Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Investigating the Correlation between Driver's Characteristics and Safety Performance

Dimitrios I. Tselentis[a,*], Katerina Folla[a], Vassiliki Agathangelou[a], George Yannis[a]

[a]*Department of Transportation Planning and Engineering, National Technical University of Athens (NTUA), 5 Heroon Polytechniou Str., Athens GR-15773,Greece*

Abstract

This research aims to correlate drivers' characteristics with their safety performance. In order to achieve this objective, two different data sets were used deriving from 12 drivers who participated on an on-road driving experiment while being assessed by a safety behaviour expert. Drivers participating in the experiment also responded to the questionnaire administered, which mainly included self-evaluation questions along with other that concerned the condition of their vehicle, demographics, driving experience etc. The analysis of drivers' behavior was carried out using the statistical methods of factor analysis and linear regression analysis. Three groups that characterize drivers' perception of careless, aggressive and cautious driving behavior were derived from factor analysis. Moreover, linear regression analysis revealed that driving experience, headways, self-reported driving skill and defensive driving positively affect the overall on-road driving performance score. More precisely, it was ascertained that driving experience leads to statistically significant increase in overall on-road driving performance score. Other factors such as defensive driving and keeping safe distances from other vehicles. The findings of this research could be exploited for providing feedback on drivers regarding their driving behavior in order to further reduce their driving risk.

© 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: on-road assessment; self-assessment; driving behaviour; factor analysis; linear regression

1. Introduction

Road safety comprises one of the most significant problems in all developed countries, with the increase of road accidents during the last years constituting road safety a public health issue of top priority. However, the vast majority of road accidents worldwide (90%) occurs in low- or middle-income countries, while these countries account for only the half of registered motorized vehicles. Despite the fact that efforts have been made for the reduction of road

* Corresponding author. Tel.: +30-210-772-2210; fax: +30-210-772-1454.
E-mail address: dtsel@central.ntua.gr

2352-1465 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

accidents, it is estimated that worldwide 1.25 million persons are killed on an annual basis, half of them concern the "vulnerable road users", i.e. pedestrians, cyclists, powered two-wheelers (PTWs) (WHO, 2015).

In 2016, about 25,600 persons were killed in road accidents in the European Union (EU), while Greece with 69 fatalities per million population is among the EU countries with the worst road safety performance (European Commission, 2018). However, significant progress has been made with Greece having recorded one of the highest reductions of road fatalities in the EU during the period 2010-2015, which is partly attributed to the economic crisis.

There is a significant number of risk factors identified in literature, which affect accident probability. In order to tackle the road safety problem, the critical factors leading to road accidents have to be identified. The most important risk factors recognized in literature (WHO 2015, Elvik 2004) concern:

- **vehicle** (technical characteristics and equipment, active and passive safety)
- **road environment** (geometric characteristics, level of road maintenance, equipment, traffic regulations and characteristics, weather conditions)
- **road user** (driving experience, patterns of behavior, mental and social particularities)

Among them, human factors are likely to be the most important cause of road traffic fatalities and injuries every year and therefore the importance of studying how these factors can affect traffic risk is high. Consequently, studying the characteristics of driving behavior and the drivers' performance in terms of road safety is deemed highly critical in order to better understand the main causes leading to road accidents and thus, improve road safety performance.

The aim of the present research is the correlation of driver's characteristics and road safety performance. More precisely, a statistical methodology is employed to identify how the elements composing driving behavior interact with each other and define the overall traffic behavior of the driver'. The results of this research aim to shed further light on how and at what extent driver's road safety performance is related to driver's characteristics and affect the overall driving performance. Consequently, the identification of the characteristics contributing to a more dangerous driving behavior compared to those of more cautious drivers is achieved in the present research.

