Proceedings of the 2nd World Congress on Civil, Structural, and Environmental Engineering (CSEE'17)
Barcelona, Spain – April 2 – 4, 2017
Paper No. ICTE 117
ISSN: 2371-5294
DOI: 10.11159/icte17.117

Predicting Speed-Related Traffic Violations on Rural Highways

Mohamed Shawky[1,2], Iyad Sahnoon[1,3], Ahmed Al-Zaidy[1]
[1]Abu Dhabi Traffic Police
Abu Dhabi, UAE
m_shawky132@hotmail.com; al_zaidy@hotmail.com
[2]Ain Shams University
Cairo, Egypt
[3]Tatweer, Abu Dhabi
UAE
isahnoon@gmail.com

Abstract - Speeding is considered as the most frequent violation type among the other traffic offences. This fact may occur due to the wide prevalence of speed enforcement among the road network. This paper aims to investigate criteria for selecting the best locations of speed cameras on rural highways. Data were collected at 76 sites of existing fixed speed cameras on rural highways of the Emirate of Abu Dhabi during a period of three months. The historical speeding violations, traffic information, aspects of the speed camera, and road characteristics were collected at each site. The statistics show that about 8,110,985 speeding violation tickets were issued between the years from 2008 till 2015 which represents around 80% of total violations. A model to predict the frequency of speeding violations was developed by using negative binomial regression approach. About fifteen independent variables/predictors were examined and are seen to significantly affect the occurrence of speeding violations frequency at a confidence level of 95%. These variables include traffic-related variables (i.e. traffic volume, average speed, and percentage of trucks); site and camera's characteristics related variables (i.e. posted speed limit, speed margin, direction of enforcement camera, straight road segment, existence of speed change zone); and types of day. These findings can be used as selection criteria to find the best locations for installing speed cameras in the future enforcement programs.

***Keywords*:** Speeding behaviour, speeding violation prediction, speed camera location, negative binomial regression, speed camera

1. Introduction

Speeding is considered as the most hazardous driving behaviour that have a strong correlation with the crash severity. Automated speeding devices such as speed cameras are the most effective tools for enforcing drivers to comply with the posted speed limits on roads. In most of countries, traffic police department is the responsible authority speed cameras allocations on roads. However, the traffic police department is facing a challenge for defining the best locations for installing the speed camera. Usually, these locations are selected based on the traffic crashes history only. This methodology may not lead to the best results in improving driver's speeding behaviour and traffic safety. So, integrating the historical speeding violations, traffic and road characteristics in selection approach of speed camera locations is essential for more effective speeding enforcement system.

The main objective of this paper is to develop criteria for selecting the best locations of speed cameras on rural highways by investigating the speeding behaviour in terms of speeding violations and find the impact of the speed cameras and road characteristics, traffic information, and days of the week on them. A model to predict the frequency of speeding violations was developed. The historical records of speeding violations at certain sites of speed cameras on rural highways in Abu Dhabi (AD) emirate, the capital of United Arab Emirate, during a period from August till October 2015 (i.e. three months) were employed as a case study in this study. Negative binomial regression approach was applied as the best prediction model used to estimate the frequency of speeding violations. The model investigated variables related to traffic information, site characteristics, and day of week as predictors affecting the occurrence of speed violations.

2. Literature Review

Despite speeding enforcement is one of the strategies that aims to improve road safety, there are very limited studies that tackled the traffic violation prediction compared to road safety and crash analysis studies in the literature. A study aimed to develop red-light running (RLR) violation prediction model adopted in random forest machine-learning technique [1]. Two data sets were employed here: observational data and driver simulator data. Both datasets included some common variables such as time to intersection (TTI), distance to intersection (DTI), and speed at the onset of the yellow signal phase. The observational data have vehicle information for different time frames. On the other hand, driver's demographic characteristics were obtained from the simulator data since the observational ones are anonymous. The results showed that TTI, DTI, deceleration parameter, and speed at the onset of the yellow signal phase were among the most important factors identified from models using observational and simulator datasets.

Other studies attempted to investigate the red-light violators and conditions in which they are likely to violate [2-3] and others were taking the site conditions in to consideration as well [4]. In addition, some researches attempted to estimate the frequency of red-light violators at signalized intersections. For example, Bonneson and Son [2] developed a regression model considering factors of traffic flow rate, cycle length, and yellow time duration. Another study was done by Zhang et al. [5] where they developed a probabilistic model to predict RLR violations taking into consideration minimizing false alarm rates and missing errors.

