This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 814761

D6.3

An integrated model of driver-vehicle-environment interaction and risk

Safe tolerance zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions

iDREAMS

Project identification

Grant Agreement No	**814761**
Acronym	*i*-DREAMS
Project Title	Safety tolerance zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions
Start Date	01/05/2019
End-Date	30/04/2023
Project URL	www.idreamsproject.eu

Document summary

Deliverable No	**6.3**
Deliverable Title	A integrated model of driver-vehicle-environment interaction and risk
Work Package	6
Contractual due date	28/04/2023
Actual submission date	28/04/2023
Nature	Report
Dissemination level	Public
Lead Beneficiary	NTUA
Responsible Author	Eva Michelaraki, Christos Katrakazas, George Yannis (NTUA)
Contributions from	Eva Michelaraki, Christos Katrakazas, George Yannis (NTUA) Evita Papazikou, Laurie Brown, Rachel Talbot (Loughborough University) Amir Pooyan Afghari, Eleonora Papadimitriou (TUD) Adnan Muhammad, Muhammad Wisal Khattak, Tom Brijs (UHasselt)

Please refer to the document as:

Michelaraki, E., Katrakazas, C., Yannis, G., Papazikou, E., Brown, L., Talbot, R., Afghari, A.P., Papadimitriou, E., Muhammad, A., Khattak, W.M., Brijs, T. (2023). *An integrated model of driver-vehicle-environment interaction and risk.* Deliverable 6.3 of the EC H2020 project i-DREAMS.

Revision history (including peer review & quality control)

Version	Issue date	% Complete	Changes	Contributor(s)
v1.0	30/10/2021	0	Initial deliverable structure	Eva Michelaraki
v1.3	03/03/2023	80%	First draft	See 'contributions from above'
v1.7	16/03/2023	100%	Full draft for review	As above
V1.9	28/04/2023	100%	Revised according to external and internal review	As above

Disclaimer

Copyright

Table of contents

List of Figures

List of Tables

Glossary and Abbreviations

Abbreviation	Description
ADAS	Advanced Driver Assistance Systems
AIC	Akaike Information Criteria
ANN	Artificial Neural Network
API	Application Programming Interfaces
BIC	Bayesian Information Criterion
CC	Coping Capacity
CFI	Comparative Fit Index
DBN	Dynamic Bayesian Network
DCM	Discrete Choice Model
FCW	Forward Collision Warning
FDR	False Discovery Rate
FN	False Negative
FP	False Positive
GFI	Goodness of Fit Index
GLM	Generalized Linear Model
GRPL	Grouped Random Parameters Binary Logit
HP	Horsepower
IBI	Inter-Beat-Interval
ICT	Information and Communication Technology
i-DREAMS	smart Driver and Road Environment Assessment and Monitoring System
ITS	Intelligent Transportation System
LDW	Lane Departure Warning
LSTM	Long Short Term Memory Network
NN	Neural Network
OPFS	Ordered Probit Fractional Split
PCW	Pedestrian Collision Warning
RMSE	Root Mean Square Error
RMSEA	Root Mean Square Error Approximation
SD	Standard Deviation
SDK	Software Development Kit
SEM	Structural Equation Model
STZ	Safety Tolerance Zone
TC	Task Complexity
TLI	Tucker Lewis Index
TN	True Negative
TP	True Positive
UK	United Kingdom
VIF	Variance Inflation Factor

Executive Summary

This Deliverable aims at **developing an integrated model** of driver-vehicle-environment interaction and risk by:

(i) identifying the most critical precursors of risk from both the task complexity and the coping capacity side,

(ii) implementing an integrated model for understanding the effect of the inter-relationship of task complexity and coping capacity with risk, and

(iii) comparing the performance of such models on different countries.

The ultimate goal of the analyses in this project was to identify the impact that the balance between task complexity and coping capacity has on the risk of a crash. For that reason, a vast library of data from naturalistic driving experiments was created in five countries (i.e. Belgium, UK, Germany, Greece and Portugal) to investigate the most prominent driving behavior indicators available, including speeding, headway, overtaking, duration, distance and harsh events (i.e. harsh acceleration and harsh braking). It is also important to investigate common behaviors and driving patterns across different countries, as well as to identify specific interventions that have been effective in improving road safety in different contexts. By understanding these factors, it may be possible to identify strategies that can be used to promote safer driving behaviors and reduce the incidence of crashes in different countries.