2. Literature review

Several studies have explored driver's behavior and attitudes and have led to driving profiles based on either observation of drivers or on self-reported driving behavior. A similar study conducted by Kontogiannis et al. (2002) aimed to classify drivers based on their self-reported offending behavior. Within the context of this study, 1,425 questionnaires were collected from 18 cities in Greece, in which drivers responded on how often they presented a specific behavior on a six-point scale. The research resulted in three types of offending behaviors with driving mistakes and road rage being the two principal ones. Two more categories were also found through factor analysis; the first refers to a situation of low readiness and negligence, while the latter refers to communication errors and "social indifference" towards the other road users. On the other hand, Vardaki et al. (2013) explored the self-reported drivers' behavior and attitudes concerning traffic infringements related to speeding, drink-driving and use of mobile phone while driving by using cluster analysis. Data come from the pan-European survey SARTRE 4, in which 601 Greek drivers participated. The results showed that differences in the drivers' attitudes and behavior are related to factors, such as age, gender and residence place. Additionally, in this research, the attitudes of Greek drivers on traffic infringements, especially on those affecting negatively road safety, are revealed. Moreover, another study (Moghaddam et al., 2014) explored the relationship between driver's characteristics and offensive driving behavior. Data were collected through questionnaires responded by 1,769 drivers (593 females and 1,176 males). The research led to four groups of drivers according to their risky behavior based on traffic violations, number of accidents and number of fines for traffic offences. A risk indicator was thus estimated which can be used for the assessment of drivers according to personality related characteristics.

Furthermore, studies using relevant methodologies in road safety analyses were explored, with emphasis on those based on drivers' personal estimations through questionnaires or interviews (self-report study) and on observation of participants' behavior under normal driving conditions. Ott et al. (2012) compared a standardized road test with naturalistic driving behavior of older drivers with cognitive impairments through factor analysis. In the first phase of the experiment, factor analysis was one-dimensional, reflecting one critical factor related to sensitivity and driving

behavior. On the contrary, the analysis of naturalistic driving led to the identification of two factors, i.e. proper lane maintenance and driver's ability to respond to traffic conditions. The proper lane maintenance is an important road safety factor, which is not widely used in the assessment tests. Another study (Amado et al., 2014) explored drivers' perceptions on their driving skills, as well as their driving behavior according to experts' judgment in real driving conditions. Low-level correlations were recorded between drivers' and experts' judgments on driving skills and behavior, mainly on those concerning mistakes and infringements related to speeding and traffic lights. Another important finding was that drivers overestimated their driving performance in comparison with experts' assessments.

Concluding, studies presented in this section cover a wide range of analyses focusing on drivers' characteristics and their relation to road safety performance. The main literature findings are summarized below:

- The various types of mistakes may pre-define a possible offending behavior of the driver.
- On that purpose, efforts have been made for the development of a driver's risk indicator with aim to assess the effects of driver's characteristics on road safety.
- Most researchers have used factor analysis, in order to minimize the dimensionality of a dataset and identify the critical factors that describe adequately driving behavior.
- Another statistical technique is cluster analysis, which classifies observations in two or more mutually exclusive unknown groups, based on combinations of the variables. Thus, drivers are classified as "less risky" and "risky" or "efficient" and "inefficient" drivers.
- Several studies showed that while drivers self-assessed as good drivers, the experts observed difficulty in compromising with speeding and traffic lights infringements.

The above mentioned findings highlight the need of further exploring road safety indicators deriving from the assessment of driving under real driving conditions and correlating them with drivers' characteristics, in order to provide feedback and recommendations on targeted road user groups on how to further enhance their driving behaviour and ultimately reduce driving risk. This would improve road safety not only on an individual level but on a national level as well.

3. Methodological framework

For the purposes of the present research, the analysis of a large number of correlated variables is necessary, while each variable is considered equally significant with the remaining variables. For this reason, the statistical analysis of the selected data was performed in two steps. In the first step, factor analysis was performed in order to reduce the dimensionality of the dataset and identify the main factors and then, a multiple linear regression model was developed.

3.1. Factor analysis

Factor analysis is a statistical technique, which aims to create underlying non-observable quantities called factors. This model is based on the assumption that the variables can be grouped based on their correlations with each other. Thus, all variables lying into the same group are highly correlated, while they present small correlations with the variables of the other groups of variables. Consequently, each group of variables represents an underlying factor, which is the cause of the observed correlations. Factor analysis attempts to correlate the unobserved variables (factors or components) and the observable variables for which measurements exist, achieving thus a grouping of the observed variable into common components.