Rosenbloom and Perlman [6] examined the drivers' tendency to commit traffic violations when they are accompanied with passengers or not. The employed data in this study include gender, age, and number of passengers. Four dependent variables were studied: wearing seatbelt, signalling, using hand-held cellular phone and keeping the distance from the vehicle in front. The results showed that both males and females, old and young, a greater proportion of drivers who were alone committed traffic violations compared to drivers who were not alone.

De Winter [7] investigated the tendency of novice license drivers on committing future violations using self-reported surveys. Respondents with a higher violations frequency, higher speed, and lower number of errors in the simulator reported completing fewer hours of on-road lessons before their first on-road driving test. The results add to the literature on the predictive validity of driving simulators, and can be used to identify at-risk drivers early in a driver-training program. Ayvaşık et al. [8] examined the effect of sensation seeking and attention on traffic violations and driver's errors. Participants were asked to respond to computerized tests (called monotonous and selective attention tests) as well as driver behaviour questioners. The results of the sensation seeking and attention tests were grouped based on the responses and analysis of variables was then conducted. The analysis outputs indicated that drivers who have higher number of traffic violations and errors are those who got high monotonous and selective attention results. Also, drivers who reported lower levels of safety skills feel overconfidence in their skills and underestimation of the hazards in traffic.

The literature review showed the lack of studies that addressed the speeding violation estimation or prediction models. Accordingly, this paper attempts to fill this gap by developing a model that predicts the frequency of speeding violations based on traffic information, road and speed cameras characteristics, and day of week. This model findings can be utilized to define the best locations of speeding camera in the future for more effective speeding enforcement programs.

3. Speeding Violations and Road Safety in AD

Figure 1 shows the trend of speeding violations in AD during the period from 2010 to 2015. The total number of traffic violations has been increased by 130% in year 2015 compared to year 2010 while the increase in the speed-related traffic violations is about 166%. This increase was occurred due to the increasing number of the installed automated speed cameras which increased from 414 in 2010 to 704 cameras in 2015 (about 70% increase). Figure 1 also shows that the speeding violations represent around 80% of total traffic violation types in average over the same period.

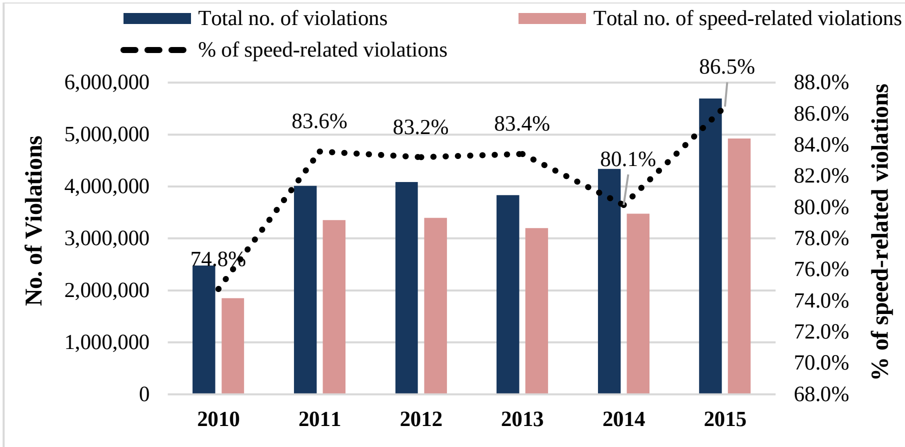

Fig. 1: Traffic violations statistics in Abu Dhabi.

Table 1 shows the speed-related severe crashes (i.e., any crash resulted in at least one injury or fatality) and the related fatalities in AD during the same period (from 2010 to 2015) and the crash severity (i.e. number of fatalities per 1000 severe crashes). This table reveals that a significant improvement in road safety has been achieved during the last five years where the numbers of the severe crashes and fatalities have been reduced by 29.1% and 34.8%, respectively, in year 2015 compared to year 2010. Although, the reduction percentage in the speed-related crashes and fatalities are relatively lower (9.7% and 1.7%, respectively), however; it is not complying with the reduction trend of the total crashes despite the increasing number of speed cameras and speeding enforcement campaigns. Therefore, a question regarding how efficient are the implemented site selection criteria to install speed camera has been raised. Towards this point, a pro-active approach to predict speed-related behaviours (represented in higher speeding violations rates) is required to maintain the improvements in traffic safety and to achieve the Emirate's strategic target in reducing the fatality rates to 3.0 fatalities per 100,000 inhabitants by 2021.

