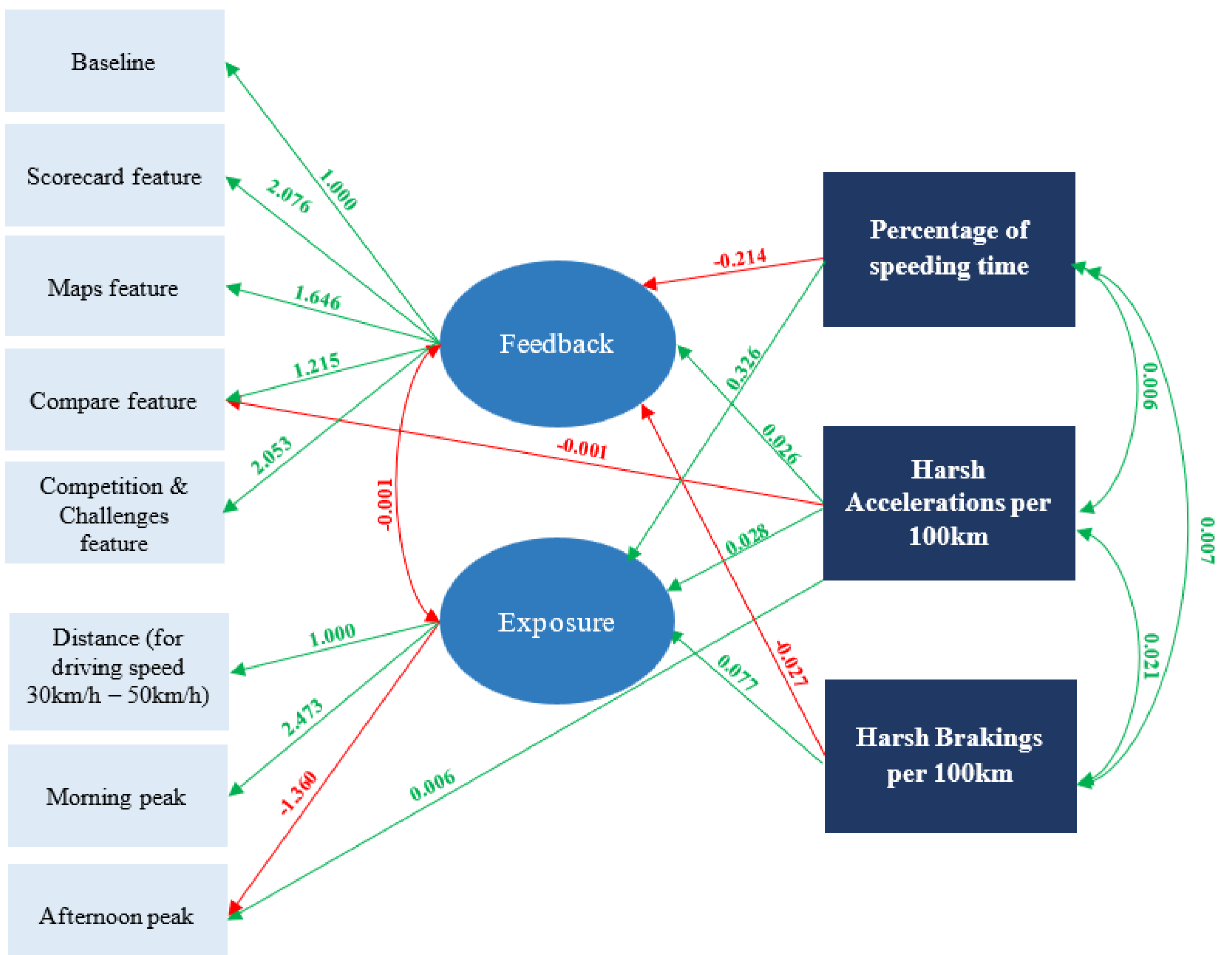

Figure 5: Path diagram of SEM model for percentage of speeding time, harsh accelerations per 100km and harsh brakings per 100km