Concerning the estimation of the main factors, there are various methods, such as principal component analysis, principal axis factoring, alpha factoring, image factoring, unweighted least-squares method, generalized least-squares method and maximum-likelihood method. The most common methods used in the factor analysis extraction are the principal component analysis and the maximum-likelihood method. For the purposes of the present research, the principal component analysis was considered the most appropriate method for the reduction of the large number of variables.

3.2. Multiple linear regression

The multiple linear regression concerns the case that a continuous variable Y has a linear relationship with one or more variables X (X_1, X_2, X_3, …, X_k). The equation describing this relationship between the dependent and independent variables is the following:

$$Y_i = \beta_0 + \beta_{1*} X_{1i} + \beta_{2*} X_{2i} + \beta_3 * X_{3i+} \ldots + \beta_{k*} X_{ki} + \varepsilon_i \quad (1)$$

The main assumption of the multiple linear regression is that there is no correlation among the independent variables ($\rho(x_i, x_j) \rightarrow 0$, for each $i \neq j$).

The accuracy of the model is assessed through the coefficient of determination R squared (R^2). R^2 shows the percentage of the variability of the dependent variable Y explained by the independent variables X included in the model. R^2 has values between 0 and 1, with 1 indicating that the variable Y is fully explained by the variables X.

4. Data collection and processing

In this section, the collection and process of the data used is presented. For the purposes of the present research, an experiment of on-road driving took place, in which 12 drivers participated and were assessed on the basis of 16 principal characteristics that compose driver's road safety performance. Then, the participants responded to a survey of 78 questions, which mainly concerned driving behavior characteristics.

4.1. Experiment

The experiment comprised of an 18.4 kilometers route of urban and interurban road network with a mean duration of approximately 27 minutes. However, only the section of the interurban road network was selected as the most appropriate for the conduction of the experiment due to the low traffic volume.

For the needs of the experiment, driver's performance was assessed based on a list of 16 principal driving indicators reflecting driver's road safety performance. The indicators included are presented in table 1:

Table 1. List of driving indicators used for the assessment of driver's performance

a/a	Indicator	a/a	Indicator
1	Speed adaptation	9	Understanding, perception and quality of traffic participation
2	Braking	10	Crossing or junction
3	Accelerating	11	Anticipation and perception of road signs and traffic signals
4	Turning	12	Joining the traffic stream
5	Headways	13	Visual behavior and communication
6	Lateral Position	14	Mirror use
7	Ability to choose the correct lane	15	Use of direction indicator
8	Lane change	16	Steering firmness

Based on the overall behavior of the driver, the expert assessed the participants for their defensive behavior while driving and thus, another indicator called "Defensive driving" was created. This indicator is defined as the total actions and overall behavior, with which the driver is allowed to forecast, identify and take all necessary measures in order to reduce accident risk involved.

Finally, drivers were assessed for each of the aforementioned indicators at a four points scale (Bad, Insufficient, Sufficient, Good), as well as for their overall behavior with an indicator called "overall on road driving performance score" at a scale from 0% to 100%.

4.2. Survey

After the on road driving experiment, the drivers were asked to respond to a survey of 78 questions. The questions concerned basic demographics and information on a wide range of characteristics related to participants' driving experience (kilometers travelled inside and outside urban areas, etc.) and potential offending behavior (number of accidents, traffic infringements, etc.). The questionnaire also included self-evaluation questions regarding their driving behavior (risky driving, steep acceleration, how cautious is as a driver, etc.).

5. Results

5.1. Factor Analysis Results

Through factor analysis, the main factors explaining the correlation of driver's characteristics and road safety performance were attempted to be defined. Ten variables were examined within this analysis, the seven of which were derived from drivers' answers in the following questions:

- In how many accidents have you been involved as a driver?
- During the last two years, how many times did you violate the traffic code while driving?
- At what extent do you keep up with the speed limits while driving on motorways?
- How aggressive would you assess yourself as a driver?
- How often do you consider your braking to be harsh?
- How often do you consider your acceleration to be harsh?
- How often do you consider your turning to be harsh?