Table 1: Speed-related traffic safety indicators in AD.

Measure		**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**% change compared to 2010**
Severe crashes	Total number of serious crashes	2,542	2,283	2,055	2,071	1,861	1,803	**-29.1%**
	speed-related crashes	236	302	340	314	303	213	**-9.7%**
	% of speed-related crashes	9.3%	13.2%	16.5%	15.2%	16.3%	11.8%	**27.0%**
Fatalities	Total number of fatalities	376	334	271	289	267	245	**-34.8%**
	Speed-related fatalities	58	57	77	65	70	57	**-1.7%**
	% of speed-related fatalities	15.4%	17.1%	28.4%	22.5%	26.2%	23.3%	**51.1%**
Crash severity (fatality / 1000	In all the serious crashes	147.9	146.3	131.9	139.5	143.5	135.9	**-8.1%**
	in speed-related crashes only	245.8	188.7	226.5	207.0	231.0	267.6	**8.9%**

4. Data Collection and Case Study

The employed data in this study were collected at 76 sites of existing fixed speed cameras in hourly basis (total hourly data = 143,541) on two selected rural highways in Abu Dhabi Emirate. The selected highways are: Abu Dhabi-Dubai (E10+E11) and Abu Dhabi-Al Ain (E22) where they are considered two of top busiest roads in the emirate since they connect Abu Dhabi City to other major cities in the country. The collected data cover a three-month period from 1st of August till 31st of October in 2015. Data were managed to be collected from the Traffic and Patrol Directorate in Abu Dhabi Police. The data are comprehensive with all information associated with 1) traffic information such as: vehicle counts, average hourly vehicle speed, vehicle speed violation counts, posted and enforcement speed limits, speed margin, percentage of trucks; 2) site characteristics such as: enforcement camera's direction (vehicle front/ vehicle rear), existence of speed change zone, spacing from the previous fixed speed camera to the current one, spacing from the current fixed speed camera to the next one, the existence of horizontal curvature around 500 meters before and after the fixed speed camera location; 3) time characteristics whether the hourly reading fills in day or night times, weekdays or weekends, and specific day of week.

Data were prepared using three different tools. The first one is the Geographic Information System (GIS) to calculate the spacing between each two consecutive speed cameras at the same direction by using the concept of Linear Referencing (LR) implemented in ArcGIS desktop application. The second tool is MS Excel where data were stored and organized in a way to be used in advanced analyses and modelling approaches. The third used tool is SPSS Statistics software version 20 where the modelling approach was conducted and results were extracted. In SPSS, negative binomial model with log link function was used with the option to estimate the dispersion parameter rather than setting it to the system's default value. It accounts for the over-dispersion that is found in the crash data and quantifies an over-dispersion parameter.

It is worth mentioning that huge efforts were put in preparing the raw database in its final shape and extensive manual data collection was conducted to collect some site characteristics such as the existence of horizontal curve before and after the speed camera's location, the camera's enforcement direction, and coding the speed cameras that comes directly after the speed change zones. A geo-video recording using a GPS-based camera was used to collect some of the data along the study roads.

5. Speeding Violation Modelling

5.1. Variables Examination and Model Development

In the modelling approach, the dependent variable was set as the speeding violation frequency and the independent variables/predictors cover the ones that are related to traffic, site and cameras characteristics, and day time and type. Tables 2 and 3 show the examined variables in the model classified as; continuous and categorical variables, respectively.

Table 2: Descriptive statistics of continuous variables.

