Structural Equation Model Analysis for the Identification of the Safety Tolerance Zone

Eva Michelaraki* [0000-0002-7167-4630], George Yannis [0000-0002-2196-2335]

National Technical University of Athens, Athens, Greece
*evamich@mail.ntua.gr

Abstract. Task complexity refers to the dynamic road and environmental conditions influencing driving, while coping capacity reflects a driver's ability to manage these challenges. The concept of the Safety Tolerance Zone (STZ) describes the point at which self-regulated control is considered safe, where the demands of the driving task are balanced with the driver's ability to cope. This study endeavours to quantify the relationship between task complexity and coping capacity in identifying the STZ. Towards that end, a driving simulator experiment involving 55 participants was conducted and the most reliable indicators were collected and analysed. Generalized Linear Models were performed to examine the impact of road, vehicle and driver characteristics on crash risk, while Structural Equation Models were applied to assess the relationship between latent and observable variables of task complexity, coping capacity and crash risk. Results indicated that increased task complexity significantly elevated crash risk, while greater coping capacity mitigated it. It was also revealed that task complexity and coping capacity were inter-related with a positive correlation, implying that drivers coping capacity increases as the complexity of driving task increases. Findings highlighted the importance of adaptive driver support systems in maintaining drivers within the STZ under varying road conditions.

Keywords: Safety Tolerance Zone, Task Complexity, Coping Capacity, Crash Risk, Structural Equation Model.

1 Introduction

Road safety is a critical issue that affects communities worldwide, as the growing number of vehicles on the roads has led to increased risks of crashes, injuries and fatalities [1]. It is essential to ensure all road users' safety (i.e. drivers, passengers, cyclists and pedestrians) which requires not only proper infrastructure and traffic management but also awareness and responsibility from individuals. Road safety is influenced by multiple factors such as the driving behaviour of individuals, the condition of vehicles, the design and maintenance of roads and the surrounding environmental conditions [2].

The concept of a Safety Tolerance Zone (STZ) is an important aspect of road safety, as it highlights the limits within which drivers, vehicles and road systems can operate safely. Every road environment has a certain level of risk, but when traffic speeds, vehicle conditions or driver behaviour go beyond the STZ, the likelihood of serious crashes and fatalities increases significantly. This zone considers factors such as the design of the road, speed limits, visibility and the ability of vehicles to withstand impacts in case of collisions.

The STZ represents the point at which self-regulated driving remains safe, balancing task complexity with the driver's coping capacity. It consists of three phases: normal driving, where crash risk is low; danger, where risk increases but a crash can still be avoided; and avoidable accident, where a collision is imminent unless timely corrective action is taken. Task complexity reflects the driving environment and demands imposed on the driver (e.g., traffic, road, weather, distractions), while coping capacity refers to the driver's ability to manage these demands through experience, skills, and strategies. Crash risk emerges when high task complexity exceeds coping capacity, reducing awareness, slowing reactions and increasing errors. Conversely, higher coping capacity helps drivers adapt and maintain safe behaviour.

This study takes a holistic approach by modelling how road, vehicle and driver-related risk factors interact within the STZ framework. Data from a simulator experiment with 55 drivers (165 trips over two months) were analysed. Key indicators included task complexity (e.g. time of day, weather), coping capacity - vehicle state (e.g. fuel type, vehicle age, gearbox) and coping capacity - driver state (e.g. speeding, headway, fatigue, harsh events).

The structure of the paper is as follows: At the beginning, the motivation and objectives of this study, along with an explanation of the STZ concept are presented. Next, a comprehensive overview of the data collection process is provided. Then, the research methodology is described, including the theoretical basis of the models. Moreover, the analysis results are presented, followed by a detailed discussion of the conclusions. Finally, suggestions and directions for future research are highlighted.

2 Data Overview

Within the framework of this study, a driving simulator experiment was implemented involving 55 drivers (with total duration of 2 months) and a database consisting of 165 trips was created. A custom simulator developed by DriveSimSolutions, based on a Peugeot 206 with Original Equipment Manufacturer (OEM) parts (dashboard, instrument cluster and seat) was used.

To minimize simulator sickness (e.g. dizziness, nausea), scenarios avoided sharp movements, drives lasted under one hour and breaks were included. Participants completed two short practice runs (5-10 minutes) for familiarization, with trials halted if discomfort occurred. Three scenarios were used, a highway, rural and urban road, each ~15 km in length, mostly under daytime and normal weather conditions. Figure 1 provides an overview of the different scenarios of the simulator experiment.