The three remaining variables concern the rating coming from the assessment of on road driving test for the following safety indicators:

- Braking
- Accelerating
- Headways

Several groups of variables were tested with a sufficient number of variables within each group. Among all variables tested, those that were not taken into account in the statistical analysis concerned mainly the demographic characteristics of participants (gender, occupation, marital status, education level) and information related to participants' vehicles (cubism, insurance).

The statistical process of the aforementioned data led to the identification of the statistically significant variables. From the initial 78 questions and the 16 variables coming from the assessment of drivers, most variables were excluded due to either low significance or high correlation with other variables.

In Table 2, the factors extracted from the analysis and the distribution of the variables included in each factor are presented. The variable loadings show the correlation between the variables and the factors. The higher is the factor loading the stronger is the relationship between this variable and the factor. It is noted that the variables with loading higher than 0.35 were considered the most significant for the estimation of the factors.

The sign of the value indicates whether each variable is responded towards the low values (negative sign) or the high values (positive sign), in order to be interpreted accordingly to the scale at which is measured. As shown below, 4 factors are extracted, from which only the three first variables will be further analyzed, since the last one cannot be adequately interpreted.

Table 2. Variable loadings on the estimated factors

Variables	Factor 1	Factor 2	Factor 3	Factor 4
In how many accidents were you involved as a driver?	.740			
During the last two years, how many times did you offend the Traffic Law while driving?		.714	-.541	
At what extent do you keep the speed limits while driving on motorways?	-.926			
How aggressive would you assess yourself as a driver?		-.667		
How often do you consider your braking to be harsh?	.860			
How often do you consider your acceleration to be harsh?	.680		.660	
How often do you consider your turning to be harsh?	.898			
Braking		-.523		
Accelerating		.589		.625
Headways			-.550	.653

Table 3 shows the percentage of the variance that is accounted for by each variable in the estimation of the factor, as well as the cumulative percentage of variance accounted for by the first n variables before and after the rotation. It is noticed that only for the first four factors, the eigenvalues are high (higher than 1.0). Consequently, only the results for these factors after the rotation are shown. As referred above, the three first factors are analyzed, which explains 65.84% of the total variance. It is noted that a cumulative percentage of the total variance over 60% is considered satisfactory for the extracted factors.

Table 3. Percentage of variance explained for each factor

Factor	Initial Eigenvalues			Rotation Eigenvalues		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	3.745	37.451	37.451	3.671	36.712	36.712
2	1.903	19.032	56.483	1.668	16.678	53.389
3	1.477	14.769	71.253	1.515	15.150	**68.539**
4	1.089	10.890	82.143	1.360	13.604	82.143
5	.791	7.912	90.055			
6	.496	4.956	95.011			
7	.339	3.391	98.402			
8	.100	1.002	99.404			
9	.052	.524	99.928			
10	.007	.072	100.000			

The Kaiser-Meyer-Olkin (KMO) Criterion is used as a measure of the sample adequacy. The KMO is $0.24 < 0.50$ (0.50 is considered an acceptable value), which is not satisfactory, but was expected due to the small number of participants. Concerning the Bartlett's test of sphericity, which examines the existence of correlations among the variables and therefore the suitability of the variables for structure detection, the significance value of the test was $0.02 (< 0.05)$, which is acceptable.

Overall, the results of the factor analysis showed that:

- The first factor explains 36.71% of the total variance and includes the number of accidents that drivers have been involved in, the extent at which the speed limits on motorways are respected and the frequency at which drivers brake, turn and accelerate harshly. Thus, this factor reflects driver's perception on careless driving.
- The second factor refers to driver's aggressiveness and explains 16.68% of the total variance. It reveals driver's aggressive behavior, since the variables included are the number of traffic law infringements, driver's self-assessment on aggressive driving and expert's assessment of the driver on harsh braking and accelerating under normal driving conditions on an interurban road.
- The last factor explains 15.15% of the total variance and concerns cautious driving, since driver's self-reported frequency of harsh braking events and expert's assessment on headways are included in its estimation.

