JOURNAL OF TRAFFIC AND TRANSPORTATION ENGINEERING (ENGLISH EDITION) 2021; 8 (5): 725–734

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/jtte

Original Research Paper

Investigating the trip configured causal effect of distracted driving on aggressive driving behavior for e-hailing taxi drivers

Muhammad Sajjad Ansar [a], *Yongfeng Ma* [a,*], *Shuyan Chen* [a], *Kun Tang* [b], *Ziyu Zhang* [a]

[a] *School of Transportation, Southeast University, Nanjing 211189, China*

[b] *School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China*

HIGHLIGHTS

- Aggressive driving and distracted driving of taxi drivers are analyzed under three different trip categories.
- Structural equation model (SEM) verdicts that distracted driving has a remarkable causal effect on aggressive driving but in different manners for each group.
- Sensitive risky driving indicators are discussed for each group separately.

ARTICLE INFO

Article history:
Received 18 May 2020
Received in revised form
13 October 2020
Accepted 27 October 2020
Available online 15 December 2020

Keywords:
E-hailing taxi driver
Distracted driving
Aggressive driving
Naturalistic driving
SEM

ABSTRACT

Risky driving behavior of taxi drivers typically evaluated for full operation or sometimes sorted into occupied and empty running trips. In this paper, we simultaneously analyze aggressive driving and distracted driving of taxi drivers under three different trip categories. Trip origin is considered a transition from without ride-order to with ride-order travelling or from with ride-order to occupied travelling, and a destination as a transition from occupied to without ride-order travelling and vice versa. Distracted driving is characterized by driver interference, driver mobile use and some entertainment aspects, while specific harmful and risky actions are considered for aggressive driving. High-resolution and real-time kinematic parameters of taxis were recorded by the in-vehicle recorder VBOX for overall 562 trips. The distracted driving parameters and aggressive driving actions were monitored through python data collector web application that was specially programmed for this particular research. Besides dual dash cam (i.e., front and inside car camera), drivers' whole day driving history from their Chinese ride-hailing DiDi smart application was used to differentiate occupied trips, unoccupied trips with ride-order and unoccupied trips without ride-order. Structural equation modeling (SEM) is practiced in this paper to understand the influence of distracted driving indicators on aggressive driving behaviors. The multi-group model analysis of SEM indicated that handling distracted risky driving could control aggressive driving behavior up to 96% and 98% in

* *Corresponding author*. Tel.: +86 13912997321.
E-mail addresses: sajjad@seu.edu.cn (M.S. Ansar), mayf@seu.edu.cn (Y. Ma), chenshuyan@seu.edu.cn (S. Chen), kuntang_seu@seu.edu.cn (K. Tang), zhangziyu@seu.edu.cn (Z. Zhang).

Peer review under responsibility of Periodical Offices of Chang'an University.
https://doi.org/10.1016/j.jtte.2020.12.001

unoccupied without ride-order trips and unoccupied trips with ride-order respectively. The model has also identified the sensitive risky driving indicators for each group separately.

1. Introduction

The world health organization (WHO) reported traffic crashes as the main cause of road fatalities around the world; of which, at least, 90% were partially caused by human errors in 2016 (Uzoudu et al., 2019; Zhao et al., 2019). Numerous efforts have been made to control these type of errors; for instant, in China, the Traffic Management Department introduced penalty points based upon driving violations (Ma et al., 2019). However, the evaluation of human errors in terms of driving behavior still needs more comprehension as it has more range aside from violations. Among others, taxi drivers are statistically more likely to engage in unsafe driving practices (Sinclair and Imaniranzi, 2015). Moreover, the capacity utilization rate (the ratio between time/distance travelled with passenger and the total time/distance) for transportation network companies is 30%–50% higher than other taxi drivers (Nie, 2017). Thus, the aggressive and distracted driving behavior of such taxi drivers are more crucial to investigate.

There are different definitions of aggressive driving. The U.S. Department of Transportation named National Highway Traffic Safety Administration (NHTSA) defines aggressive driving as the driver driving behavior that endangers or is likely to endanger other road users or property (McCartt et al., 2001). While Tasca (2000) stated that aggressive driving corresponds with deliberate behaviors that may likely increase the risk of collision. However, this present study is based on a complete naturalistic driving experiment in which aggressive driving state is describing more of the driver driving behavior that could be affected by traffic conditions and the surrounding environment. Fountas et al. (2019) listed it as observed aggressive driving behavior.

E-hailing taxi drivers, as part of their driving task, provide transportation services using digital interfaces frequently, which causes distraction at a higher level relative to other drivers' scenarios. Several studies proved that distraction caused by driving behavior affects remarkably on driving style of taxi drivers (Stavrinos et al., 2010). On the contrary, studies have also demonstrated no statistically significant impact on driving performance during distracted driving (Papantoniou et al., 2019; Wang et al., 2019). The fact is that there are three general modalities to examine distraction (Zhang and Kaber, 2016): visual distraction, cognitive distraction and manual distraction. In this context, Fountas et al. (2019) identified internal and external sources of distractions. It will be quite uncertain to proclaim on distracted driving effects without keeping possible modalities and sources of distraction under consideration. So, this study has investigated general possible distractions.

Taxi driving behavior can be different when the taxi is vacant or occupied, or when the taxi driver is experienced or novice (Ding et al., 2015). Instead of complete day operation without trip categorization, several researches analyzed the classified trips to obtain a deep understanding of the driving behavior of taxi drivers (Liu et al., 2010). Similarly, this study aims to assess the causal effect of distracted driving on aggressive driving behavior in three different trip categories for e-hailing taxi drivers. Volatile effect assessment of underlying parameters under different trips was initially overlooked and particularly unheeded in the naturalistic field condition as well. To fill this gap, this paper attempted the SEM approach by considering naturalistically observed risky driving parameters. The model has also separately identified the risky driving indicators for each group.