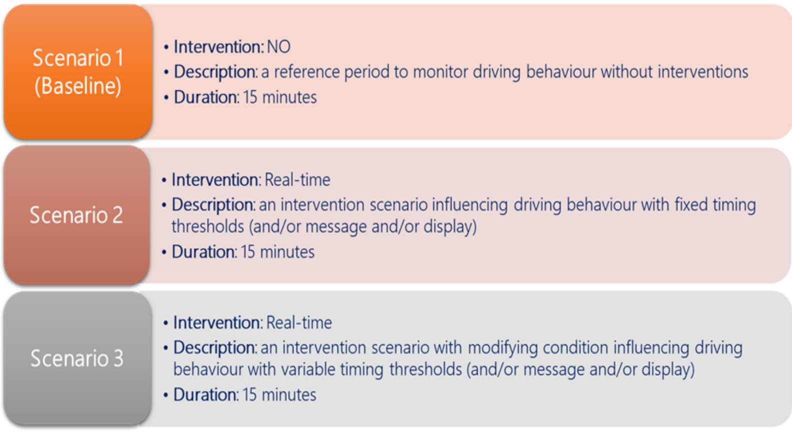

Fig. 1. Overview of the different scenarios of the simulator experiment

3 Methodology

3.1 Generalized Linear Models (GLM)

Numerous studies have examined the relationship between crashes and contributing factors such as roadway geometry, environment, traffic and human behaviour [3, 4]. The Generalized Linear Model (GLM) extends traditional linear regression by allowing error distributions beyond the normal and introducing a link function to relate predictors to the response variable [5]. Each outcome is assumed to follow a distribution from the exponential family (e.g., normal, binomial, Poisson, gamma), with the expected response value μ expressed as a function of the predictors X [6].

$$E(Y|\mathrm{X}) = \mu = g^{-1}(X\beta) \qquad (1)$$

where: E(Y|X) is the expected value of Y conditional on X; Xβ is the linear predictor, a linear combination of unknown parameters β; g is the link function. In regression modelling, it is also important to address multicollinearity, which occurs when independent variables are highly correlated. This can be assessed using the Variance Inflation Factor (VIF), with values <5 generally indicating suitability for inclusion.

3.2 Structural Equation Models (SEM)

Structural Equation Models (SEMs) are widely used in road safety research to analyse the complex interactions between observed variables and latent constructs. They are particularly valuable because they separate measurement error from true scores, enabling researchers to assess both direct and indirect effects of multiple factors on crash risk [7]. SEM, also known as path analysis, is a multivariate technique for testing hypotheses about relationships among observed and latent variables [8]. SEM consists of two parts: the measurement model, which assesses how well observed indicators capture latent constructs while accounting for error and the structural model, which examines causal relationships among variables, including both direct and indirect effects [9]. The general formulation of SEM, as outlined by Washington et al. [10], is as follows:

$$\eta = \beta\eta + \gamma\xi + \varepsilon \tag{2}$$

where η and ξ denote endogenous and exogenous variable vectors, respectively, β and γ are coefficient matrices and ε represents error terms. Measurement relationships are defined as $x = \Lambda x\xi + \delta$ and $y = \Lambda y\eta + \zeta$, where observed variables (x, y) are linked to latent constructs through factor loadings (Λx, Λy) with associated measurement errors (δ, ζ).

3.3 Goodness-of-Fit Measures

Model evaluation was conducted using multiple goodness-of-fit indices that assess how well the hypothesized model reproduces the observed variance–covariance matrix. Absolute and incremental fit measures included the Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) and Tucker–Lewis Index (TLI), where values above 0.90 (and preferably 0.95 for CFI) indicate acceptable to excellent fit and RMSEA values ≤ 0.05 suggest close fit. In addition, model selection criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were employed to compare competing models, both of which incorporate a penalty for model complexity to avoid overfitting; lower values of AIC and BIC indicate a more parsimonious and better-fitting model.

4 Results

4.1 Generalized Linear Model (GLM)

Headway, the distance or time gap between vehicles, is a critical factor in rear-end crashes, one of the most common collision types. Maintaining sufficient headway gives drivers more time to react to sudden stops or traffic changes. In this study, headway is defined as the time gap between the lead vehicle and the following vehicle (time headway).

To explore this, a GLM examined the relationship between headway and selected indicators of task complexity and coping capacity. The dependent variable was a binary indicator of headway events (1 = event, 0 = none). In particular, the "headway event" binary variable was defined based on time-headway thresholds derived from STZ phases, where values below 1.4 seconds were classified as critical (event = 1) and values above this threshold as non-critical (event = 0), with intermediate ranges used to distinguish driving phases. Task complexity was represented by the time indicator (i.e. daytime and night-time driving conditions), while coping capacity variables included trip duration, average speed, time-to-collision (TTC), fatigue and hands-on events (reflecting potential distraction when drivers frequently return their hands to the wheel).