Feedback

The SEM analysis reveals that feedback mechanisms, including the scorecard, maps, compare, and competition & challenges features, significantly impact driver behavior. The scorecard feature has the highest positive estimate at 2.076 ($p < 0.001$), indicating its crucial role in modifying driving habits. Similarly, the maps feature shows a strong influence with an estimate of 1.646 ($p < 0.001$), suggesting that providing drivers with map-based feedback can effectively encourage safer driving practices. The compare feature, with an estimate of 1.215 ($p < 0.001$), helps drivers assess their performance relative to others, positively influencing behavior. Additionally, the competition & challenges feature, with an estimate of 2.053 ($p < 0.001$), motivates drivers through competitive elements, reinforcing safe driving behaviors. Overall, these feedback mechanisms are effective in reducing the percentage of speeding time (feedback estimate: -0.214, $p < 0.001$) and harsh braking incidents (feedback estimate: -0.027, $p < 0.001$), although there is a slight increase in harsh accelerations (feedback estimate: 0.026, $p < 0.001$), which may require further refinement of the feedback system.

Exposure

Exposure factors, particularly the times of day, play a significant role in driving behaviors. Morning peak exposure is associated with increased driving aggressiveness, as indicated by the significant positive estimate of 2.473 ($p < 0.001$). This suggests that drivers are more likely to engage in risky behaviors, such as speeding and harsh braking, during morning peak hours. In contrast, afternoon peak exposure has a negative estimate of -1.360 ($p < 0.001$), indicating that driving behaviors may be less aggressive during this time. The distance driven at speeds between 30km/h and 50km/h serves as a reference parameter with an estimate of 1.000. Understanding these exposure patterns can help in designing targeted interventions to mitigate risky driving behaviors during specific times of the day.

Regressions

The regression analysis provides insights into how exposure and feedback influence specific driving behaviors. The percentage of speeding time is positively associated with exposure, with an estimate of 0.326 ($p < 0.001$), indicating that increased driving time leads to more speeding. However, feedback mechanisms significantly reduce speeding, with an estimate of -0.214 ($p < 0.001$), highlighting their effectiveness. Harsh accelerations per 100km show a slight positive association with exposure (estimate: 0.028, $p = 0.006$) and feedback (estimate: 0.026, $p < 0.001$), suggesting that while feedback reduces some risky behaviors, it may inadvertently increase others. The competition & challenges feature slightly reduces harsh accelerations (estimate: -0.001, $p < 0.001$), demonstrating its potential in moderating aggressive driving. Afternoon peak exposure increases harsh accelerations (estimate: 0.006, $p = 0.002$), further emphasizing the need for targeted interventions. Harsh brakings per 100km are positively influenced by exposure (estimate: 0.077, $p < 0.001$) but significantly reduced by feedback (estimate: -0.027, $p < 0.001$), reinforcing the importance of feedback in promoting safer driving.

Covariances

The covariance analysis reveals strong interrelationships between various driving behaviors. There is a positive correlation between the percentage of speeding time and harsh brakings per 100km (estimate: 0.007, $p < 0.001$), indicating that drivers who speed are also more likely to brake harshly. Similarly, there is a positive correlation between speeding and harsh accelerations (estimate: 0.006, $p < 0.001$), suggesting that these behaviors often co-occur in aggressive driving patterns. The strong positive correlation between harsh brakings and harsh accelerations (estimate: 0.021, $p < 0.001$) further supports this finding. Additionally, a negative correlation between feedback and exposure (estimate: -0.001, $p < 0.001$) indicates that increased feedback is associated with decreased exposure to risky driving conditions. These covariance relationships highlight the interconnected nature of different driving behaviors and the importance of comprehensive intervention strategies to address multiple aspects of driver behavior simultaneously.