After making a short summary of the project's aims and objective, the naturalistic driving experiment procedure in all of the countries involved was described along with the data acquisition, data cleaning and data aggregation procedures followed to extract the datasets that were used in the analyses. These strategies aimed to comprehend how the data were stored in the back-end database, how to deal with missing values, how to impute missing values taking into account the natural meaning of the recorded variables and how to best exploit the data for developing the models applied. The volume, diversity and noise included in the dataset, due to the different experimental difficulties faced in each of the countries led to extensive efforts to acquire clean data. The total number of drivers, trips and minutes per country and transport mode is presented in Table below:

Drivers	Belgium (cars)	Belgium (trucks)	UK (cars)	Germany (cars)	Greece (cars)	Portugal (buses)	Total
Phase 1	39	23	53	28	65	29	**237**
Phase 2	43	22	54	28		29	**176**
Phase 3	51	22	53	27	65	26	**244**
Phase 4	49	23	54	28	65	22	**241**
Max	**51**	**23**	**54**	**28**	**65**	**29**	**250**

Trips	Belgium (cars)	Belgium (trucks)	UK (cars)	Germany (cars)	Greece (cars)	Portugal (buses)	Total
Phase 1	1173	1448	3073	1397	2937	2459	**12487**
Phase 2	1549	1691	3317	1322		1363	**9242**
Phase 3	1973	1440	3417	1129	3935	1411	**13305**
Phase 4	2468	1767	4594	1496	2194	2098	**14617**
Summary	**7163**	**6346**	**14401**	**5344**	**9066**	**7331**	**49651**

Duration (minutes)	Belgium (cars)	Belgium (trucks)	UK (cars)	Germany (cars)	Greece (cars)	Portugal (buses)	Total
Phase 1	23725	117160	56853	23617	51786	202532	**475673**
Phase 2	31414	146315	58458	19469		123132	**378788**
Phase 3	40121	139245	59556	17704	69962	145934	**472522**
Phase 4	52077	187636	93974	23644	39695	232323	**629349**
Summary	**147337**	**590356**	**268841**	**84434**	**161443**	**703921**	**1956332**

It should be noted that **Structural Equation Model (SEM)** is used to explore how the model variables are inter-related, allowing for both direct and indirect relationships to be modeled. In particular, observed variables are measurable, whereas unobserved variables are latent constructs. These models are often represented by a path analysis, showing how a set of 'explanatory' variables can influence a 'dependent' variable. In this Deliverable, particular emphasis was given in SEM analysis as it was found to be the **most widely used and appropriate for modeling complex and multi-layered relationships** between observed (e.g. number of speeding and headway events) and unobserved variables (e.g. crash risk).

The next section of the deliverable describes in detail, the **methodologies followed** throughout the analyses. Apart from SEMs, Generalized Linear Models (GLMs), Neural Networks (LSTMs and shallow), as well as Grouped Random Parameters Binary Logit and Ordered Probit Fractional Split Models are described.

Ultimately, the goal of these analyses was to identify the **impact that the balance between task complexity and coping capacity** has on the risk of a crash. The results of those analyses are thoroughly described in Chapter 4 of the current Deliverable.

Through the application of SEM models, the analyses revealed that **higher task complexity levels lead to higher coping capacity** by drivers. Additionally, the influence of task complexity on risk was greater than the effect of coping capacity in Belgium, Greece and Germany and mixed results were observed in the UK and Portugal. Models fitted on data from different phases of the experiments validated that interventions had a positive influence on risk compensation, increasing drivers' coping capacity and reducing dangerous driving behavior.

Furthermore in Chapter 5, predictive **real-time analyses** (NNs and LSTMs) demonstrated that it is possible to predict the level of Safety Tolerance Zone (STZ) with up to 95% accuracy, while **post-trip explanatory** studies (GRPL and OPFS) showcased the capacity of state-of-the-art econometric models to shed light on the complex relationship of risk with task complexity and coping capacity. The comparison of models fitted on data from the different phases of the experiments, validated that in the majority of the countries the interventions had a positive influence on risk compensation, increasing the coping capacity of the operators and reducing the risk of dangerous driving behavior. Moreover, predictive real-time analyses demonstrated that it is possible to predict the level of STZ with an accuracy of up to 95%, while post-trip explanatory studies showcased the capacity of state-of-the-art econometric models to shed light on the complex relationship of risk with the interdependence of task complexity and coping capacity.

An overview of the effects found for task complexity and coping capacity on risk among all available data can be found in Table below. A positive sign means a positive correlation of task complexity or coping capacity with risk while a negative sing indicates a negative relationship between task complexity or coping capacity and risk.