Findings derived from Table 1 demonstrated that all the explanatory variables were statistically significant at a 95% confidence level. In addition, there was no issue of multicollinearity as the VIF values are much lower than 5. With regards to the coefficients, it was found that time of the day was negatively correlated with headway, which means that drivers tend to keep safer distances from the vehicle in front of them during the night.

Coping capacity indicators showed that higher average speed and shorter TTC increased the likelihood of headway events. Fatigue and hands-on events were also positively correlated, suggesting that impaired alertness or re-engagement with the steering wheel often leads to reduced following distances. In contrast, trip duration was negatively associated with headway events, indicating that longer drives promote steadier patterns and fewer adjustments in following distance.

Table 1. Parameter estimates and multicollinearity diagnostics of the GLM for headway

Variables	Estimate	Std. Error	z-value	Pr(|z|)	VIF
(Intercept)	0.859	0.221	3.896	< .001	-
Time indicator	-0.690	0.318	-7.443	< .001	1.209
Average speed	0.742	0.080	9.231	< .001	1.020
Time to collision	0.004	3.116	14.300	< .001	1.018
Duration	-5.658	1.395	-4.057	< .001	1.040
Fatigue	5.088	1.587	3.206	0.001	1.114
Hands on wheel	5.369	2.311	2.323	0.020	1.076

Summary statistics

AIC : 4546.08; BIC: 4141.62; Degrees of freedom: 33820

4.2 Structural Equation Model (SEM)

Risk was measured by means of the STZ levels for headway (level 1 refers to 'normal driving' used as the reference case; level 2 refers to 'dangerous driving' while level 3 refers to 'avoidable accident driving'). Task complexity, represented by trip duration and distance travelled, was positively correlated with the construct. Coping capacity, measured through TTC, average speed, hands-on events and fatigue, showed mixed effects: fatigue and hands-on events correlated positively, suggesting drivers may compensate by adopting more cautious behaviour, while higher TTC and faster speeds were negatively associated, indicating reduced ability to manage driving demands under such conditions. However, it should be noted that higher TTC values are generally associated with safer conditions, while the observed relationship in the model reflected context-specific interactions with other variables rather than a direct reduction in coping capacity. The SEM path diagram is presented in Figure 2.

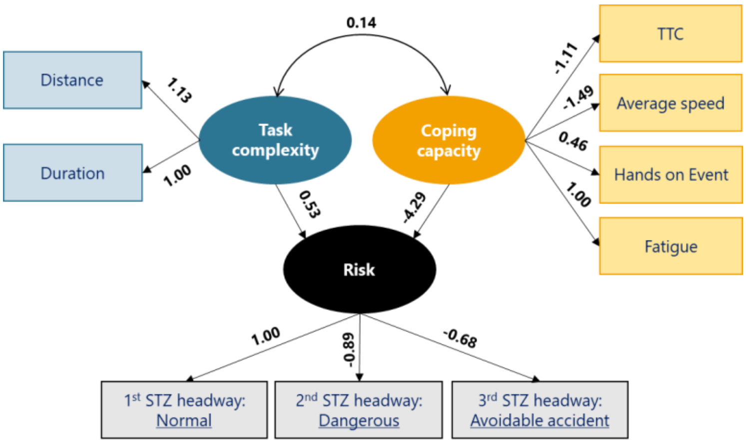

Fig. 2. SEM results of task complexity and coping capacity on risk (STZ headway)

The CFI of the model is equal 0.966; TLI is 0.944 and RMSEA is 0.079. Table 2 summarizes the SEM applied, while residual variances are presented in Table 3.

Table 2. Model Fit Summary for headway

Model Fit measures	Value
CFI	0.966
TLI	0.944
RMSEA	0.079
GFI	0.973
Hoelter's critical N (α = .05)	247.929
Hoelter's critical N (α = .01)	300.037
AIC	65281.042
BIC	65445.959

Table 3. Residual variances for headway

Variable	Estimate	Std. Error	z-value	P(>|z|)
Distance	0.108	0.024	4.576	< .001
Duration	0.107	0.023	4.542	< .001
Fatigue	0.950	0.024	39.002	< .001
Time to collision	0.939	0.025	38.280	< .001
Average speed	0.890	0.026	33.990	< .001
HandsOnEvent	0.989	0.024	40.565	< .001
Headway_STZ_level_0	-0.242	0.059	-4.082	< .001
Headway_STZ_level_1	0.177	0.049	3.652	< .001
Headway_STZ_level_2	0.422	0.029	14.344	< .001

The structural model revealed a positive correlation between task complexity and coping capacity ($\beta = 0.14$), suggesting that drivers tend to enhance their coping strategies as driving demands increase. Task complexity was positively associated with crash risk ($\beta = 0.53$), while coping capacity showed a strong negative relationship with risk ($\beta = -4.29$), indicating its protective role in reducing the likelihood of unsafe outcomes.

