Analysis of adolescent traffic safety behavior based on multiple regression model

Hyun-Jun Cho[1, a, *], Samuel Adams[1, b],

[1]University of Exeter, Streatham Campus, Exeter, Devon, EX4 4QJ, United Kingdom
a.hyunjun_368124rok@yahoo.com, b.adsamuel_652@outlook.com
***Corresponding Author**

Abstract: In order to enhance the traffic safety level of adolescents in Guangzhou and ensure their safety, this study investigates the factors influencing adolescent traffic safety behavior based on the safety survey data of adolescents in Guangzhou. Through quantitative analysis, four indicators most closely related to adolescent traffic safety behavior were selected (education, awareness, attitude, and personal factors). A survey questionnaire was designed and used to obtain scores for traffic safety education, awareness, and attitude. SPSS 19.0 was employed to conduct Pearson correlation analysis and multiple regression model analysis for empirical analysis. The results indicate that education, awareness, attitude, and personal factors can well explain the scores of adolescent traffic safety behavior. The significance level sig. for all variables is less than 0.05, which suggests that all variables have a significant impact and can be effectively applied in the evaluation of adolescent traffic safety behavior.

Keywords: Traffic Engineering, Behavior Analysis, Multiple Regression Model, Adolescents, Traffic Safety

1. Introduction

With the in-depth study of traffic safety, people are increasingly recognizing the importance of the interrelationship between the safety of adolescents and their traffic behavior. Scholars at home and abroad have used methods such as causal chain analysis, multiple representation and multi-center methods, and risk analysis methods to conduct a comprehensive evaluation of the traffic environment and have established models for the safety and reliability of road traffic systems [1-4]. Literature [5-9] has conducted surveys on the traffic behavior and current situation of primary and secondary school students, statistically analyzed and quantified the collected data, identified relevant influencing factors, and conducted some exploratory analyses.

Due to the gap between China and developed countries in the systematic and targeted nature of primary school traffic safety education, the behavior of primary school students in urban areas in China is worse than that of primary school students abroad when it comes to safely crossing the road. Faced with complex traffic conditions, domestic research also believes that short-term behavior and temporary measures cannot fundamentally solve this problem [10]. This paper, based on travel behavior analysis theory, analyzes adolescent safety survey data. Currently, there is a lack of in-depth research on several key factors of adolescents' traffic behavior in real traffic environments domestically and internationally. The paper aims to strengthen the understanding of adolescent traffic behavior and explore the impact of education, awareness, attitude, and personal factors on adolescent safety behavior. By scoring adolescent traffic safety behavior and conducting further exploratory analysis, a multiple regression model is established, providing a theoretical basis for the prevention and control of traffic injuries and targeted traffic safety education for adolescents.

2. Survey and analysis

2.1. Survey subjects

Based on the different sources of students, schools in the main urban area of Guangzhou are divided into two major categories: schools for children of migrant workers and urban student schools. From each

https://doi.org/10.62852/ytr/2025/141

category, 2 primary schools and 2 secondary schools were randomly selected. Then, using cluster sampling methods, 5 classes were randomly selected from each of the 4 schools, and all students from the 40 classes were taken as the survey subjects.

Tianhe District Tianze Middle School has 5 classes each in the first and second years, totaling 430 people; Tianhe District Dong fang Primary School has 5 classes each in the fifth and sixth grades, totaling 490 people; Tianhe District Hui Jing Experimental School has 5 classes each in the first and second years, totaling 440 people; Tianhe District Wu Shan Primary School has 1 class each in the fifth and sixth grades, totaling 440 people. The total number of surveyed individuals is 1,800.

2.2. Survey methods

2.2.1. Questionnaire development

When designing the survey questionnaire, it is first necessary to analyze the factors affecting adolescent traffic safety, which are people, vehicles, roads, and environment. These four aspects are analyzed for their impact on adolescent traffic safety and are divided into five major sections for the questionnaire design: traffic safety education, traffic safety awareness, traffic safety attitude, traffic safety behavior, and student basic information. Ultimately, some quantifiable indicators are selected to set questionnaire questions, and scores are designed with gradient options.

The questionnaire is designed based on the cognitive characteristics of adolescents themselves. The questionnaire consists of 5 parts with a total of 35 questions.