Country (transport mode)	Indicator	Phase 1		Phase 2		Phase 3		Phase 4	
		TC	CC	TC	CC	TC	CC	TC	CC
Belgium (cars)	speeding	-	+	-	+	-	+	+	+
	headway	-	+	-	+	-	-	-	+
Belgium (trucks)	speeding	-	-	-	-	-	-	-	-
	harsh acceleration	+	-	+	-	+	-	+	-
	headway	-	-	-	-	-	-	+	-
UK (cars)	headway	-	-	+	-	-	-	-	-
Germany (cars)	speeding	+	-	+	-	+	-	+	+
Greece (cars)	speeding	+	-			+	-	+	-
Portugal (buses)	headway	+	-	-	-	+	-	+	-
Overall (cars)	speeding, headway, overtaking, fatigue	+	-	+	-	+	-	+	-

TC refers to Task Complexity and CC refers to Coping Capacity

The difference in the relationship between variables across different countries could be due to a variety of factors, such as cultural differences, economic factors, or variations in driving behaviors and infrastructure.

In the final part of the Deliverable, **conclusions are drawn** for the relationship between task complexity, coping capacity and risk, while explanations for the model drawbacks are given.

On the basis of the i-DREAMS results, a set of policy recommendations at different levels (EU, national and local authorities, industry, etc.) can be provided. The i-DREAMS system itself can directly improve safety once launched, but also additional safety benefits can be envisaged in the medium and long term as it is built on and further adapted to different contexts and industry needs, thanks to its modular nature. The effectiveness of the i-DREAMS system may depend on a variety of factors, including the specific context in which it is implemented, the quality and accuracy of the data used to train the system, and the degree of integration and adoption by drivers and other stakeholders.

The integrated treatment of task complexity, coping capacity and risk can improve behavior and safety of all travelers and all transport modes, through the unobtrusive and seamless monitoring of behavior. Thus, authorities may use data systems at population level to plan mobility and safety interventions, set up road user incentives, optimize enforcement and enhance community building on safe traveling.

1 Introduction

The goal of this section is to provide a brief outline of the objectives of the specific deliverable, how those are aligned and relevant with the overall project, and which approach was followed in order to achieve them.

1.1 About the project

The overall objective of the i-DREAMS project is to setup a framework for the definition, development, testing and validation of a context-aware safety envelope for driving ('Safety Tolerance Zone'), within a smart Driver, Vehicle & Environment Assessment and Monitoring System (i-DREAMS). Taking into account driver background factors and real-time risk indicators associated with the driving performance as well as the driver state and driving task complexity indicators, a continuous real-time assessment will be made to monitor and determine if a driver is within acceptable boundaries of safe operation (i.e. Safety Tolerance Zone). Moreover, the to-be-developed i-DREAMS platform will offer a series of in-vehicle interventions, meant to prevent drivers from getting too close to the boundaries of unsafe operation and to bring them back into the Safety Tolerance Zone (STZ) while driving. The safety-oriented interventions will be developed to inform or warn the driver real-time in an effective way as well as on an aggregated level after driving through an app- and web-based gamified coaching platform, thus reinforcing the learning of safer driving habits/behaviors. Consequently, the i-DREAMS platform will allow the implementation of the two aforementioned safety interventions, meant to motivate and enable human operators to develop the appropriate safety-oriented attitude.

Specifically, the in-vehicle interventions are meant to assist and support vehicle operators in real-time (i.e. while driving). Depending on how imminent crash risks are, a distinction can be made between a 'normal driving' phase, a 'danger' phase, and an 'avoidable accident' phase. In the normal driving phase, no abnormalities in a vehicle operator's driving style are detected by the monitoring pillar of the i-DREAMS platform, and no sign of a crash course initiating is present. Consequently, no real-time intervention is required. In the danger phase, abnormal deviations from the vehicle operator's driving style are detected by the i-DREAMS monitoring module, and the potential for a crash course to unfold is present. A warning signal is to be issued in that case. In the avoidable accident phase, deviations from normal driving have evolved even further, and the risk for a crash to occur will become imminent if the vehicle operator does not adapt appropriately and immediately to the present circumstances. A more intrusive warning signal is provided to support vehicle operators in avoiding a collision.

With regards to post-trip interventions, these are not operational while driving, but they are based on what happens during a trip. They hinge upon all the raw data that is captured by the i-DREAMS sensors, which is further processed and fused into information about a vehicle operator's driving style, how it evolved during a trip, how many (safety-critical) events occurred, and in which circumstances these events happened. This information can be further translated into feedback consultable for vehicle operators via an app in a pre- or post-trip setting. To establish a longer-term relationship with individual vehicle operators, app-supported feedback can be combined with the use of a web-based coaching platform, containing gamification features meant to motivate drivers to work on a gradual and persistent improvement of their driving.