CONCLUSIONS

Rapid technological advances, especially in telematics and Big Data analytics, as well as the increasing penetration and use of information technology by drivers (e.g. smartphones), provide new capabilities for monitoring and analyzing driving behavior. In this paper, we aimed to examine the effect of driver feedback via a smartphone application on driving behavior risk indicators within a multiphase naturalistic driving experiment. First, a preliminary analysis highlighted the beneficial effects of upgraded feedback on key risk indicators across experiment phases. Subsequently, SEM analysis on a 73,869 trip dataset provided significant insights into how feedback mechanisms and exposure factors influence driving behaviors.

Feedback features, namely scorecard, maps, compare, and competition & challenges seem to be effective in reducing risky driving behaviors like speeding and harsh braking, although they may slightly increase harsh accelerations, at least some of the feedback features. Morning peak exposure is associated with more aggressive driving, while afternoon peak exposure tends to be less risky. Additionally, the strong positive correlations between speeding, harsh braking, and harsh accelerations highlight the interconnected nature of aggressive driving behaviors, confirming previous studies (46) and showcasing the importance of driver behavior analysis.

The present findings come with some practical implications, as well. First, further in depth examination of driver feedback is necessary to quantify the complex relationships involved in various driving tasks, taking also into account the 2 other road safety pillars; environment and vehicle. On that note, modifications to vehicle and mobile phone interfaces could be beneficial, particularly with the expected rise in connectivity and automation in the future. Additionally, incorporating eco-driving feedback is crucial to understanding and enhancing the effectiveness of feedback mechanisms.

The ultimate goal of providing feedback to drivers is to activate the process of learning and self-assessment, enabling them to gradually improve their performance and monitor their progress. This process involves establishing detailed cause-and-effect relationships between aggressive driving and associated risks, offering valuable insights for improving road safety. Such information is beneficial for insurance companies, fleet management applications, and identifying hazardous geographical locations on the road network. Additionally, feedback can serve as a tool for objectively proving driving behavior, allowing users to gain benefits from their insurance companies or to regain their driver's license after revocation.

ACKNOWLEDGMENTS

The authors would like to thank OSeven Telematics for providing all necessary data exploited to accomplish this study.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: A. Kontaxi, A. Ziakopoulos, G. Yannis; data collection: A. Kontaxi, A. Ziakopoulos; analysis and interpretation of results: A. Kontaxi, A. Ziakopoulos; draft manuscript preparation: A. Kontaxi, A. Ziakopoulos, G. Yannis. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. *Global Status Report on Road Safety 2023. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO.*

2. Singh, S. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. *DOT HS 812 115*, 2015.

3. Koesdwiady, A., R. Soua, F. Karray, and M. S. Kamel. Recent Trends in Driver Safety Monitoring Systems: State of the Art and Challenges. *IEEE Transactions on Vehicular Technology*, Vol. 66, No. 6, 2017, pp. 4550–4563. https://doi.org/10.1109/TVT.2016.2631604.

4. Chan, T. K., C. S. Chin, H. Chen, and X. Zhong. A Comprehensive Review of Driver Behavior Analysis Utilizing Smartphones. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 10, 2020, pp. 4444–4475. https://doi.org/10.1109/TITS.2019.2940481.

5. Ziakopoulos, A., D. Tselentis, A. Kontaxi, and G. Yannis. A Critical Overview of Driver Recording Tools. *Journal of Safety Research*, Vol. 72, 2020, pp. 203–212. https://doi.org/10.1016/J.JSR.2019.12.021.

6. Castignani, G., T. Derrmann, R. Frank, and T. Engel. Driver Behavior Profiling Using Smartphones: A Low-Cost Platform for Driver Monitoring. *IEEE Intelligent Transportation Systems Magazine*, Vol. 7, No. 1, 2015, pp. 91–102. https://doi.org/10.1109/MITS.2014.2328673.

7. Mantouka, E. G., E. N. Barmpounakis, and E. I. Vlahogianni. Identifying Driving Safety Profiles from Smartphone Data Using Unsupervised Learning. *Safety Science*, Vol. 119, No. January, 2019, pp. 84–90. https://doi.org/10.1016/j.ssci.2019.01.025.