The literature about underlying parameters has been discussed in the next section. After that, the experimental design of the naturalistic field test is explained. Following the methodological approach and results, discussion and conclusion are illustrated.

2. Literature review

In the past literature, three different methods were used to classify driver behavior. Driving style questionnaire generates subjective measurements in which every driver evaluates their driving behavior on a self-reported scale (Eboli et al., 2017). Secondly, driving simulators generates the artificial nature of the environment and controlled conditions of the experiment. However, it can lead to conclusions that are, sometimes, not easy to translate into real-world situations. Thirdly, capturing objective measurements in a real vehicle that provides the highest level of accuracy (Gonzalez et al., 2014; Ma et al., 2020). Although trips on highways and city roads are indeed very different in their characteristics and a model built for one might not work well on the other (Hong et al., 2014), but pure naturalistic data collection does not force drivers to drive predefined routes.

A huge amount of work has been done in several ways on aggressive driving and distracted driving of taxi drivers. There is a positive association between injury severity and aggressive driving (Wang et al., 2015). The study in South Africa reported that taxi drivers are statistically more likely to engage than other drivers to participate in aggressive driving practices such as improper passing (overtaking) and signal violation, etc. (Sinclair and Imaniranzi, 2015) The report on taxi crashes in Australia revealed that aggressive driving behavior such as running red lights, unsafe lane changes etc., could be the primary cause for high crash rate (Shi et al., 2014). Moreover, the naturalistic field experiment on research of driver aggression has taken horn honking as a measure of aggression (Turner et al., 1975).

Furthermore, driving performance reduces dramatically by distracted driving behavior. In addition to distraction due to phone use and driver entertainment aspects, the conduct of passenger interference behavior is only examinable in the occupied trips. Regardless of age differences, distractions such as texting, calling on the phone have a significantly negative impact on safety and traffic operations (Stavrinos et al., 2010). Besides the fact that hand-held use of a cellular phone is forbidden in the car while driving in many countries, this evidence is still witnessed in many studies. Traveling under a text-reading condition may increase the chances of lane deviation more than 300 times relative to normal condition traveling. Mintu Miah and Prevedouros Panos (2019) also supported the fact that reading a text while driving certainly lessens driving performance. The performance of secondary task about using cell phone has been investigated for distraction analysis which include dialing, talking, texting and reaching for phone (Klauer et al., 2014). The effect of dialing and reading text maybe even more detrimental because such distractions take the driver's eyes off the road, adversely affecting more than the effect of conversing. Mintu Miah and Prevedouros Panos (2019) observed that the vehicle slightly tends to veer the direction (left or right) in which the driver's head tilt to read the information from the mobile. Moreover, the questionnaire assessing distracted driving (QUADD) had considered the mobile phone usage aspects and the driver entertainment aspects as the core indicators for the assessment of distracted driving (Welburn et al., 2011). The driver entertainment aspects were measured for distraction analysis during travelling by items such as smoking, listening music, eating and radio listening adapted from the literature (Horvath et al., 2012; Klauer et al., 2014; Welburn et al., 2011).

While for investigation of behavioral interaction between drivers and passengers, passenger interference behavior is very important from the safety perspective due to the non-uniform effect of passengers on drivers (Regan and Mitsopoulos, 2001). Investigation of the effect of passengers on distracted driving behavior is relatively complex (Laberge et al., 2004). Findings from Amado and Ulupınar (2005) have revealed that the driver's conversation either with in-vehicle passengers or on cell-phone has a negative effect on peripheral detection and attention of the driver while driving.

Mintu Miah and Prevedouros Panos (2019) explained that distracted traveling has hundreds of times greater impact on aggressive driving relative to the normal condition of traveling. Several studies have listed distraction as a causal factor in aggressive driving behavior. In structural equation model (SEM), it is common to use the term latent variables to imply factors. This underlying causal effect of distracted driving construct on aggressive driving behavior construct has been investigated through SEM (Papantoniou et al., 2019). In the first measurement step of SEM, exploratory factor analysis (EFA) classify the factors on the basis of the interrelationships among observed indicators. For instance, EFA has determined two factors of driving behavior from eight observed indicators of drivers' characteristics (Eboli et al., 2017). Alternatively, in the case of theoretically oriented factors, confirmatory factor analysis (CFA) used to develop an acceptable measurement model. For example, CFA solely described the relationship of latent constructs such as in-vehicle distraction and risky driving behavior with the observed variables in the model prepared by Hassan and Abdel-Aty (2013).

Prevailing safety interventions unfolded many aspects regarding the impacts of aggressive driving and distracted driving separately. But the association between them and naturalistic assessment of these risky driving indicators has not been investigated in depth. Some studies analyze taxi driving behavior only in unoccupied trips (Ding et al., 2015), while some made a comparison of taxi drivers' behavior under-occupied and empty running trips (Liu et al., 2010). But the unoccupied trip with ride-order acceptance is disregarded throughout too.

3. Experimental design and data description

This study hired 15 professional drivers, each for a single whole day. The data collection and experimental design details are as follows.

3.1. *Participants*

Fifteen healthy professional participants (sex: male only, average age ± SD: 36.05 ± 7.9 years, range: 28–52 years, average driving experience ± SD: 10.2 ± 4.0 years, range: 4–22 years) from DiDi Kuaidi (i.e., a Chinese transportation company) were hired. It included local residents who have a profession of taxi driving. Sex and age effects were consistent (Rhodes and Pivik, 2011). So, it does not imply any assumption of discrepancies among the subjects selected. The use of thirty cases is recommended for examining more than eight latent variables in the data analysis by Zhao et al. (2019). In this research, we used thirteen valid cases to examine two latent variables. The reason for dismissing two cases is a showcase of difficulties and problems faced during the naturalistic data execution. Moreover, after completing the driving task, each driver was offered 250 Chinese Yuan as a reward. The detailed sample characteristics (M for average/mean and SD is depicting standard deviation) are mentioned in Table 1.