The first part is an investigation of adolescents' traffic safety education, awareness, and attitude. This part has a unique operation, with a total of 35 items. First, adolescents are required to select their behaviors in road traffic for each item to assess the safety of adolescent traffic behavior; second, adolescents are required to select what they consider safe behaviors in road traffic for each item to assess adolescent traffic safety awareness; third, adolescents are required to select what they consider dangerous behaviors in road traffic for each item to assess adolescent traffic danger awareness.

The second part is about adolescents' road traffic behavior in the past 30 days, with a total of 10 questions. It includes the frequency of occurrence of various common incorrect traffic behaviors.

The third part involves demographic information, including: grade, gender, age, mode of travel, escort, distance from school, and channels of safety education, with a total of 7 questions.

The fourth part is a survey of adolescents' parents, with a total of 9 questions.

It takes about 15 minutes to complete each questionnaire.

2.2.2. Data collection

The entire questionnaire was completed by the adolescents themselves. For items that the adolescents did not understand, investigators provided explanations to ensure the accuracy of the survey data. A total of 1800 questionnaires were distributed, and 102 questionnaires were discarded due to improper filling or partial omissions, resulting in a total of 1698 valid questionnaires collected.

2.2.3. Analysis method

Data analysis was conducted using SPSS 19.0: First, score the first part of the questionnaire. The questionnaire used a Likert 5-point scale, and 2 to 3 measurable items were designed for each latent variable of each dimension. The scoring for traffic safety education, traffic safety behavior, and traffic safety awareness was consistent. The higher the score, the better the traffic safety education, awareness, and attitude. Second, frequency and descriptive statistical analysis were performed on the second and third parts of the questionnaire. Then, frequency and descriptive statistical analysis, Pearson correlation analysis, and multiple regression analysis were conducted.

2.3. Descriptive statistical analysis of survey data

Through a series of surveys conducted in the four schools, it was found that the comprehensive traffic safety scores of urban schools were better than those of schools for children of migrant workers, and the comprehensive traffic safety scores of middle schools were better than those of primary schools. As can be seen from Figures 1 to 5, due to factors such as family conditions, age, parents' education level, and the

https://doi.org/10.62852/ytr/2025/141

distance from home to school, students who live close to school mainly walk, while those who live far from school mainly use public transportation. The higher the parents' education level, the better the family's safety education, and the higher the overall safety awareness and attitude scores of the school. Family conditions also affect the overall level of safety awareness and attitude.