Figure 1 summarizes the conceptual framework, which will be tested in a simulator study and three stages of on-road trials in Belgium, Germany, Greece, Portugal and the United Kingdom (UK) with a total of 600 participants representing car, bus, truck and tram/train drivers. For the purpose of the current research, data from 250 drivers (car, trucks and buses) were analyzed.

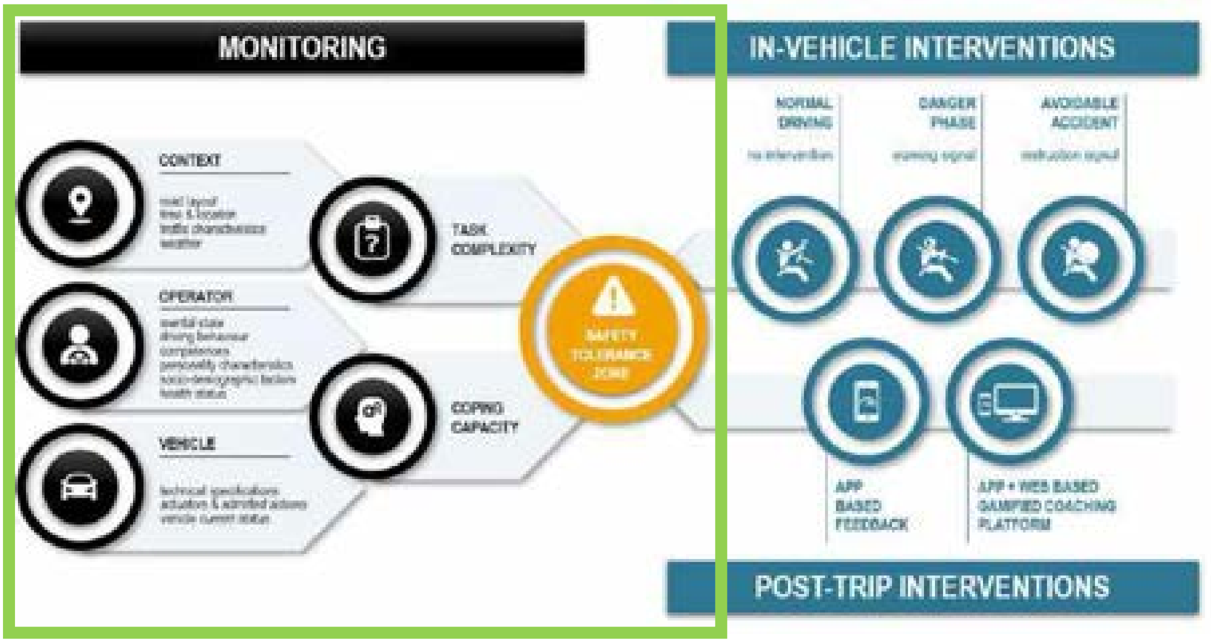

Figure 1: Conceptual framework of the i-DREAMS platform. The green frame indicates the thematic scope of this deliverable (see section 1.2)

Expected by the end of the project in 2023, the key output of the project will be an integrated set of monitoring and communication tools for intervention and support, including in-vehicle assistance and feedback and notification tools as well as a gamified platform for self-determined goal setting working with incentive schemes, training and community building tools. Furthermore, a database with anonymized data with regards to human factors (e.g. speeding, harsh events, fatigue) from the simulator and field experiments will be developed. It should be noted that only the monitoring would be assessed in this Deliverable and the impact of both real-time and post-trip intervention will be investigated in Deliverable 7.2 (Brown et al., 2023).

1.2 About this report

The work presented in this deliverable relates to the left part of Figure 1 (see green box), i.e. the determination of Safety Tolerance Zone via monitoring of task complexity and coping capacity. Staying within the STZ, vehicle operators avoid situations in which a collision becomes unavoidable. As can be seen in Figure 1, the Safety Tolerance Zone is subdivided in three segments, i.e. ‘normal driving’, the ‘danger phase’, and the ‘avoidable accident phase’. For the real-time determination of this Safety Tolerance Zone, the monitoring module in the i-DREAMS platform continuously registers and processes data for all the variables related to the context and to the vehicle. Regarding the operator however, continuous data registration and processing are limited to mental state and behavior. Data related to operator competence, personality, socio-demographic background, and health status, are collected via survey questionnaires.