3.2. *Apparatus*

The experimental driving test system includes VBOX-IISX10 GPS Acquisition System manufactured by RACELOGIC and dual-lens driving recorder, shown Fig. 1(a). GPS based VBOX equipment is an in-vehicle sensing technology that provides continuous and objective data e.g., GPS coordinates, speed, lateral and longitudinal accelerations, heading angle, etc. (Hong et al., 2014), portrayed in Fig. 1(b). Each DiDi taxi was equipped with these two apparatuses for a single complete day after making appropriate communication with DiDi drivers. The drivers were asked to take rides and travel as usual.

Table 1 – Sample characteristics.

Sample size (N)	Driver age (years)		Driver experience (years)		Number of occupied trips (TWP)		Unoccupied trip with ride order (TWO)		Unoccupied trip without ride order (TWT)	
	M	SD	M	SD	M	SD	M	SD	M	SD
13	36.05	7.90	10.20	4.00	18.25	10.48	15.42	7.62	13.25	10.28

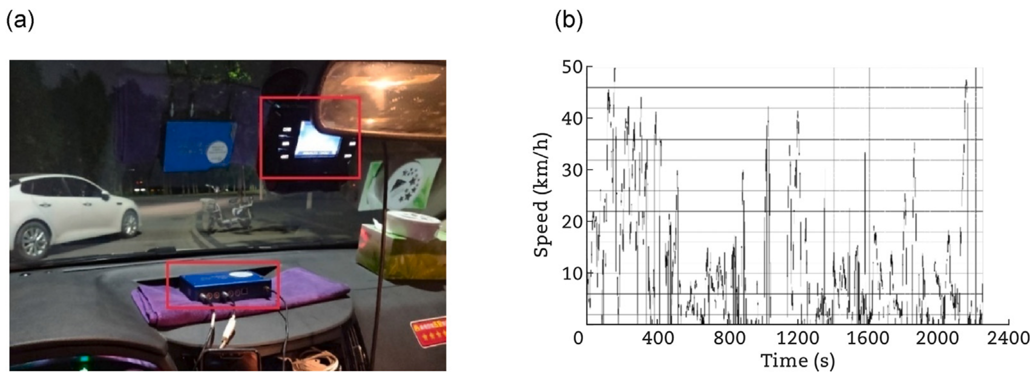

Fig. 1 – Components of driving style test experimental system. (a) Apparatuses. (b) Sample data in VBOX software.

3.3. *Data collection procedure*

It is stated in acceleration studies (Eboli et al., 2016), that driving dynamics deal with the properties of the vehicle during the vehicle motion. The specification of the vehicle condition constrained by DiDi service operation ensures the good properties of the vehicle. So, both types of equipment were mounted in these vehicles taking sunny weather into considerations. In the next stage, we let them go while clarifying the cause and intent of this activity to ensure their routine driving. Every equipment was removed at night time when the driver finished the job. The details of rides from drivers' mobile DiDi application were also captured.

3.4. *Data pre-processing*

Altogether, there could be four possible conditions for taxi drivers; traveling with a passenger, traveling without passenger with ride-order acceptance, traveling without passenger without ride-order acceptance and static (or non-traveling) condition. Trips with passengers (TWP) elapsed from when the passenger gets on to when the passenger got off period. While trips without passenger with order (TWO) start when the driver accepts the ride order and ended when he reached the boarding station. Trips without passenger without order (TWT) were quickly identified from the database through the simple linear equation in which we have a single unknown (i.e., TWT) and all other driving and static states are known. The trips include TWP, TWO and TWT were recognized using a dual cam driving recorder and from the trip details of the whole day that were recorded from the DiDi drivers' mobile application. After that, video data was processed through a python-built data collector web application. The data from VBOX was analyzed through VBOX Tools (i.e., software by RACELOGIC). Moreover, the static condition in the form of waiting for passenger at boarding station or sleeping time of driver does not lie in between any of these underlying three groups (i.e., TWP, TWT and TWO). The events happen during signal waiting time were also excluded.

Performance indicators were linked with distracted driving and aggressive driving constructs. Driver entertainment aspects such as smoking, listening to music, eating and driver cell phone usage patterns such as reading text, dialing, recording voice and talking on a phone call during traveling are undertaken for distracted driving measurement in this study. Furthermore, distracted driving construct in this current research also focuses merely on certain effects of passengers such as making a loud phone call, navigation assistance, asking questions and making requests. Except for passengers' loud phone calls, all other aspects are related to communication between passenger and driver. Consequently, parameters of aggressive driving behavior were monitored by the visual and computational way. These parameters in this study include overtaking (OVT), yellow double-line violation (LC), signal violation (SVL), illegal honking (HNK) and external points (EP) through acceleration-based driving style diagram.

Table 2 represents the descriptive statistics and scaling patterns of each variable. We count lane change (LC) as an event when a driver crossed the double-yellow solid line for a considerable elapsed period (>3 s). On behalf of quantity size, a total of three indicators were treated as dichotomous variables, including SVL, HNK and OVT, which were rarely identified events. Except for dichotomous indicators, every event recognized during traveling (not static) condition was collected in the form of frequency. If any variable elapsed about more than half of the particular trip's traveled time, we gave that indicator the highest score. For example, three-times smoking in one trip has the same scale (i.e., 3) as smoking once that lasts at least half-time span of that trip. Similarly, an illegal honking event defines horns in the

Table 2 – Details of indicators and descriptive statistics.