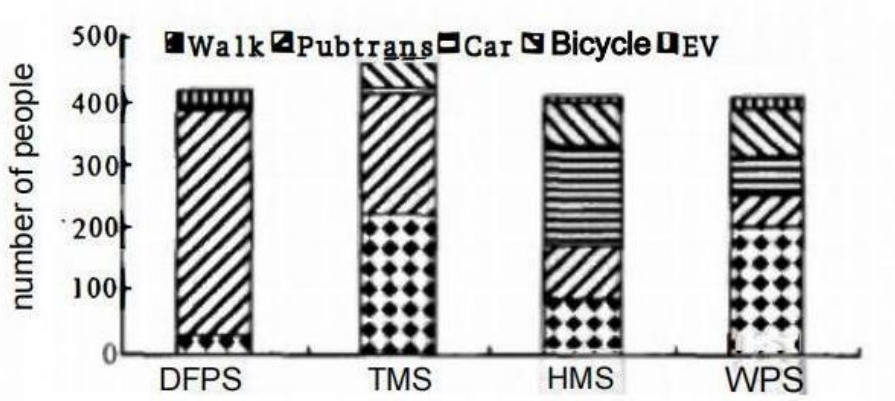

Figure 1: The most common mode of transportation to school

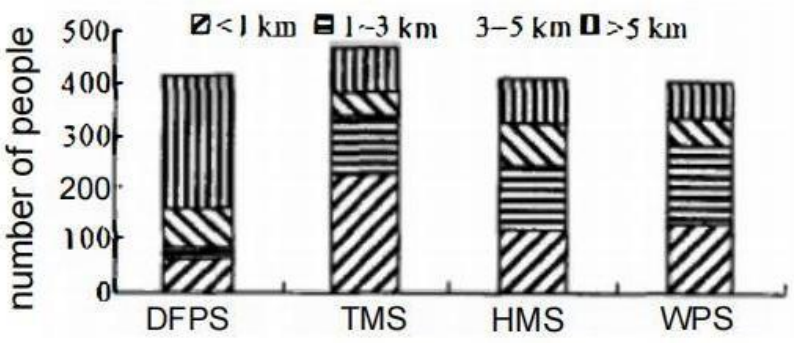

Figure 2: The distance from home to school

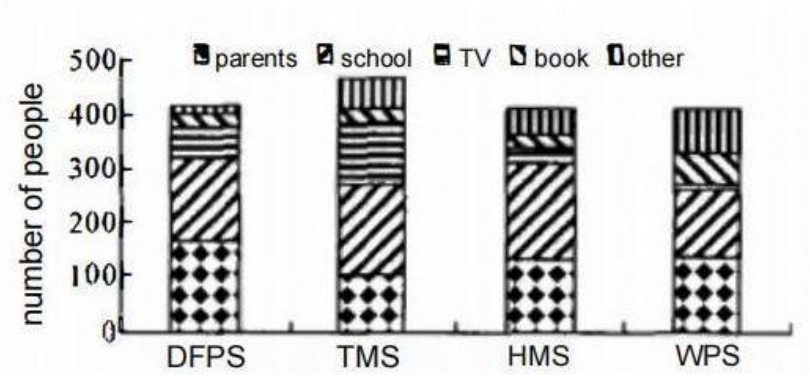

Figure 3: The main source of traffic safety knowledge

https://doi.org/10.62852/ytr/2025/141

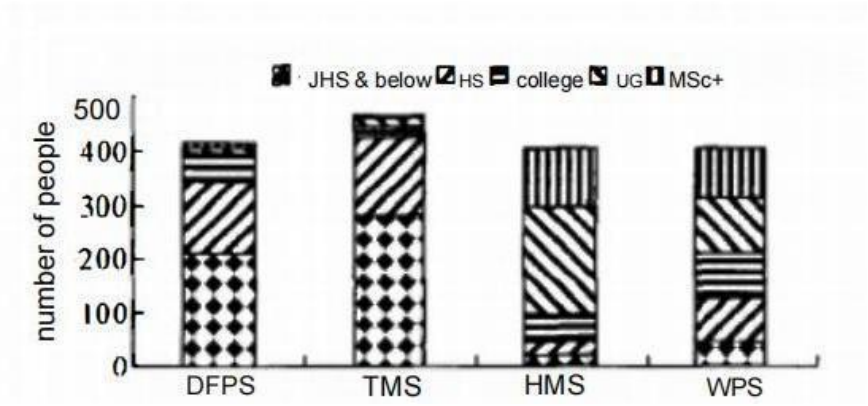

Figure 4: The highest level of education of parents

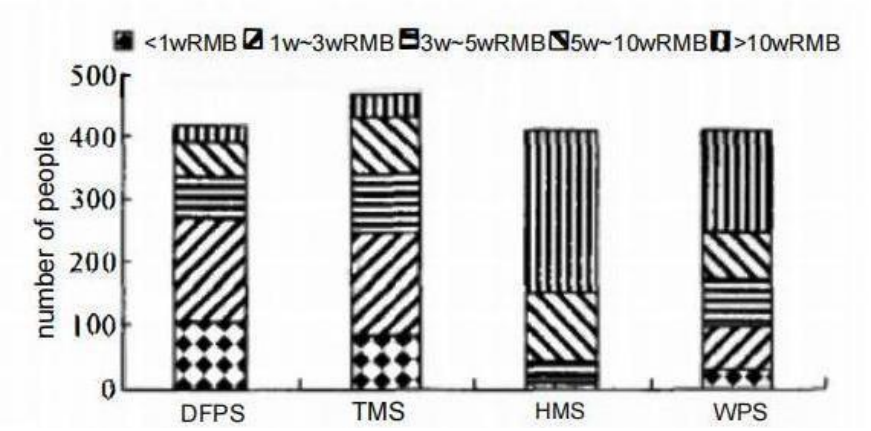

Figure 5: Annual household income

3. Analysis of factors influencing adolescent traffic safety behavior

Adolescent traffic safety education, adolescent traffic safety awareness, adolescent traffic safety attitude, and other influencing factors are used as explanatory variables. The other influencing factors refer to the scores obtained by combining age, mode of travel, travel distance, and parents' cultural level. The dependent variable is the score of adolescent traffic safety behavior. Pearson correlation analysis and multiple regression model analysis using SPSS 19.0 are employed for empirical analysis.