Variable	Description	Min.	Max.	Mean	Standard Deviation
Sp_Vio	Total hourly speed violation counts at camera site (violation/hr)	0	84	2.27	4.648
Avg_Speed	Average hourly speed for vehicles at camera site, (km/hr)	31	152	113.83	11.765
Ln(Veh_Count)	Natural logarithm of hourly vehicle counts at camera site (veh/hr)	0	8.72	6.7143	1.0322
Car_Posted	Posted speed limit for passenger cars at camera site, (km/hr)	80	140	123.26	16.510
Car_Enf	Enforcement speed limit for passenger cars at camera site, (km/hr)	101	161	146.30	14.642
Dist_Before	Distance between the current speed camera and the previous one, (km)	2	21	5.34	4.121
Dist_To	Distance between the current speed camera and the next one, (km)	2	21	5.42	4.089
Trucks	Truck percentage of the total traffic volume at camera site, (%)	0	0.571	0.010	0.024

The negative binomial regression model was used to predict the speed violation and has the form shown in Equation 1:

$$Ln\ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \qquad (1)$$

Where, Y is the dependent variable; $X_1, X_2, \ldots X_n$ are the independent variables; and $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are the regression coefficients.

The SPSS statistical software package was employed to estimate the model using the customized negative binomial with log link function with the option to estimate the dispersion parameter rather than setting it to the system's default value. It accounts for the over-dispersion that is found in the crash data and quantifies an over-dispersion parameter. All independent variables were tested for their influence on the dependent variable and the statistically significant ones were selected at 95% confidence level ($p \leq 0.05$).

Table 3: Descriptive statistics of categorical variables.

Variable	Description	Categories	Frequency (#)	Percent (%)
Cam_ED	The direction of the enforcement speed camera	vehicle rear	137,090	95.5
		vehicle front	6,451	4.5
Curve_Before	Existence of horizontal curve around 500 meters before the current speed camera	curve existed	23,451	16.3
		no curve existed	120,090	83.7
Curve_After	Existence of horizontal curve around 500 meters after the current speed camera	curve existed	36,950	25.7
		no curve existed	106,591	74.3
Day_Night	The time of day each hourly data fill in	Day-time	71,240	49.6
		Night-time	72,301	50.4
Day_of_Week	The day each hourly data fill in	Monday	20,355	14.2
		Tuesday	20,010	13.9
		Wednesday	20,234	14.1
		Thursday	20,553	14.3
		Friday	20,410	14.2
		Saturday	21,891	15.3
		Sunday	20,088	14.0
Weekday/Weekend	The type of day each hourly data fill in	Weekday	101,240	70.5
		Weekend	42,301	29.5
Cam_SCZ	Existence of speed camera after speed change zone	Yes	13,175	9.2
		No	130,366	90.8
Speed_Margin	Difference between the enforcement speed and posted speed limits at camera site	Speed margin=20 kph	66,489	46.3
		Speed margin=40 kph	77,052	53.7

5.2. Model Results and Discussion

Table 4 shows the results of the negative binomial model with log link function which represent the varibales that have significant impact at a confedance level of 95% ($p \leq 0.05$) on the occurance of speeding violations. The significant variables are: *Cam_ED, Curve_Before, Cam_SCZ, Speed_Margin, Dist_Before, Dist_To, Car_Posted, Weekend, Ln(Veh_Count), Avg_Speed, Trucks*.

The ***enforcement cameras that are directed to the vehicle front side*** were found as a significant variable which lead to higher speeding violations. One explanation of this finding is that the drivers are more familiar with the cameras directed to the vehicle rear side, since they are more common in Abu Dhabi, and hence facing a front directed camera may confuse them and does not let them reduce speed properly before reaching it which increase their chance to get a speeding violation ticket. However, the rear-aligned cameras let the drivers have more decision time to reduce speed before reaching the camera's point and hence lower probability to be involved in a speeding violation.

The results also showed that the ***non-existence of horizontal curve 500 meters before the location of speed camera*** is a significant variable with a positive estimated parameter. Drivers moving on a straight road segment have higher chance to get a speeding violation ticket compared to those arriving from a horizontal curve section where they may be driving with relatively lower speeds.

Table 4: Results of the model (significant variables only at 95% confidence level).

