An Roinn Iompair
Department of Transport

10th Transport Research Arena Conference
Advancing Sustainable and Inclusive Mobility
Dublin, Ireland, April 15-18, 2024

Exploitation of naturalistic driving data to estimate crash risk through machine learning

Spyridon Tsigkos
Transportation Engineer

Together with:
Eva Michelaraki, Elena Provatari, George Yannis
National Technical University of Athens

Introduction

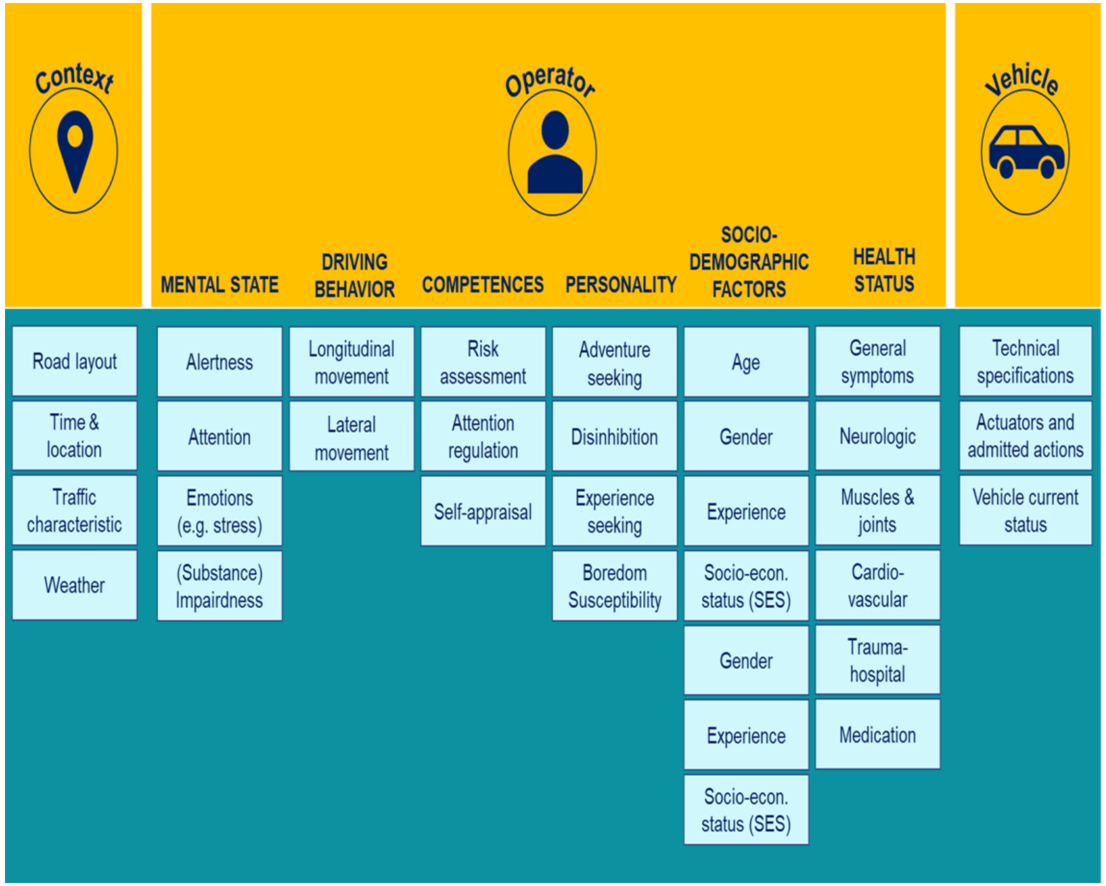
- In modern society and reality, road transport consists an important part of people's daily lives, as they use it for all their activities during the course of a day

- With the ever-increasing need for daily traffic movements, road crashes are also increasing, bringing to light the prominent issue of road safety

- Annual road traffic fatalities amount to 1.19 million and are the leading cause of death in people aged 5-29 years

Data Description

- Task complexity relates to the current status of the real world context in which a vehicle is being operated:
  - ✓ road layout (i.e. highway, rural, urban)
  - ✓ time and location
  - ✓ traffic volumes (i.e. high, medium, low
  - ✓ weather conditions
- Coping capacity is dependent upon two underlying factors and it consists of several aspects of vehicle state:
  - ✓ technical specifications
  - ✓ current vehicle status
- and operator state:
  - ✓ mental state
  - ✓ driving behavior
  - ✓ competencies
  - ✓ personality
  - ✓ sociodemographic profile
  - ✓ health status