3.1. Pearson correlation analysis

The main purpose of the correlation analysis is to study the degree of closeness between various variables. Based on the data obtained from the distributed questionnaires, the relationships among the variables of adolescent safety education, adolescent safety awareness, adolescent safety attitude, other factors, and adolescent safety behavior scores are analyzed. The Pearson correlation matrix for these 5 variables using SPSS 19.0 software is presented in Table 1.

https://doi.org/10.62852/ytr/2025/141

Table 1 Correlation Matrix of Factors Influencing Adolescent Traffic Safety

		Education	Awareness	Attitude	Other	Total Score
education	correlation	1.00	0.391**	0.481**	0.277**	0.518*
	significance		0.000	0.000	0.000	0.000
awareness	correlation	0.391**	1.00	0.377**	0.247**	0.744**
	significance	0.000		0.000	0.000	0.000
attitude	correlation	0.481**	0.377**	1.00	0.204**	0.654**
	significance	0.000	0.000		0.000	0.000
other	correlation	0.277**	0.247**	0.204**	1.00	0.373**
	significance	0.000	0.000	0.000		0.000
total score	correlation	0.518**	0.744**	0.654**	0.373**	1.00
	significance	0.000	0.000	0.000	0.000	

From Table 1, it can be seen that the traffic safety behavior score is positively correlated with adolescent safety education, adolescent safety awareness, adolescent safety attitude, and other factors, and the Pearson correlation coefficient R for all five variables has reached a significant level, meaning that the correlation coefficients have passed the test and the obtained R values can be used. It can also be observed that the correlation coefficient R values with the adolescent traffic safety behavior score, arranged from largest to smallest, are adolescent traffic safety awareness, adolescent traffic safety attitude, adolescent traffic safety education, and other factors. The R values are 0.744, 0.654, 0.518, and 0.373, respectively. This indicates that the most effective way to improve adolescent traffic behavior is by enhancing their traffic safety awareness, and the impact of adolescent traffic safety attitude and education is also significant. The R value for other influencing factors is only 0.373, which is relatively low compared to the other variables but still reaches a significant level, suggesting that other factors have a potential impact on the traffic safety behavior of adolescents.

3.2. Regression analysis of adolescent traffic safety behavior

After the correlation analysis of adolescent traffic safety behavior, to further analyze the influencing factors of adolescent traffic safety behavior, a regression analysis of adolescent traffic safety behavior is required. Regression analysis is the process of finding a functional relationship between independent variables and dependent variables, thereby approximating the correlation between variables. From the correlation analysis, it is known that there is an approximate linear relationship between adolescent traffic safety behavior and each variable, hence linear regression analysis is chosen. The general mathematical model for linear regression is:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e$$

In the equation: (Y) represents the score of adolescent traffic safety behavior; (X1) represents adolescent traffic safety education; (X2) represents adolescent traffic safety awareness; (X3) represents adolescent traffic safety attitude; (X4) represents other influencing factors; (alpha) is the constant term; (e) is the random variable. SPSS 19.0 software is used to perform multiple linear regression analysis on these 5 variables and to establish a model. The following data were obtained, see Tables 2 and 3.

Table 2 Model Fit Statistics

model	R	R^2	Adjusted R^2	Estimate Standard Error
L	0.862	0.743	0.74	0.869

https://doi.org/10.62852/ytr/2025/141

Table 3 Mean Squares, F-values, and Significance Tests for the F-values

model		sum of squares	F	Sig.
L	regression	21702.54	228.806	0.000
	residual	7493.261		
	total	29195.801		

Table 2 lists the model's R, R^2, adjusted R^2, and estimated standard error. R is referred to as the multiple correlation coefficient; R^2 is the degree of fit of the model, meaning that the larger the value of R^2, the higher the rate of covariation between the independent and dependent variables, and the better the model's fit. The multiple correlation coefficient R obtained from this model is 0.862, the determination coefficient R^2 for fitting the multiple linear regression is 0.743, the adjusted determination coefficient is 0.740, and the estimated standard error is 0.869. This indicates that the model has a relatively good fit.