Latent constructs		Observable indicator	Abb.	Max	Min	TWP		TWO		TWT	
						M	SD	M	SD	M	SD
Aggressive driving (AD)	–	Signal violation	SVL	2	1	1.04	0.19	1.03	0.16	1.03	0.16
		Honk	HNK	2	1	1.09	0.29	1.07	0.25	1.03	0.16
		Overtaking	OVT	2	1	1.21	0.41	1.06	0.25	1.09	0.28
		Lane change	LC	4	1	1.42	0.75	1.26	0.58	1.20	0.47
		External points	EP	4	1	2.28	1.02	1.98	1.13	1.90	1.12
Distracted driving (DD)	Driver's entertainment	Eating	ET	3	1	1.07	0.30	1.06	0.29	1.09	0.33
		Smoking	SM	3	1	1.03	0.22	1.24	0.61	1.24	0.61
		Listen to music	LM	3	1	1.22	0.61	1.21	0.60	1.26	0.67
		Radio listening	RL	3	1	1.15	0.50	1.13	0.49	1.11	0.45
	Driver's mobile phone usage	Phone call	PC	4	1	1.13	0.41	1.30	0.60	1.38	0.78
		Typing	TP	4	1	1.26	0.60	1.36	0.71	1.61	1.04
		View SMS	VS	4	1	1.54	0.90	1.48	0.80	1.78	0.96
		Send voice	SV	4	1	1.10	0.41	1.10	0.45	1.24	0.64
	Passenger interference behavior	Loud phone call	LPC	4	1	1.41	0.88				
		Navigation assistance	NA	3	1	1.27	0.53				
		Asking question	AQ	4	1	1.55	0.98				
		Making request	MR	4	1	1.10	0.34				

honking-prohibited zone. The non-honking area has been recognized from the road signs via video recorder as it is said that Nanjing is one of the well-planned cities in which roads are properly marked and signed (Fei, 2016). The VS indicator of mobile phone usage is only nominated with the condition of reaching towards the phone.

Two latent constructs were built by relevant indicators in total. The above-mentioned events that happened in each traveling state (i.e., TWT, TWO and TWP) got monitored and processed through the common procedure. The overall database contains more than 500 trips, their associated events second-by-second and conformably, ten-Hertz GPS data. Data Analysis was carried out using complete data for all 215 trips for TWP, 187 TWO trips, and 160 TWT trips. The discussion on acceleration-based external points (EP) has carried out in the methodology part. However, the complete details of indicators and descriptive statistics are illustrated in Table 2.

4. Methodology

This current research collects possible visual and computational events for aggressive driving and distracted driving alike. In this section, we have discussed the only computational parameter for aggressive driving (i.e., EP) and the SEM model.

4.1. *Definition of EP indicator*

The EP indicator for aggressive driving (AD) latent construct computed through the following *g*-*g* diagram. For naturalistic driving data, speed is not a good way to distinguish aggressive and calm drivers (Hong et al., 2014). Among other vehicle dynamic parameters, vehicle's acceleration is a rich source of driver behavior assessment (Eboli et al., 2016; Paefgen et al., 2012; Vaiana et al., 2014). Moreover, volatility in instantaneous driving decisions can be quantified by variability in vehicular movement (Wang et al., 2015). So, a way for characterizing driver behavior is the *g*-*g* diagram. The acceleration threshold in the underlying *g*-*g* diagram is related to the edges of the friction circle which corresponds with tire characteristics and road surface conditions. The experience of driver influenced a lot on the lateral and longitudinal variations besides road surface condition and tire characteristics.

The *g*-*g* diagram, shown in Fig. 2, renowned as driving style diagram (DSD) is created by overlapping the driving experience area and the friction circle. Vaiana et al. (2014) have classified safe and unsafe behavior based on external points of DSD. Moreover, the considerations of specific road type, road surface condition, and test route are crucial for the reliable and well-understood measurements of acceleration threshold in the categorization of safe and unsafe driver behavior. Although, some researchers take exemplary thresholds (Paefgen et al., 2012). As like, Eboli et al. (2016) considered the longitudinal threshold based on literature studies. In this sense, for this current study, considerations of external points of DSD in the evaluation of risky driving events will be reasonable. Fig. 2 is depicting the EPs of single day data for three different trip categories. The safe and unsafe behaviors were classified on the basis of external points of this underlying *g*-*g* diagram. We divide the number of DSD external points of every single trip to its traveled time and get the score. Then the scale is generated according to mean and standard deviation of scores of all trips in a specific category generated by a particular driver. This scaling approach shown in Table 3, was followed by Wang et al. (2015). It is helpful to rectify the effect of instantaneous changes in surrounding circumstances.

4.2. *Structural equation model (SEM) approach*

SEM methodology was adopted in various fields of research. Eboli et al. (2017) used SEM to explore the hypothesized relationship between driving style and a certain driver's conditions. While Zhao et al. (2019) chosen SEM to analyze

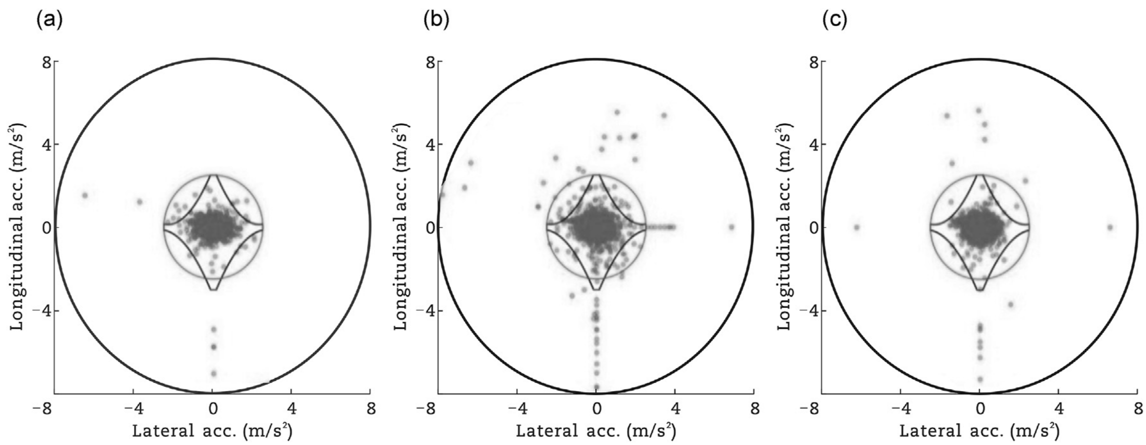

Fig. 2 – External points (EPs) through *g*-*g* diagram. (a) Day-1 TWO *g*-*g* diagram. (b) Day-1 TWP *g*-*g* diagram. (c) Day-1 TWT *g*-*g* diagram.