8. Kontaxi, A., D. M. Tzoutzoulis, A. Ziakopoulos, and G. Yannis. Exploring Speeding Behavior Using Naturalistic Car Driving Data from Smartphones. *Journal of Traffic and Transportation Engineering (English Edition)*, Vol. 10, No. 6, 2023, pp. 1162–1173. https://doi.org/10.1016/J.JTTE.2023.07.007.

9. Richard, C., J. Campbell, J. Brown, M. Lichty, S. Chrysler, and R. Atkins. Investigating Speeding Behavior with Naturalistic Approaches. *Transportation Research Record*, Vol. 1, No. 2365, 2013, pp. 58–65. https://doi.org/10.3141/2365-08.

10. Papantoniou, P., A. Kontaxi, G. Yannis, and P. Fortsakis. Investigating the Correlation of Mobile Phone Use with Trip Characteristics Recorded Through Smartphone Sensors. No. 1278, 2021, pp. 935–944.

11. Ziakopoulos, A., A. Kontaxi, and G. Yannis. Analysis of Mobile Phone Use Engagement during Naturalistic Driving through Explainable Imbalanced Machine Learning. *Accident Analysis and Prevention*, Vol. 181, No. July 2022, 2023, p. 106936. https://doi.org/10.1016/j.aap.2022.106936.

12. Frantzola, E.-K., A. Kontaxi, and G. Yannis. Impact of Road and Traffic Characteristics on Driver Behaviour and Safety Using Data from Smartphones. 2022.

13. Precht, L., A. Keinath, and J. F. Krems. Effects of Driving Anger on Driver Behavior – Results from Naturalistic Driving Data. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 45, 2017, pp. 75–92. https://doi.org/10.1016/J.TRF.2016.10.019.

14. Dingus, T. A., V. L. Neale, S. G. Klauer, A. D. Petersen, and R. J. Carroll. The Development of a Naturalistic Data Collection System to Perform Critical Incident Analysis: An Investigation of Safety and Fatigue Issues in Long-Haul Trucking. *Accident Analysis & Prevention*, Vol. 38, No. 6, 2006, pp. 1127–1136. https://doi.org/10.1016/J.AAP.2006.05.001.

15. Wang, X., R. Xu, S. Zhang, Y. Zhuang, and Y. Wang. Driver Distraction Detection Based on Vehicle Dynamics Using Naturalistic Driving Data. 2022. https://doi.org/10.1016/j.trc.2022.103561.

16. Shahverdy, M., M. Fathy, R. Berangi, and M. Sabokrou. Driver Behavior Detection and Classification Using Deep Convolutional Neural Networks. *Expert Systems with Applications*, Vol. 149, 2020, p. 113240. https://doi.org/10.1016/J.ESWA.2020.113240.

17. Shi, X., Y. D. Wong, M. Z. F. Li, C. Palanisamy, and C. Chai. A Feature Learning Approach Based on XGBoost for Driving Assessment and Risk Prediction. *Accident Analysis & Prevention*, Vol. 129, 2019, pp. 170–179. https://doi.org/10.1016/J.AAP.2019.05.005.

18. Zhang, X., X. Wang, Y. Bao, and X. Zhu. Safety Assessment of Trucks Based on GPS and In-Vehicle Monitoring Data. *Accident Analysis and Prevention*, Vol. 168, 2022. https://doi.org/10.1016/j.aap.2022.106619.

19. Braun, D., E. Reiter, and A. Siddharthan. Creating Textual Driver Feedback from Telemetric Data. 2015.

20. Gu, Y., Q. Wang, and S. Kamijo. Intelligent Driving Data Recorder in Smartphone Using Deep Neural Network-Based Speedometer and Scene Understanding. *IEEE Sensors Journal*, Vol. 19, No. 1, 2019, pp. 287–296. https://doi.org/10.1109/JSEN.2018.2874665.

21. Archer, J. C. State of the Science in Health Professional Education: Effective Feedback. *Medical Education*, Vol. 44, No. 1, 2010, pp. 101–108. https://doi.org/10.1111/J.1365-2923.2009.03546.X.