Consequently, the three factors extracted from the factor analysis explain 68.54% of the total variance and describe drivers' characteristics, their perceptions on their driving behavior, as well as their road safety performance under normal driving conditions, as assessed by an expert.

5.2. Multiple linear regression model

A large number of tests performed resulted to the development of the multiple linear regression model that appears in Table 4, which includes the following variables:

Dependent variable:

Overall on road driving performance score: Indicator of driver's overall performance under normal driving conditions, as assessed by an expert, at a scale from 0% to 100%, which is used as a dependent variable.

Independent variables:

- **Years of driving**: This variable concerns how many years the participant drives and is the answer in the question "How many years do you drive?" (continuous variable).
- **Headways 4**: The participant keeps a safe distance to the vehicle in front without a further correction (Score 4= Good), with reference value being the variable "Headways 2".
- **Efficient driver 3**: This variable concerns the participants that were self-assessed as "medium efficient drivers" when they were asked "How efficient driver do you consider yourself?". As a reference value was used the variable "Efficient driver 2".
- **Defensive driving 3**: The participant presents a sufficiently defensive driving behavior by forecasting, identifying and taking all necessary measures in order to safely avoid potential accident risks (Score 3= Sufficient). As a reference value was used the variable "Defensive driving 2".

Table 4. Parameter estimates and fit of the linear regression model

	Unstandardized Coefficients		Standardized Coefficients	t	p-value
	B	Std. error	Beta		
(Constant)	.449	.056		8.067	.000
Years of driving	.010	.003	.569	3.250	.014
Headways 4	.142	.044	.555	3.198	.015
Efficient driver 3	.141	.049	.554	2.857	.024
Defensive driving 3	.108	.044	.424	2.444	.045
Adjusted R^2	0.721				
F statistic	8.155				
df	11				

All the parameter estimates are statistically significant at 95% confidence intervals (p-values <0.05), the F statistic is 8.155 for 11 degrees of freedom (p-value=0.09) and the R^2 is 0.721, which leads to the acceptance of the model. The value of R^2 indicates that the variables included in the model explain the 72.1% of the variance of the "Overall on road driving performance score" variable.

The modelling results can be analyzed as follows:

- The variable "Years of driving" has a positive relationship with the dependent variable, showing that as driving experience increases, the overall road safety performance score increases as well.
- The positive sign of the variable "Headways 4" shows that the overall on road driving performance score of drivers assessed as Good for the indicator Headways is higher by 0.142 compared to drivers assessed as insufficient (Headways2) for this indicator.
- Respectively, users who consider themselves drivers of medium efficiency have a higher overall driving performance score by 0.141 compared to those who consider themselves less efficient.
- The overall on road driving performance of drivers who present a sufficiently defensive driving behavior is higher by 0.108 compared to those assessed as insufficient.

The elasticity analysis conducted led to the following results:

- The influence of the variable "Years of driving" is higher by 4.06 times compared to the variable "Efficient driver3" and by 1.70 times compared to the variable "Defensive driving 3", which shows the significance of driving experience when estimating the road safety performance of a driver.
- From the remaining three independent variables, the "Defensive driving" has the highest influence on road safety performance. More specifically, its influence is by 2.39 and 2.34 times higher compared to "Efficient driver 3" and "Headways4" respectively. This may be attributed to the fact that the latter indicators are individual characteristics of driver's behavior, while the aggressive driving includes a number of characteristics and consequently affects more the overall road safety performance score.

6. Conclusions

The objective of the present research is to investigate the relationship between drivers' characteristics and their road safety performance. To this end, a statistical methodology is applied to identify the characteristics of drivers that contribute to a more risky driving behavior compared to those of more cautious drivers that reduce accident risk. An on-road driving experiment was conducted, in which 12 drivers participated, while further necessary data were collected through a survey administered to those participated in the experiment.