Table 3 shows that the model's regression sum of squares is 21702.540, the residual sum of squares is 7493.261, the total sum of squares is 29195.801, the F-value is 228.806, and the F-value is relatively large, indicating that the regression model is significant; moreover, the sig. value is approximately 0.000, sig. < 0.05, therefore, it can be considered that the multiple linear regression equation model built in this study is effective. Table 4 presents the coefficients of this multiple linear regression equation model.

Table 4 Regression Coefficients

Model	Non-standardized Coefficients		Standard Coefficient	t	sig.
	bi	Standard Error			
Constant	6.845	3.249		2.107	0.036
Safety Awareness	0.098	0.030	0.128	3.230	0.001
Safety Education	0.485	0.038	0.473	12.661	0.000
Safety Attitude	0.292	0.031	0.363	9.339	0.000
Other Factors	0.060	0.020	0.107	3.068	0.002

In this simulation, the non-standardized coefficients for the regression of the dependent variable Y (traffic safety behavior score) on the four independent variables X_1 (safety education), X_2 (safety awareness), X_3 (safety attitude), and X_4 (other factors) are 0.092, 0.485, 0.292, and 0.060, respectively. The corresponding t-values for the significance tests are 3.230, 12.661, 9.339, and 3.068. The significance levels sig. for the regression coefficients b_1, b_2, b_3, and b_4 are at most 0.001, and all sig. values are less than 0.05. Therefore, it can be concluded that the independent variables X_1 (safety education), X_2 (safety awareness), X_3 (safety attitude), and X_4 (other factors) all have a significant impact on the dependent variable Y (traffic safety behavior score). Based on the data in the above tables, the complete multiple linear regression equation can be obtained.:

$$Y = 6.845 + 0.098X_1 + 0.485X_2 + 0.292X_3 + 0.060X_4$$

It is evident from regression model (2) that the score of adolescent traffic safety behavior is positively correlated with adolescent traffic safety education, adolescent traffic safety awareness, adolescent traffic safety attitude, and other corresponding factors, among which the impact of adolescent traffic safety awareness and adolescent traffic safety attitude is the greatest.

Figure 6 is used to study the histogram of residuals. It is one of the tools and methods to observe whether the residuals conform to the normal distribution. In this figure, it can be clearly seen that the shape of the residuals on the graph basically conforms to the normal distribution.

https://doi.org/10.62852/ytr/2025/141

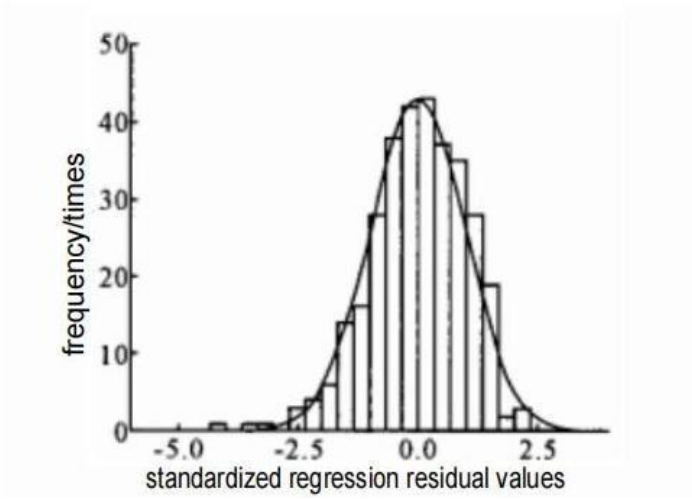

Figure 6 Histogram of Traffic Safety Behavior Scores

As shown in Figure 7, it can be clearly observed that the residual plot is a scatter plot, and it can be seen that the scatter points are randomly distributed in the horizontal band centered at (e = 0). Therefore, it can be proven that the model is usable.