Objectives

- Identification of the most critical indicators of risk from both the task complexity and the coping capacity (vehicle and operator state) side

- Explanatory variables of risk, such as time head-way, distance traveled, speed, time of the day (lighting indicators), or weather conditions were assessed

- Development of SEMs (Structural Equation Models) for estimating the effect of different factors (speeding, harsh braking etc.) in crash risk

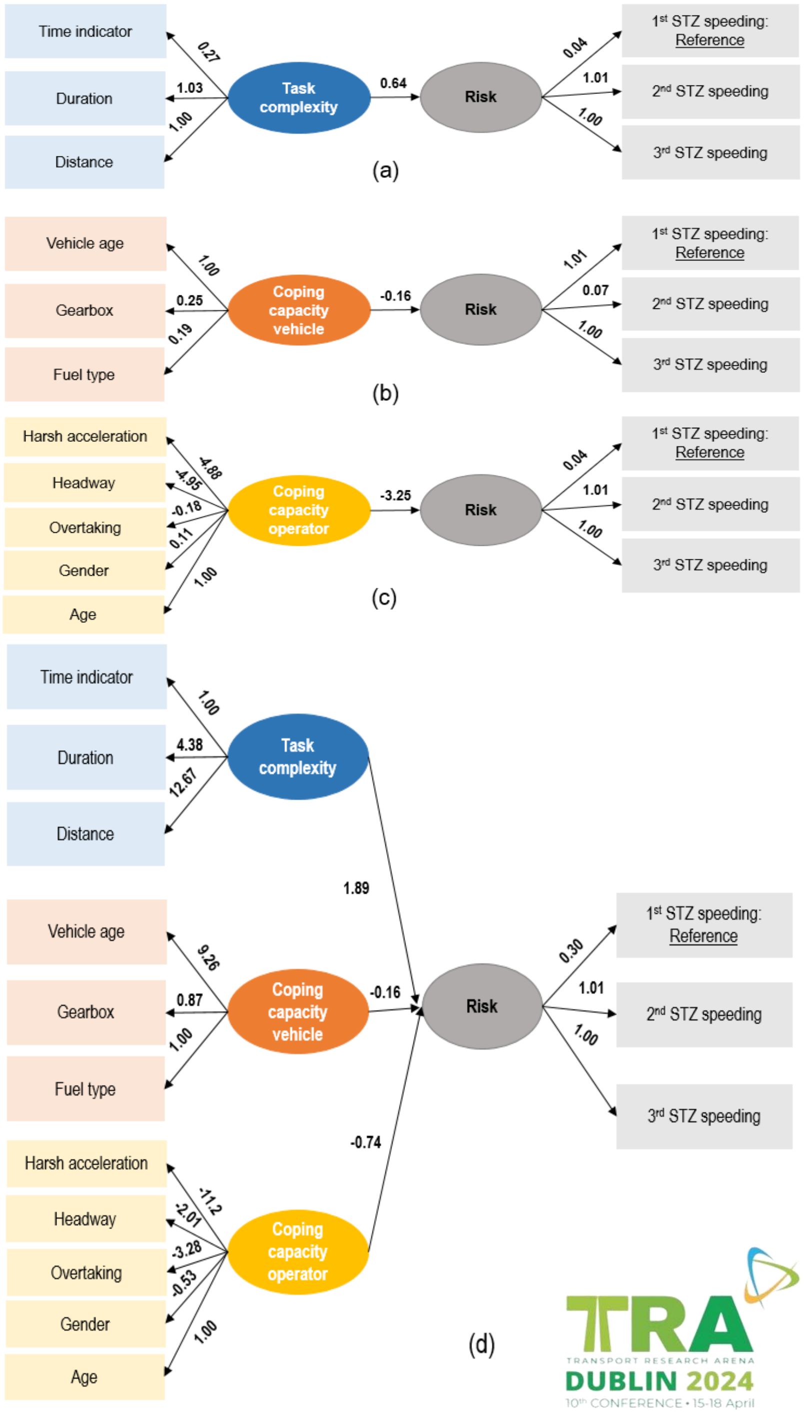
Methodology

- A naturalistic driving experiment was carried out for 4 months, involving 30 car drivers from the UK and a large database was collected and analyzed
- Structural Equation Model (SEMs) were developed in order to identify the relationship between observed (i.e. number of speeding or harsh braking events) and latent or unobserved variables (i.e. task complexity)
- The results of the models were evaluated by satisfying the following statistical tests: $p\text{-value}<0.001$, $CFI>0.90$, $TLI>0.90$, $SRMR<0.05$, $RMSEA<0.05$
- Through tests carried out, the dependent variables for which the most significant results were obtained, were speeding and harsh braking