Variable	*β*	*Standard Error*	*Wald Chi-Square*	*p-value*	*Odds Ratio*
Intercept	-0.852	0.0563	229.095	0.000	0.427
Cam_ED (vehicle front)	0.067	0.0138	23.655	.000	0.935
Cam_ED (vehicle rear)	0[a]	--	--	--	--
Curve_Before (curve not existed)	0.155	0.0090	299.871	0.000	1.168
Curve_Before (curve existed)	0[a]	--	--	--	--
Cam_SCZ (No)	-0.190	0.0117	263.585	0.000	0.827
Cam_SCZ (Yes)	0[a]	--	--	--	--
Speed_Margin (20kph)	1.949	0.0087	49912.315	0.000	7.024
Speed_Margin (40kph)	0[a]	--	--	--	--
Weekend	0.101	0.0068	222.387	0.000	1.106
Weekday	0[a]				1
Dist_Before (km)	0.016	0.0010	240.607	0.000	1.016
Dist_To (km)	-0.006	0.0010	32.765	.000	0.994
Car_Posted (kph)	-0.116	0.0006	33762.315	0.000	0.891
Ln(Veh_Count), vph	0.372	0.0040	8432.645	0.000	1.450
Avg_Speed, (kph)	0.095	0.0007	21233.498	0.000	1.100
Trucks (%)	13.257	0.1851	5131.062	0.000	572246.344
Over-dispersion parameter	0.597	0.0051	--	--	--

[a] set to zero because the parameter is redundant.
Sample size = 143,541

Drivers are more likely to be involved in speeding violation at speed cameras locations that are next ***speed change zones (speed limit reduction)***. This can be explained as the drivers may not be aware about the change is posted speed limits at some locations and hence their violation rates are relatively higher at the first speed cameras locations just after the speed change zone.

Allowable speed margin of 20 kph over the posted speed limit is considered as a significant variable and contribute to higher violation counts. This can be justified as the drivers are feeling tight to a relatively smaller speed margin compared to the other allowable value of 40 kph so their change to violate at the early value is higher.

Weekdays were found significant in higher speeding violation frequencies. The weekend days (Fri, Sat) were found to be more directly proportional to the dependent variable compared to the weekdays (i.e. Sat, Mon, Tue, Wed, Thur). This finding can be explained as that the road-based trips are relatively higher on rural roads during the weekends which increases the probability of getting speeding violations especially on the rural highways that connects the Capital with the other major cities.

The results also showed a direct influence between the predicted speed violation variable and the ***natural logarithmic of vehicle counts, average vehicle speed, and percentage of trucks***. While the ***posted speed limits for passenger cars*** variable was found inversely proportional with the predicted speed violation variable.

The ***spacing between the previous speed camera and the current one*** was found significant in predicting speeding violations. The higher the distance between the current camera and the upstream camera the higher the predicted speed violation. However, the predicted speeding violation decrease with increasing the distance between the current camera and the downstream one. This means that installing speed cameras with relatively lower distances after the current site improves the traffic safety level on that road section.

To define the proper spacing between speed cameras, the same mode was tested by using distance variable (i.e., distance from/to the previous/next speed camera) on the speeding violations prediction. The same developed models were used here to find the relation between the speeding violation (independent variable) and before/ after spacing variables (dependent variables). The models results showed that for each one-unit increase on spacing distance from the previous speed camera, the expected natural logarithmic count of the speeding violations prediction increases by 0.056. Also, a one-unit increase on the spacing distance to the next speed camera, the expected natural logarithmic count of the speeding violations prediction increases by 0.089. A decision should be set on what average violation rate per hour is accepted for certain road to find the proper speed cameras spacing range that best maintain the traffic safety level on road section. One point to stress is that all other contributing variables (e.g. traffic volume, speed, % of trucks, etc.) should be fixed when selecting the desirable spacing. For example, to maintain an average violation rate per hour of 1.5 the spacing between speed cameras should be within a range from 4.5 to 7 kilometres. This finding can help the traffic police department or any other responsible authority in speed cameras allocations that sustain their strategic traffic safety levels as defined.

It is worth mentioning that Poisson regression model has been tested against negative binomial and their goodness-of-fit results were compared as shown in Table 5. The Poisson model was found to be relatively over-dispersed compared to the negative binomial model where the deviance value per degree of freedom is larger. For the other statistical tests, Akaike's information criterion (AIC); Finite Sample Corrected AIC (AICC); and Bayesian Information Criterion (BIC), they were relatively smaller at the negative binomial model which concludes that it is over perform the other one.

Another finding of the model is the over dispersion parameter. A zero dispersion means that the Poisson model would be more appropriate. As shown in Table 4, the calculated over-dispersion parameter is different from 0 and hence the negative binomial model is more appropriate to the data than the Poisson model.