Table 3 – Scale's definition for external points (EPs).

Indicator value	1	2	3	4
External point score for configured trips associated with a particular driver	< mean−std. dev	> mean−std. dev, ≤ mean	> mean, ≤ mean + std.dev	> mean + std. dev

the impact of different factors on illegal driving action and driver performance. The general SEM can be decomposed into the measurement model and the structural model. Confirmatory factor analysis (CFA) is used in this paper for testing structure between observed indicators and latent variables. Exploratory factor analysis (EFA) is somewhat like an arbitrary model that is free from any theoretical hypothesis. In contrast, CFA is restricted to a clear hypothesis about the factor structure. CFA model specifies the pattern by which each indicator loads on a particular factor. Whereas the structural model defines the causal relationship among unobserved variables (i.e., exogenous latent variables and endogenous latent variables). In this article, the structural model is composed of two constructs that include distracted driving and aggressive driving. Distracted driving parameters belong to passenger interference behavior (PIB) and driver's behavioral characteristics such as driver's mobile phone usage and driver's entertainment aspects. The latent construct of aggressive driving was drawn based on harmful and risky actions. Distracted driving was treated as an exogenous latent construct, while aggressive driving was considered as an endogenous latent construct. Each latent construct was linked with more than three observed indicators in every model. Single-headed arrows are regression coefficients that represent factor loadings of observed variables on the corresponding latent construct. The analysis of moment structures (AMOS) 21 module IBM SPSS is used in this paper to create the SEM. Moreover, our analysis has been made on the adequate and acceptable fit based on indices such as Q, GFI, PGFI, RMSR, and RMSEA of the SEM model shown in Table 4.

Table 4 – SEM model fit indices.

Fit index	TWP	TWO	TWT
Chi-square (CMIN)	24.24	13.44	28.60
DF	24	13	25
Probability level	0.45	0.41	0.28
CMIN/DF (Q-value)	1.01	1.03	1.14
RMR	0.02	0.02	0.03
CFI	0.989	0.979	0.939
GFI	0.975	0.979	0.964
AGFI	0.954	0.956	0.936
RMSEA	0.007	0.013	0.030
NFI	0.578	0.682	0.698

5. Results

5.1. *Test for normality*

Following descriptive statistics Table 2, the Kolmogorov-Simonov method is practiced to test the normality in the very first step on these three groups of data samples in which the normality hypothesis was rejected by Skewness and Kurtosis *z*-values and the Shapiro-Wilk's *p*-value. No single indicator has Skewness and Kurtosis value near to 0. The significant value of both Kolmogorov-Simonov and Shapiro-Wilk tests is less than 0.001 for each variable. So, we presume there is a statistically significant difference between each variable and normal distribution. Thus, a non-parametric test named Kruskal-Wallis test was performed to investigate group differences. This test is also famous as one-way ANOVA on ranks. Observed variables that have at least 0.03 asymptotic significance values or minimum were chosen.

Table 5 – KW test for group differences.

	SV	LC	HNK	OVT	ET	SM	LM	RL	PC	TP	VS	SV	EP
Kruskal-Wallis H	0.591	11.778	6.955	23.494	1.065	22.086	0.640	1.295	15.728	10.278	12.026	9.688	18.670
DF	2	2	2	2	2	2	2	2	2	2	2	2	2
p-value	0.744	0.003	0.031	0	0.587	0	0.726	0.523	0	0.006	0.002	0.008	0

Table 5 demonstrates a total of 13 parameters where 9 indicators display highly statistically significant differences among these three groups. Thus, for multiple group analysis, only nine comparable parameters were in our sight for the establishment of a measurement model. Among all selected indicators, "HNK" has the minimum KW score (i.e., 6.955). Except for one all other chosen variables were reported having significant *p*-values lower than 0.8%. Smoking, as a driver entertainment indicator, is the only comparable parameter identified for distracted driving.

The most often used estimator of parameters is the maximum likelihood (ML) method. ML is preferable for continuous data and based on the assumption of multivariate normality. But underlying observed variables fail to exhibit multivariate normality. However, it is good news that a recent advancement argues that linear structural relation (LISREL) or so-called SEM does not need to be linear anymore as the possibility of SEM extended well beyond the original LISREL program (Hasman, 2015). We perform bootstrapping to validate and assess the stability of the modeling results under the ML approach.

Efron (1979) introduced the bootstrap technique in which any nonparametric estimator can do a good estimation job without any prior distribution restriction. Thus, this underlying model has used bootstrapping to identify better predictors among the three groups.

5.2. *SEM estimation*

The two-factor model is specified with nine indicators loadings; four aggressive driving indicators and five indicators for distracted driving. The aim was to categorize and identify the powerful indicators in each different group. However, nine variables that have significant differences among groups, still need justification of significance in each particular group too. This fact was revealed when we perform CFA. So, confirmatory factor analysis is performed for the validity of the measurement model. The path diagram shown in Fig. 3 is depicting the directional effect of distracted driving on aggressive driving behavior.