TRA TRANSPORT RESEARCH ARENA DUBLIN 2024 10th CONFERENCE • 15-18 April

Results - Speeding

- Risk is measured by means of the STZ (Safety Tolerance Zone) levels for speeding (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’)
- The numbers on the arrows indicate the correlation among the variables
- The most important finding is the positive correlation between the complexity of the driving task and risk, meaning that increased task complexity relates to increased risk according to the model
- The structural model between coping capacity and risk shows a negative coefficient, which means that increased coping capacity relates to decreased risk according to the model



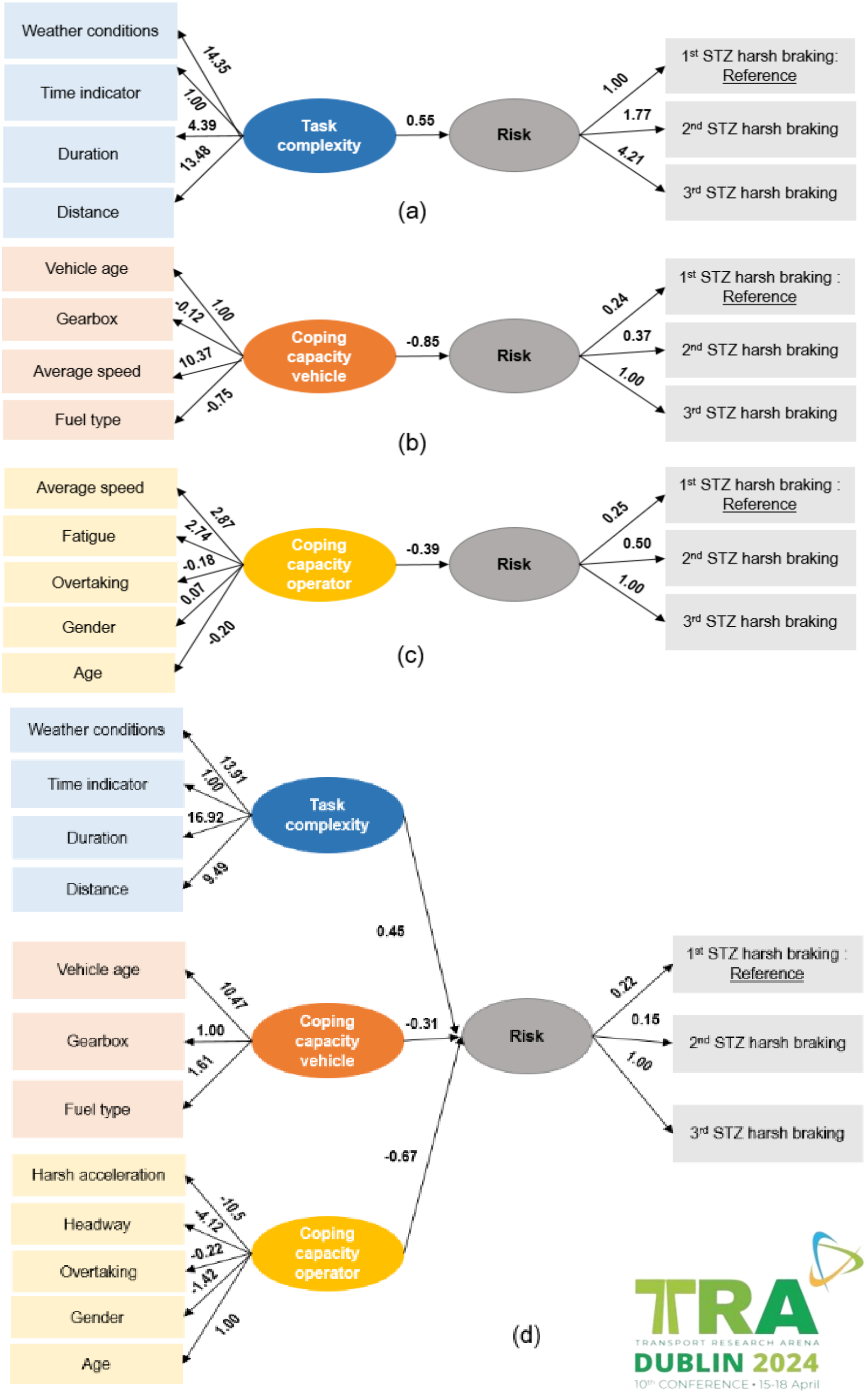
Results - Harsh Braking

- Risk was measured through the types of harsh events and the level of severity at which they occur

- For harsh braking, there are high severity events (1st STZ level), moderate severity events (2nd STZ level) and low severity events (3rd STZ level).

- Overall, there is a positive correlation between task complexity and risk (regression coefficient=0.45), while there is a negative correlation between vehicle and operator state and risk

- On the other hand, SEM between coping capacity and risk shows a negative coefficient, which means that increased coping capacity relates to decreased risk (coefficient=-0.31 (vehicle state) / -0.67 (operator state))



Results - Overall

- It was found that p-values <0.001 across all metrics revealed strong evidence of model fit for both speeding and harsh braking
- For speeding, all metrics demonstrated high levels of fit, with CFI values ranging from 0.997 to 0.981, TLI values ranging from 0.994 to 0.969, SRMR values between 0.032 and 0.066, and RMSEA values between 0.052 and 0.101.
- For harsh braking, significant p-values alongside CFI values ranging from 0.992 to 0.707, TLI values between 0.987 and 0.613, SRMR values from 0.045 to 0.090, and RMSEA values spanning 0.080 to 0.148 affirm the robustness of the model fit across various dimensions, highlighting its reliability in capturing the complexities associated with both driving behaviors

Model fit metrics	Task complexity	Coping capacity operator state	Coping capacity vehicle state	Synthesis
Speeding				
p-value	<0.001	<0.001	<0.001	<0.001
CFI	0.997	0.992	0.730	0.952
TLI	0.994	0.987	0.601	0.936
SRMR	0.048	0.045	0.153	0.074