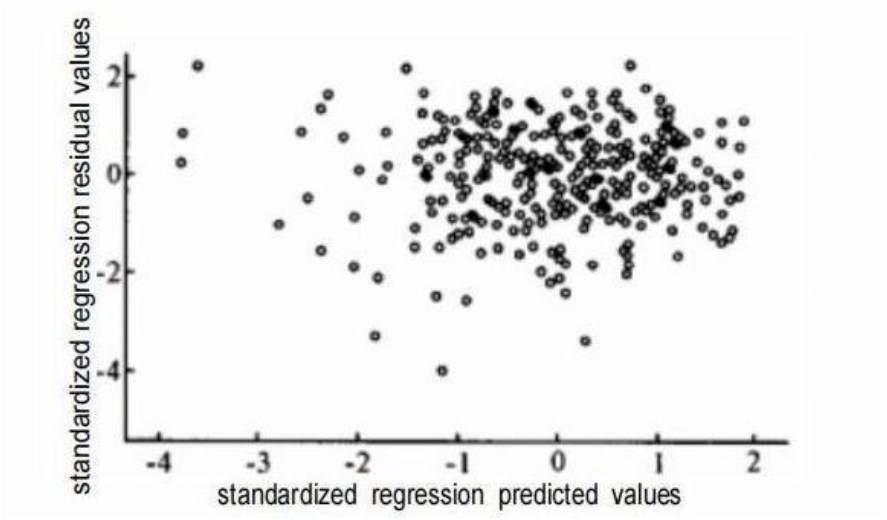

Figure 7 Residual Plot of Traffic Safety Behavior Scores

Based on the model fit analysis results from Figures 6 and 7, it can be determined that the derived multiple linear regression model is feasible.

4. Conclusion

The author, through quantitative analysis combined with safety survey data from adolescents in the main urban areas of Guangzhou, has established and calibrated an adolescent traffic safety regression analysis model, leading to two conclusions.

The model not only reflects the analysis of the traffic safety situation of adolescents but can also be used to estimate and evaluate traffic safety conditions.

https://doi.org/10.62852/ytr/2025/141

The traffic safety awareness and attitude of adolescents largely dictate their traffic safety behavior; it is only by enhancing their awareness and attitude that the traffic safety of adolescents can be fundamentally guaranteed.

5. References

[1] ELVIK R. Assessing the validity of road safety evaluation studies by analysing causal chains[J]. Accident Analysis &. Prevention,2003,35(5):741-748.

[2] WANG Xiang, PAN Xiao Dong, CONG Hao Zhe, Children road traffic safety education effect evaluation based on multiple representations[C]. Management Science and Industrial Engineering (MSIE),2011 International Conference on. Har bin: IEEE,2011:89-92.

[3] VIET Hung K, HUYEN L T. Education influence in traffic safety: A case study in Vietnam [J]. IA TSS research,2011,34(2):87-93.

[4] [4]LI Shu Qing, LI Zhe. Construction of the traffic safety reliability models in Road Traffic System[C]. 11¹International Conference of Chinese Transportation Professionals (ICCTP). Nanjing: Southeast U- niversity,2011:2086-2091

[5] YAO Yu Hua, CHEN Dao Yong, ZHOU Feng. Traffic injury knowledge, attitude and behaviors of primary school students in Hong Kou district in shanghai[J]. Chinese Journal of School Health,2008(11):991- 992. (in Chinese)

[6] SHI Qi, WANG Xue Song, YANG Dong Yuan, et al. Analysis of traffic safety among middle school students and safety education improvements [J]. China Safety Science Journal,2011,21(12):143- 152. (in Chinese)

[7] WANG Wei, HUANG Ying, GAO Ning Bo. A road safety evaluation model based on fuzzy definition theory[J]. Journal of Transport Information and Safety,2013(5):87-92. (in Chinese)

[8] WANG Li, WANG Xue Song, YANG Dong Yuan. Precise education of pupils' traffic safety considering schools and grades differences[J]. Journal of Transport Information and Safety,2012,30(1):103-110. (in Chinese)

[9] LI Hai Qiong, QIN Ya qin, ZHAO Lin Lin, et al. Research on traffic safety behavior of children pedestrian in Kunming [J]. Chinese Journal of Ergonomics, 2012(1):45-49. (in Chinese)

[10] SONG Yang, WANG Xue Song. Research on improving traffic safety education for pupils and middle school students in China[J]. China Safety Science Journal,2013(2):153-159. (in Chinese)

https://doi.org/10.62852/ytr/2025/141