According to the outcomes, presented in Table 6 VS is the relatively most significant indicator that affects DD construct in each group. Overtaking (OVT) affects AD highly in TWP and TWT groups. In the TWO group, TP, HNK, and EP are statistically insignificant. While HNK and EP indicators were found insignificant (*p*-value $>$ 0.05) in the TWT category. Beyond the scope of multiple group analysis, we added passenger interference related indicators in the TWP SEM model. Only five indicators are statistically significant up to *p*-value $\leq$ 5% in the TWP group. Among all other distracted indicators in TWP, conversation starts from a passenger in the form of making request (MR) weighing

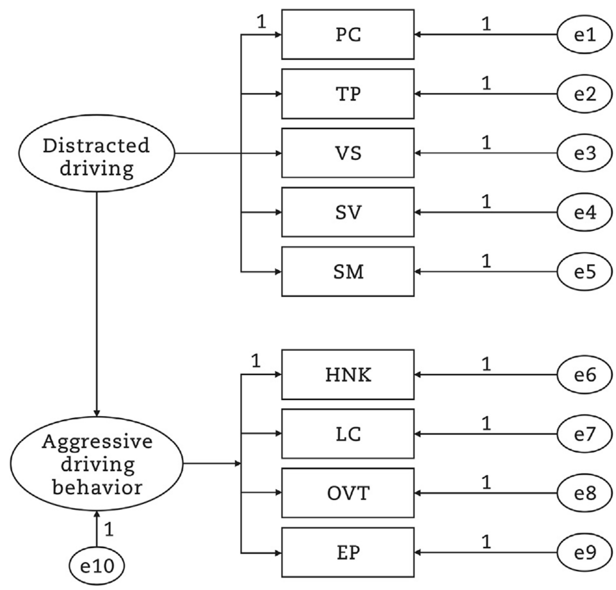

Fig. 3 – The path diagram of the estimated two-factor SEM.

higher (factor load is 0.46) on DD, which supports past results (Laberge et al., 2004). Moreover, DD has an overall highly significant impact on AD in TWT and TWO groups. The *p*-value mentioned in the table comes after fixing a 95% confidence interval for bootstrap multivariate analysis. The observed variables that had negative error variance were deleted from this model. *B* is stated in the replacement of unstandardized regression weights while Beta is presenting standardized regression weight.

6. Discussion

The evaluation of unoccupied trips describes the overall same causal effect of DD on AD in both TWT and TWO groups. Although we found volatile weightage of each variable in these groups. For instance, talking on the phone (TP) has 37% significant weightage on DD in TWT but in TWO it has 8% insignificant weightage. In the TWT group, contrary to TWO, VS and SV are more reactional than PC. It can be interpreted that handling SV, SM, and VS could control overtaking action while traveling without ride-order. Distracted driving significantly affecting the aggressive driving about 96% in the TWT group and 98% in TWO. This causal impact is 2% greater in the case of TWO in comparison with TWT, which justify that aggressive driving is also contributed by time pressure in TWO (McCartt et al., 1998).

The insignificant structural model in TWP demonstrates that the indicators used in TWT and TWO, could not be able to find the effect of DD on AD. Occupied trips may need

Table 6 – SEM results.

Observed variable	←	Latent variable	TWP			TWO			TWT		
			B	Beta	*p*-value	*B*	Beta	*p*-value	*B*	Beta	*p*-value
PC	←	DD	–	–	–	0.94	0.55	0.01	0.51	0.34	0.07
SM	←	DD	−0.11	−0.15	0.01	–	–	–	0.51	0.44	0.01
VS	←	DD	1.00	0.32	0.05	1.00	0.44	0.02	1.00	0.55	0.00
SV	←	DD	−0.76	−0.05	0.15	0.41	0.32	0.02	0.56	0.46	0.03
TP	←	DD	–	–	–	0.17	0.08	0.37	0.72	0.37	0.03
HNK	←	AD	−0.25	−0.25	0.14	0.39	0.21	0.12	0.01	0.01	0.79
LC	←	AD	–	–	–	1.00	0.24	0.01	1.00	0.25	0.08
EP	←	AD	1.00	0.28	0.12	0.65	0.08	0.31	0.10	0.01	0.88
OVT	←	AD	−0.67	−0.46	0.03	–	–	–	1.00	0.41	0.01
LPC	←	DD	0.97	0.32	0.05						
MR	←	DD	0.55	0.46	0.01						
AQ	←	DD	−0.23	−0.07	0.49						
Latent variable	←	Latent variable									
AD	←	DD	−0.15	−0.15	0.56	0.38	0.98	0.00	0.21	0.96	0.01

Note: "←" means directional effect.

further observed indicators to understand the effect of DD on AD. The insignificant results of underlying variables showing less integrity for getting these proposed objectives. Although, we can understand some confident variables for AD and DD in occupied trips. For example, overtaking (OVT) and making a request (MR) both weighting about 46% on AD and DD in TWP.

The approach for acceleration-based external points through driving style diagram did not work well throughout. It seems that such evaluation should be road-specific and require to fix spatiotemporal conditions. Even it can be assumed that every insignificant parameter has spatiotemporal concerns or there could be some other reasons. For instance, cell phone usage events encountered lower in numbers due to the usage of proper hangers for their phone and Bluetooth devices by some genius drivers. Similarly, trips with passengers (TWP) were relatively longer than TWO and TWT. Although, we took the frequency of each indicator for the establishment of scale. But still, this fact might have weakened the underlying comparisons among groups.

Although, previous literature has pointed out the extent of distraction and aggressiveness of drivers contributing in traffic accidents. However, the analysis of the causal effect of distracted driving on aggressive driving would be much crucial to ensure traffic safety. Aggressive driving which is the pre-collision state (Tasca, 2000), could be well restricted by identifying the highly influencing factors of distracted driving. The authors believe that the study outcomes would have similar importance like evaluation and analysis of traffic conflicts besides traffic crashes.