RMSEA	0.062	0.080	0.440	0.100
Harsh braking				
p-value	<0.001	<0.001	<0.001	<0.001
CFI	0.981	0.707	0.897	0.832
TLI	0.969	0.613	0.848	0.845
SRMR	0.032	0.090	0.066	0.064
RMSEA	0.052	0.148	0.101	0.085

Discussion

- Higher task complexity was associated with an increased crash risk

- Drivers could probably become overwhelmed by the demands of complex tasks, leading to reduced attention to the road and other traffic participants

- Increased coping capacity relates to decreased risk according to the model

- This means that by increasing the coping capacity of both the operators and the vehicles, we can create a safer driving environment

Conclusions

- Results demonstrated that coping capacity and risk displayed a negative relationship across all phases, indicating that higher coping capacity was associated with reduced risk

- The relationship among task complexity, coping capacity and risk, may depend on the specific context and the type of task or activity involved

- Higher task complexity may increase the potential for errors or crashes, as it can lead to greater cognitive or physical demands on the individual performing the task

- Similarly, a higher coping capacity may help to reduce the risk of crashes or errors, as it can provide individuals with the resources or strategies needed to effectively manage challenging or stressful situations

Future Research

- Future research should take into account other critical factors affecting the driver's capability to cope with the complexity of driving. For instance, road type and the volume of traffic composition can be critical parameters in defining task complexity
- Develop alternative techniques for examining feature importance
- Further investigation of the significance of variables can more accurately determine the relationship among variables
- Further improvement of existing or creation of new applications for improving driving behaviour to obtain more accurate data and variables to be exploited

An Roinn Iompair
Department of Transport

10th Transport Research Arena Conference
Advancing Sustainable and Inclusive Mobility
Dublin, Ireland, April 15-18, 2024

Exploitation of naturalistic driving data to estimate crash risk through machine learning

Spyridon Tsigkos
Transportation Engineer

Together with:
Eva Michelaraki, Elena Provatari, George Yannis
National Technical University of Athens