22. Hattie, J., and H. Timperley. The Power of Feedback. *Review of educational research*, Vol. 77, No. 1, 2007, pp. 81--112.

23. London, M., and J. W. Smither. Feedback Orientation, Feedback Culture, and the Longitudinal Performance Management Process. *Human Resource Management Review*, Vol. 12, No. 1, 2002, pp. 81–100. https://doi.org/10.1016/S1053-4822(01)00043-2.

24. Meuleners, L., M. Fraser, M. Stevenson, and P. Roberts. Personalized Driving Safety: Using Telematics to Reduce Risky Driving Behavior among Young Drivers. *Journal of Safety Research*, Vol. 86, 2023, pp. 164–173. https://doi.org/10.1016/J.JSR.2023.05.007.

25. Chen, W., and B. Donmez. A Naturalistic Driving Study of Feedback Timing and Financial Incentives in Promoting Speed Limit Compliance. *IEEE Transactions on Human-Machine Systems*, Vol. 52, No. 1, 2022, pp. 64–73. https://doi.org/10.1109/THMS.2021.3117234.

26. Peer, S., A. Muermann, and K. Sallinger. App-Based Feedback on Safety to Novice Drivers: Learning and Monetary Incentives. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 71, 2020, pp. 198–219. https://doi.org/10.1016/j.trf.2020.04.005.

27. Soleymanian, M., C. B. Weinberg, and T. Zhu. Sensor Data and Behavioral Tracking: Does Usage-Based Auto Insurance Benefit Drivers? *Marketing Science*, Vol. 38, No. 1, 2019, pp. 21–43. https://doi.org/10.1287/mksc.2018.1126.

28. Ghamari, A., N. Rezaei, M. R. Malekpour, S. Azadnajafabad, A. Jafari, N. Ahmadi, K. Bhalla, S. Shahraz, D. Bose, and F. Farzadfar. The Effect of Non-Punitive Peer Comparison and Performance Feedback on Drivers' Behavior Using the Telematics: The First Randomized Trial in Iran. *Journal of Safety Research*, Vol. 82, 2022, pp. 430–437. https://doi.org/10.1016/j.jsr.2022.07.010.

29. Merrikhpour, M., B. Donmez, and V. Battista. A Field Operational Trial Evaluating a Feedback-Reward System on Speeding and Tailgating Behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 27, No. PA, 2014, pp. 56–68. https://doi.org/10.1016/j.trf.2014.09.002.

30. Hari, D., C. J. Brace, C. Vagg, J. Poxon, and L. Ash. Analysis of a Driver Behaviour Improvement Tool to Reduce Fuel Consumption. *Proceedings - 2012 International Conference on Connected Vehicles and Expo, ICCVE 2012*, 2012, pp. 208–213. https://doi.org/10.1109/ICCVE.2012.46.

31. Husnjak, S., D. Peraković, I. Forenbacher, and M. Mumdziev. Telematics System in Usage Based Motor Insurance. *Procedia Engineering*, Vol. 100, No. January, 2015, pp. 816–825. https://doi.org/10.1016/J.PROENG.2015.01.436.

32. Mazureck, U., and J. van Hattem. Rewards for Safe Driving Behavior. *https://doi.org/10.1177/0361198106198000106*, Vol. 1980, No. 1, 2006, pp. 31–38. https://doi.org/10.1177/0361198106198000106.

33. Kontaxi, A., A. Ziakopoulos, and G. Yannis. Trip Characteristics Impact on the Frequency of Harsh Events Recorded via Smartphone Sensors. *IATSS Research*, Vol. 45, No. 4, 2021, pp. 574–583. https://doi.org/10.1016/J.IATSSR.2021.07.004.

34. Toledo, T., and T. Lotan. In-Vehicle Data Recorder for Evaluation of Driving Behavior and Safety. *Transportation Research Record*, No. 1953, 2006, pp. 112–119. https://doi.org/10.3141/1953-13.