7. Conclusions

This study describes the causal effect of distracted driving on aggressive driving in three different trip categories. Although several studies have explored the effect of factors like personality trait, attitude, and driver's characteristics on aggressive and distracted driving behavior. But few studies have focused directly on the relationship of aggressive and distracted driving behavior of taxi drivers. The primary difference this study makes from past studies is investigating the causal effect of aggressive and distracted driving behavior in a specific trip framework. These trip categories include occupied trips (TWP), unoccupied trips with ride-order (TWO) and unoccupied trips without ride-order (TWT). Distracted driving and aggressive driving have been defined based on relevant indicators investigated in naturalistic driving experiment. All indicators were assessed on the scale except for three rarely identified indicators. Structural equation modeling is practiced to understand the causal effect. The outcomes of descriptive statistics and the structural model argued that a comparative study based upon descriptive statistics would not be reliable without consideration of a comprehensive statistical approach like SEM.

Structural model results verdicts that distracted driving (DD) has a remarkable causal effect on aggressive driving (AD) but in different manners for each group. It highlights two main implications. First is the volatile effect of each indicator in different groups that were initially disregarded. The second one is regarding the overall same causal response in TWT and TWO. It means that unoccupied trips are easily comparable. More than 95% causal effect of DD on AD verdicts that handling DD parameters could control AD which has a direct influence on road accidents. It can be argued that If we control DD accurately, it will alleviate AD automatically. It recommends that DD parameters should provide special considerations in advanced driving-assistance system (ADAS) and education of drivers.

Indeed, simultaneously working on a lot of indicators under different groups is not so easy. That is the reason we experienced non-significance issues in a bunch of parameters. Although, the tests for reliability and validity of indicators before training SEM model would be helpful to specify particular indicators and it could enhance the overall significance of the model. Future research could be made on factors affecting AD and DD separately in each group to validate the similarity of each indicator's effect, as such group classification is new according to the author's knowledge. However, this way certainly has very limitations but still it opens a novel approach to control aggressive driving by the factors which need preliminary fixation. Lastly, for the analysis of nonparametric indicators in multiple-group analysis, our research advocates the KW test to find group differences before the implementation of SEM. Conclusively, we suggest the abovementioned considerations to future researchers before analyzing driver behavior.

Conflict of interest

The authors do not have any conflict of interest with other entities or researchers.

Acknowledgments

This research is supported by the National Key R&D Program of China (2018YFB1601600).

REFERENCES

Amado, S., Ulupınar, P., 2005. The effects of conversation on attention and peripheral detection: is talking with a passenger and talking on the cell phone different? Transportation Research Part F: Traffic Psychology and Behaviour 8 (6), 383–395.

Ding, L., Fan, H., Meng, L., 2015. Understanding taxi driving behaviors from movement data. AGILE 17, 219–234.

Eboli, L., Mazzulla, G., Pungillo, G., 2016. Combining speed and acceleration to define car users' safe or unsafe driving behaviour. Transportation Research Part C: Emerging Technologies 68, 113–125.

Eboli, L., Mazzulla, G., Pungillo, G., 2017. How drivers' characteristics can affect driving style. Transportation Research Procedia 27, 945–952.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics 7 (1), 26.

Fei, S., 2016. Parking versus public transport subsidies: case study of Nanjing, China. Transportation Letters 8 (2), 90–97.

Fountas, G., Pantangi, S.S., Hulme, K.F., et al., 2019. The effects of driver fatigue, gender, and distracted driving on perceived and observed aggressive driving behavior: a correlated grouped random parameters bivariate probit approach. Analytic Methods in Accident Research 22, 100091.

Gonzalez, R.A.B., Wilby, M.R., Diaz, J.J.V., et al., 2014. Modeling and detecting aggressiveness from driving signals. IEEE Transactions on Intelligent Transportation System 15 (4), 1419–1428.

Hassan, H.M., Abdel-Aty, M.A., 2013. Exploring the safety implications of young drivers' behavior, attitudes and perceptions. Accident Analysis & Prevention 50, 361–370.

Hong, J.-H., Margines, B., Dey, A.K., 2014. A smartphone-based sensing platform to model aggressive driving behaviors. In: The SIGCHI Conference on Human Factors in Computing Systems, Toronto, 2014.

Horvath, C., Lewis, I., Watson, B., 2012. Peer passenger identity and passenger pressure on young drivers' speeding intentions. Transportation Research Part F: Traffic Psychology and Behaviour 15 (1), 52–64.

Hasman, A., 2015. An introduction in structural equation modeling. Studies in Health Technology and Informatics 213, 3–6.

Klauer, S.G., Guo, F., Simons-Morton, B.G., et al., 2014. Distracted driving and risk of road crashes among novice and experienced drivers. New England Journal of Medicine 370, 54–59.

Laberge, J., Scialfa, C., White, C., et al., 2004. Effects of passenger and cellular phone conversations on driver distraction. Transportation Research Record 1899, 109–116.

Liu, L., Andris, C., Ratti, C., 2010. Uncovering cabdrivers' behavior patterns from their digital traces. Computers, Environment and Urban Systems 34 (6), 541–548.

Ma, Y., Tang, K., Chen, S., et al., 2020. On-line aggressive driving identification based on in-vehicle kinematic parameters under naturalistic driving conditions. Transportation Research Part C: Emerging Technologies 114, 554–571.

Ma, Y., Yu, Y., Chen, S., et al., 2019. 2019 Factors affecting driver's perceptual bias of aggressive driving behavior in China. In: Transportation Research Board 98th Annual Meeting, Washington DC, 2019.

McCartt, A.T., Hammer, M.C., Fuller, S.Z., 1998. 1998 Research on the scope and nature of fatigued driving among truck drivers in New York State. In: Traffic Safety on Two Continents, Malmo, 1998.

McCartt, A.T., Leaf, W.A., Witkowski, T.L., et al., 2001. Evaluation of the Aggression Suppression Program, Milwaukee, Wisconsin. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington DC.

Mintu Miah, M., Prevedouros Panos, D., 2019. 2019 Effects of reading text while driving: analysis of 200 Honolulu taxi drivers on a simulator. In: The Transportation Research Board (TRB) 98th Annual Meeting, Washington DC, 2019.