Table 5: Goodness-of-fit results for negative binomial model.

Criterion	Poisson		Negative Binomial	
	Value	Value/DF	Value	Value/DF
Deviance	293092.992	2.042	147348.987	1.027
Scaled Deviance	293092.992	2.042	147348.987	1.027
Pearson Chi-Square	1808251.054	12.599	1079590.602	7.522
Scaled Pearson χ^2	1808251.054	12.599	1079590.602	7.522
Log Likelihood	-262969.629	--	-232894.280	--
AIC (smaller is better)	525977.257	--	465828.561	--
AICC (smaller is better)	525977.263	--	465828.567	--
BIC (smaller is better)	526164.871	--	466026.049	--

6. Conclusion

Prior studies proved a strong positive relationship between the speed and crash severity. Speed enforcement plays a main role in improving driver's speeding behaviour which lead to the improvement of road safety. Selecting the proper locations of speed cameras ensures more effective speed enforcement system. This paper investigated the relationship between speeding violations and the traffic information, speed camera, road characteristics, and day of the week as an approach to find the best locations to install speed camera in order to maintain the strategic traffic safety levels.

A model to predict the frequency of the speeding violation was developed by employing Negative binomial regression approach. Fifteen independent variables were examined to predict the frequency of speeding violations which cover the data related to traffic information, road and camera characteristics and day of week. The results showed that the traffic-related variables (i.e. traffic volume, average speed, and percentage of tucks); site-related variables (i.e. posted speed limit, speed margin, direction of enforcement camera, straight road segment, existence of speed change zone); and type of day (weekday/ weekend) are significantly influencing the occurrence of speeding violations at confidence level of 95%.

Based on the study findings, it is recommended to take into consideration the following points in the site selection process of speed cameras:

- Avoid locations just after horizontal curves and focus on the ones at straight segments and before the curves.
- Align the speed camera's face next to the traffic flow direction (i.e. face-to-face).
- Avoid locations of speed change zones.
- Installing the speed cameras with spacing depending on the average expected or desirable speeding violation rate per hour on a subject road. for example, at 1.5 average speeding violation, the spacing between radars should be in the range between 4.5 and 7.0 Km.
- Unify the margin between the posted speed and the enforced speed limits at all locations.
- Prioritize roads with high traffic volume with high average speed.

Further research is needed to investigate more related factors in the occurrence of speeding and other types of violations and to integrate the historical crashes that had been occurred near the speed cameras where such data were not available during the study time.

Acknowledgment

The authors would like to thank the traffic and patrols directorate, Abu Dhabi police, for providing the required data in this study. All opinions and results are those of the authors.

References

[1] A. Jahangiri, H. Rakha, T. A. Dingus, “Red-light running violation prediction using observational and simulator data,” *Accident Analysis & Prevention*, 2016.

[2] J. Bonneson, H. Son, “Prediction of expected red-light-running frequency at urban intersections,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1830, pp. 38-47, 2003.

[3] R. A. Retting, S. A. Ferguson, C. M. Farmer, “Reducing red light running through longer yellow signal timing and red light camera enforcement: results of a field investigation,” *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 327-333, 2008.

[4] B. E. Porter, K. J. England, “Predicting red-light running behavior: a traffic safety study in three urban settings,” *Journal of Safety Research*, vol. 31, no. 1, pp. 1-8, 2000.

[5] L. Zhang, K. Zhou, W. B. Zhang, J. Misener, “Prediction of red light running based on statistics of discrete point sensors,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2128, pp. 132-142, 2009.

[6] T. Rosenbloom, A. Perlman, “Tendency to commit traffic violations and presence of passengers in the car,” *Transportation research part F: traffic psychology and behaviour*, vol. 39, no. 10-18, 2016.

[7] J. C. F. De Winter, “Predicting self-reported violations among novice license drivers using pre-license simulator measures,” *Accident Analysis & Prevention*, vol. 52, pp. 71-79, 2013.

[8] H. B. Ayvaşık, N. Er, N. Sümer, “Traffic violations and errors: the effects of sensation seeking and attention,” in *Proceedings of the third international driving symposium on human factors in driver assessment, training and vehicle design*, pp. 395-402, 2005.