35. Bell, J. L., M. A. Taylor, G. X. Chen, R. D. Kirk, and E. R. Leatherman. Evaluation of an In-Vehicle Monitoring System (IVMS) to Reduce Risky Driving Behaviors in Commercial Drivers: Comparison of in-Cab Warning Lights and Supervisory Coaching with Videos of Driving Behavior. *Journal of Safety Research*, Vol. 60, 2017, pp. 125–136. https://doi.org/10.1016/j.jsr.2016.12.008.

36. Aidman, E., C. Chadunow, K. Johnson, and J. Reece. Real-Time Driver Drowsiness Feedback Improves Driver Alertness and Self-Reported Driving Performance. *Accident Analysis & Prevention*, Vol. 81, 2015, pp. 8–13. https://doi.org/10.1016/J.AAP.2015.03.041.

37. Kontaxi, A., A. Ziakopoulos, and G. Yannis. Investigation of the Speeding Behavior of Motorcyclists through an Innovative Smartphone Application. *Traffic Injury Prevention*, Vol. 22, No. 6, 2021, pp. 460–466. https://doi.org/10.1080/15389588.2021.1927002/SUPPL_FILE/GCPI_A_1927002_SM3969.DOCX.

38. Papantoniou, P., C. Antoniou, G. Yannis, and D. Pavlou. Which Factors Affect Accident Probability at Unexpected Incidents? A Structural Equation Model Approach. *Journal of Transportation Safety & Security*, Vol. 11, No. 5, 2019, pp. 544–561. https://doi.org/10.1080/19439962.2018.1447523.

39. Wu, K. F. (Ken), and L. Wang. Exploring the Combined Effects of Driving Situations on Freeway Rear-End Crash Risk Using Naturalistic Driving Study Data. *Accident Analysis & Prevention*, Vol. 150, 2021, p. 105866. https://doi.org/10.1016/J.AAP.2020.105866.

40. Roussou, S., T. Garefalakis, E. Michelaraki, C. Katrakazas, M. Adnan, M. W. Khattak, T. Brijs, and G. Yannis. Examination of the Effect of Task Complexity and Coping Capacity on Driving Risk: A Cross-Country and Transportation Mode Comparative Study. *Sensors 2023, Vol. 23, Page 9663*, Vol. 23, No. 24, 2023, p. 9663. https://doi.org/10.3390/S23249663.

41. Ram, T., and K. Chand. Effect of Drivers' Risk Perception and Perception of Driving Tasks on Road Safety Attitude. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 42, 2016, pp. 162–176. https://doi.org/10.1016/J.TRF.2016.07.012.

42. Zhao, X., W. Xu, J. Ma, H. Li, and Y. Chen. An Analysis of the Relationship between Driver Characteristics and Driving Safety Using Structural Equation Models. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 62, 2019, pp. 529–545. https://doi.org/10.1016/J.TRF.2019.02.004.

43. Jöreskog, K. G. ., and Dag. Sörbom. LISREL 8 : User's Reference Guide. 2001, p. 378.

44. Newnam, S., I. Lewis, and A. Warmerdam. Modifying Behaviour to Reduce Over-Speeding in Work-Related Drivers: An Objective Approach. *Accident Analysis and Prevention*, Vol. 64, 2014, pp. 23–29. https://doi.org/10.1016/j.aap.2013.10.032.

45. Camden, M. C., S. A. Soccolich, J. S. Hickman, and R. J. Hanowski. Reducing Risky Driving: Assessing the Impacts of an Automatically-Assigned, Targeted Web-Based Instruction Program. *Journal of Safety Research*, Vol. 70, 2019, pp. 105–115. https://doi.org/10.1016/j.jsr.2019.06.006.

46. Su, Z., R. Woodman, J. Smyth, and M. Elliott. The Relationship between Aggressive Driving and Driver Performance: A Systematic Review with Meta-Analysis. *Accident Analysis & Prevention,* Vol. 183, 2023, p. 106972. https://doi.org/10.1016/J.AAP.2023.106972.