Nie, Y., 2017. How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China. Transportation Research Part C: Emerging Technologies 79, 242–256.

Paefgen, J., Kehr, F., Zhai, Y., et al., 2012. 2012 Driving behavior analysis with smartphones: insights from a controlled field study. In: The 11th International Conference on Mobile and Ubiquitous Multimedia, Ulm, 2012.

Papantoniou, P., Yannis, G., Christofa, E., 2019. Which factors lead to driving errors? A structural equation model analysis through a driving simulator experiment. IATSS Research 43 (1), 44–50.

Regan, M.A., Mitsopoulos, E., 2001. Understanding Passenger Influences on Driver Behaviour: Implications for Road Safety and Recommendations for Countermeasure Development. #180-2001. Monash University, Melbourne.

Rhodes, N., Pivik, K., 2011. Age and gender differences in risky driving: the roles of positive affect and risk perception. Accident Analysis & Prevention 43 (3), 923–931.

Shi, J., Tao, L., Li, X., et al., 2014. A survey of taxi drivers' aberrant driving behavior in Beijing. Journal of Transportation Safety & Security 6 (1), 34–43.

Sinclair, M., Imaniranzi, E., 2015. Aggressive driving behaviour: the case of minibus taxi drivers in Cape Town, South Africa. In: 34th Southern African Transport Conference, Cape Town, 2015.

Stavrinos, D., Jones, J.L., Garner, A.A., et al., 2010. Impact of distracted driving on safety and traffic flow. Accident Analysis & Prevention 61, 63–70.

Tasca, L., 2000. A Review of the Literature on Aggressive Driving Research. Ontario Advisory Group on Safe Driving Secretariat, Road User Safety Branch. Ontario Ministry of Transportation, Toronto.

Turner, C.W., Layton, J.F., Simons, L.S., 1975. Naturalistic studies of aggressive behavior: aggressive stimuli, victim visibility, and horn honking. Journal of Personality and Social Psychology 31 (6), 1098–1107.

Uzoudu, C., Jamson, S., Lai, F., 2019. Investigating unsafe behaviours in traffic conflict situations: an observational study in Nigeria. Journal of Traffic and Transportation Engineering (English Edition) 6 (5), 482–492.
Vaiana, R., Iuele, T., Astarita, V., et al., 2014. Driving behavior and traffic safety: an acceleration-based safety evaluation procedure for smartphones. Modern Applied Science 8, 88–96.
Wang, C., Li, Z., Fu, R., et al., 2019. What is the difference in driver's lateral control ability during naturalistic distracted driving and normal driving? A case study on a real highway. Accident Analysis & Prevention 125, 98–105.
Wang, X., Khattak, A.J., Liu, J., et al., 2015. What is the level of volatility in instantaneous driving decisions? Transportation Research Part C: Emerging Technologies 58 (Part B), 413–427.
Welburn, S.C., Garner, A.A., Franklin, C.A., et al., 2011. Psychometric Validity of the Questionnaire Assessing Distracted Driving (QUADD). Translational Research for Injury Prevention Laboratory, University of Alabama at Birmingham (UAB), Birmingham.
Zhang, Y., Kaber, D., 2016. Evaluation of strategies for integrated classification of visual-manual and cognitive distractions in driving. Human factors 58, 944–958.
Zhao, X., Xu, W., Ma, J., et al., 2019. An analysis of the relationship between driver characteristics and driving safety using structural equation models. Transportation Research Part F: Traffic Psychology and Behaviour 62, 529–545.

Muhammad Sajjad Ansar received the B.S. degree in transportation engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2018. Chinese Scholarship Council (CSC) nominated him as a master's degree scholar on the full scholarship for the period fall-2018 to 2020. He is currently a graduate research assistant with the Jiangsu Key Laboratory of Urban ITS and the Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transporation, Southeast University. His research interests include traffic safety, machine learning, data mining, and intelligent transportation system.

Yongfeng Ma received the PhD degree in transportation engineering at School of Transportation, Southeast University in China in 2007. He has been working at the position of visiting scholar in the School of Civil Engineering, Center for Road Safety, Purdue University, from August 2014 to September 2015. He has participated in a number of Chinese projects related to road safety, and has published over 50 articles in scientific peer-reviewed journals, books, and conferences. His current research interests include traffic safety, driving behavior analysis, intelligent transportation systemsand data mining.

Shuyan Chen received the B.E. and M.E. degrees in computer science and engineering from Southeast University, Nanjing, China, in 1989 and 1992, and the PhD degree in system engineering from Southeast University, Nanjing, China, in 2003. She is currently a professor in the School of Transportation, Southeast University, Nanjing, China. She has received two National Science and Technology Progress Awards of China. Her research interests include traffic incident detection, traffic signal control, travel time prediction, route guidance, data mining, machine learning and its applications in intelligent transportation systems. She has authored or co-authored more than 90 articles in international journals and conferences.

Kun Tang received the B.S. degree in electronic information science and technology from Hebei Polytechnic University, Tangshan, China, in 2011, the M.S. degree in electrical and communication engineering from Jiangsu University of Science and Technology, Zhenjiang, China, in 2014, the PhD degree in transportation engineering at School of Transportation, Southeast University, Nanjing, China, in 2019. He is currently an assistant professor in School of Automation, Nanjing University of Science and Technology, Nanjing, China. His research interests include intelligent transportation systems, machine learning, traffic data mining, driving behavior and traffic safety. He has authored or co-authored more than 10 articles in international journals and conferences.

Ziyu Zhang received the B.S. degree in traffic and transportation engineering from Central South University, Changsha, China, in 2017. He is currently a research assistant in the Jiangsu Key Laboratory of Urban ITS and the Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University. His research interests include intelligent transportation systems, driving behavior, traffic data mining, machine learning, and deep learning.