One of the most important findings is that driving experience is the most significant predictor of driver's safety and driving performance, as expressed by the overall road safety performance score. The second most important factor of driver's road safety performance is the "defensive driving" variable, as assessed under naturalistic driving conditions, including those actions and behaviors, with which driver is able to forecast, identify and take measures in order to reduce the risk of an accident. Other behavioral factors affecting the overall road safety performance of a driver found are "headways" and "efficiency". The first one accounts for the expert’s assessment of the driver' s ability to keep adequate distance from the vehicle in front and increases driver's road safety performance. The latter represents the self-reported driving efficiency that is found to be correlated with driver's overall safety performance.

Furthermore, three patterns of driving behavior are defined through factor analysis, based on data collected from the questionnaires and the most significant safety performances. The first driving pattern is mainly characterized by driver's perception on careless driving, including characteristics related to risky driving behavior, such as the number of accidents that the driver has involved in, violation of speed limits as well as the number of harsh braking/ acceleration/ turning events occurred. The second driving behavior pattern is composed by drivers' perception on aggressive driving, as expressed by the number of traffic law infringements, the self-reported aggressive driving, as well as the part of the driver’s assessment that concerns braking and accelerating. The last pattern is based on driver's perception regarding cautious driving, which includes the self-reported frequency of harsh braking events headwayas kept by drivers, as assessed by an expert.

It should be noted that both factor analysis and linear regression, led to statistically clear and reasonably interpretable results, which are in full accordance with literature findings. The conclusions of the present study could contribute towards the improvement of driving behavior and road safety level if appropriate measures are taken. Finally, the results of this paper could be generalized under certain conditions, so that they can be used in studies with similar objective and data.

However, it is noted that a strong limitation of this research is the small driving sample used. It is therefore recommended to explore the same indicators and their correlation with drivers' road safety performance by using a larger sample. In addition, the correlation between reported driver's characteristics and observed through experiments of naturalistic driving or simulator could also be examined. Similar studies could also be conducted under different traffic conditions or road environments (urban roads, highways, high/low traffic volume, daytime/night-time, different weather conditions etc.), as well as for specific road user groups (young drivers, elderly, etc.). Finally, the use of advanced in-vehicle technologies recording data on driver's behavior, such as more accurate GPS, radar measuring vehicle's side position, reaction time, distance from the vehicle in front and cameras inside and outside the vehicle would contribute to the collection of more detailed data and the better exploration of driver's behavior and road safety performance.

References

Amado S., Arıkan E., Kaça G., Koyuncu M., Turkan B.N. (2014). "How accurately do drivers evaluate their own driving behavior? An on-road observational study." Accident Analysis & Prevention 63 (2014): 65-73.

Elvik, R., 2014. Rewarding safe and environmentally sustainable driving: systematic review of trials. Transportation Research Record: Journal of the Transportation Research Board 2465, 1-7.

European Commission, 2018, https://ec.europa.eu/commission/index_en

Kontogiannis T., Kossiavelou Z., Marmaras N. (2002). "Self-reports of aberrant behaviour on the roads: errors and violations in a sample of Greek drivers." Accident Analysis & Prevention 34.3 (2002): 381-399.

Moghaddam A.M., Ayati E. (2014). "Introducing a risk estimation index for drivers: A case of Iran." Safety science 62 (2014): 90-97.

Ott, B.R., Papandonatos G.D., Davis J.D., Barco P.B. (2012). "Naturalistic validation of an on-road driving test of older drivers." Human Factors: The Journal of the Human Factors and Ergonomics Society 54.4 (2012): 663-674.

Vardaki S., Yannis G. (2013). "Investigating the self-reported behavior of drivers and their attitudes to traffic violations." Journal of safety research 46 (2013): 1-11.

WHO (2015) Global status report on road safety 2015. World Health Organization, Geneva.

Yannis G., Weijeramrs W., Gitelman V., Vis M., Chaziris A., Papadimitriou E., Azevedo C.L. (2014). "Road safety performance indicators for the interurban road network." Accident Analysis & Prevention 60 (2013): 384-395.