5 Discussion

Results of both the GLM and SEM analyses highlighted the importance of headway as a critical risk factor in driving safety. The GLM findings confirmed that drivers' behavioural and situational characteristics directly influence following distances. Specifically, higher average speeds, shorter TTC, fatigue and hands-on events were positively associated with unsafe headway events, suggesting that both performance and distraction-related factors elevate crash risk. Conversely, time of day and trip duration were protective, with night-driving and longer trips linked to safer headway.

SEM analysis further enriched these insights by modelling the relationships between latent constructs. Task complexity, represented by exposure indicators such as distance and duration, showed a positive correlation with coping capacity, implying that drivers tend to adjust their behaviour to manage increasing task demands. However, while greater task complexity increased overall crash risk, stronger coping capacity was associated with reduced risk. Importantly, the negative correlation between coping capacity and risk highlights the protective role of adaptive behaviour in maintaining safe driving performance.

A key limitation of this study is the relatively small simulator sample size, which may affect the generalizability of the findings. In addition, the dataset was limited to simulator conditions and may not fully reflect the diversity of driving behaviours across different populations, regions, or transport modes. Participant health or medical status was not considered, which could also influence driving performance. Future research should include larger and more diverse samples, with attention to age, education and driving experience. Comparative studies across countries and transport modes could provide broader insights. Expanding the range of task complexity and coping capacity indicators and applying advanced methods (e.g. deep learning) would further strengthen the analysis and enhance the predictive power of the STZ concept.

6 Conclusions

The objective of this study was to identify the STZ using a structural equation modelling framework. A driving simulator experiment with 55 participants was conducted to capture behavioural, vehicle-related and environmental indicators and to assess how task complexity and coping capacity interact to influence crash risk.

Results demonstrated that headway is a critical determinant of safety, with observable factors such as speed, fatigue, time-to-collision and driver distraction significantly influencing unsafe following distances. SEM analysis further revealed that increased task complexity was associated with higher crash risk, while greater coping capacity had a protective effect, enabling drivers to adapt their behaviour and mitigate risk. Recommendations arising from this research include the need to incorporate task complexity and coping capacity indicators into driver assistance systems and real-time monitoring technologies, ensuring that drivers receive timely feedback and support when their STZ is shifting towards danger. Training programmes should also emphasize awareness of headway, fatigue or distraction as key determinants of safety.

Acknowledgments

The research was funded by the European Union's Horizon 2020 i-DREAMS project (Project Number: 814761) funded by European Commission under the MG-2-1-2018 Research and Innovation Action (RIA).

References

1. Ahmed, S. K. et al. Road traffic accidental injuries and deaths: A neglected global health issue. Health science reports, 6(5), e1240 (2023). https://doi.org/10.1002/hsr2.1240
2. Wang, C., Quddus, M. A., & Ison, S. G. The effect of traffic and road characteristics on road safety: A review and future research direction. Safety science, 57, 264-275 (2013). https://doi.org/10.1016/j.ssci.2013.02.012
3. Michelaraki, E. et al. Investigating the effect of driver-vehicle-environment interaction with risk through naturalistic driving data. In Transport Research Arena Conference, pp. 147-161. Cham: Springer Nature Switzerland (2024). https://doi.org/10.1007/978-3-031-88974-5_22
4. Papantoniou, P. Structural equation model analysis for the evaluation of overall driving performance: A driving simulator study focusing on driver distraction. Traffic injury prevention, 19(3), 317-325 (2018). https://doi.org/10.1080/15389588.2017.1398825
5. Hastie, T. J., & Pregibon, D. Generalized linear models. In Statistical models in S (pp. 195-247). Routledge (2017). https://doi.org/10.1201/9780203738535
6. Dunn, P. K., & Smyth, G. K. Chapter 10: Models for Counts: Poisson and Negative Binomial GLMs. In Generalized Linear Models With Examples in R (pp. 371-424). New York, NY: Springer New York (2018). https://doi.org/10.1007/978-1-4419-0118-7_10
7. Yuan, K. H., & Bentler, P. M. 5. Finite Normal Mixture SEM Analysis by Fitting Multiple Conventional SEM Models. Sociological methodology, 40(1), 191-245 (2010). https://doi.org/10.1111/j.1467-9531.2010.0122
8. Harrison, L., & Friston, K. J. Effective connectivity. Statistical parametric mapping: the analysis of functional brain images, 508-521 (2007).
9. Stein, C. M., Morris, N. J., & Nock, N. L. Structural equation modeling. In Statistical human genetics: Methods and protocols (pp. 495-512). Totowa, NJ: Humana Press (2011).
10. Washington, S., Karlaftis, M., Mannering, F., & Anastasopoulos, P. Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC (2020).