It should be noted that the current deliverable is directly related to the Deliverables 6.1 (Papazikou et al., 2023) and 6.2 (Michelaraki et al., 2023). In particular, Deliverable 6.1 focuses on the relationship between task complexity and risk and Deliverable 6.2 deals with the interaction of coping capacity on risk, without investigating the potential interaction between both latent concepts on risk, while this report mainly focuses on the **development of an integrated model of driver behavior and safety**, based on the interaction of 'task complexity' on the one hand, and 'coping capacity' on the other hand, with 'risk'. A complete Structural Equation Model (SEM) will be developed to describe the interactions between task complexity and coping capacity (i.e. related to both vehicle state and operator state factors). From the parameter estimates of the integrated model, a comprehensive set of quantitative effects of indicators will be created, describing the impacts of vehicle, operator and context characteristics on risk under different conditions. Lastly, comparisons among different countries and transport modes will be made.

1.2.1 Aims and objectives

This deliverable has following aims and objectives:

- **Identification of the most critical precursors of risk** from both the task complexity and the coping capacity(vehicle and operator state) side
- **Examination of the effect of task complexity and coping capacity** (i.e. vehicle and operator state) on risk across the four phases of i-DREAMS road-trial on a transport mode basis. A detailed description of the phases can be found on Table 1.
- **Implementation of an integrated model** for understanding the effect of the aforementioned inter-relationship with risk.
- **Extraction of a comprehensive set of quantitative effects of indicators**, describing the impacts of vehicle, operator and context characteristics on risk under different conditions.
- **Comparison of the performance** of such models on different countries.

1.2.2 Structure

The organization of the Deliverable is the following:

Chapter 2 provides a **detailed description of the field trial study design**. In particular, an overview of the obtained dataset, the questionnaire data collected as well as the procedure followed for data aggregation and cleaning is clearly explained. In addition, the definition of task complexity and coping capacity (i.e. vehicle and operator state) is provided and the variables used to define task complexity and coping capacity along with some descriptive statistics are presented.

This is followed by a **description of the methodological anlaysis** (Chapter 3) in which the purpose of this analysis along with the concept of Multivariate Regression Analysis (e.g. Generalized Linear Modeling technique) and latent variables analysis (e.g. Structural Equation Models) are highlighted. In addition, a methodological overview of real-time techniques, such as Neural Networks, classification and Long Short-Term Memory Networks as well as post-trip approaches, such as Grouped Random Parameters Binary Logit models and Ordered Probit Fractional Split models is given. The key performance indicators and appropriate metrics that are commonly used for model evaluation and selection are also descripted.

The major part of this Deliverable is dedicated to the **mathematical modeling of the STZ** (Chapter 4), where Generalized Linear Models and Structural Equation Models are implemented in order to turn the available measurements into meaningful information on the Safety Tolerance Zone level.

Chapter 5 aims to **develop an integrated model** of driver behavior and safety, based on the interaction of task complexity and coping capacity with risk. To that end, real-time (i.e. Neural Networks, classification and Long Short-Term Memory Networks) and post-trip (i.e. Grouped Random Parameters Binary Logit models and Ordered Probit Fractional Split models) analyses are implemented in order to examine the impact of vehicle, operator and context characteristics on risk under different conditions. Comparisons among the examined countries (i.e. Belgium, UK, Germany, Greece, Portugal) and different transport modes (i.e. cars, trucks and buses) are also provided.

Lastly, Chapter 6 draws the **main findings along with practical conclusions** and gives recommendations for further research.

2 *i*-DREAMS data collection

2.1 Experiment description

Within the i-DREAMS project, a **naturalistic driving experiment** was carried out involving 250 drivers from Belgium, UK, Germany, Greece and Portugal and a large database of 49,651 trips and 1,956,332 minutes was created. A detailed description of the on-road driving trials for identifying STZ and the performance of in-vehicle interventions can be found in previous Deliverable 5.3 (Hancox et al., 2021).

It should be highlighted that the i-DREAMS field trials are the first time that all components of the complete i-DREAMS system are combined in a real-world setting, where it can be used by individuals and organisations outside of the i-DREAMS project.

The objectives of the on-road trials in i-DREAMS are to:

- **test the driving behavior** and validate the STZ mathematical model
- test if the i-DREAMS system **influences driver safety**
- **assess the effect of the interventions** (developed as part of the i-DREAMS system) for both real-time and post-trip warnings and
- obtain the user feedback about the **acceptance and acceptability** of the i-DREAMS system

The on-road trials in i-DREAMS were designed based on several proven principles derived from previous literature focusing on testing interventions in order to assist drivers in maintaining the STZ. As the first stage of the field trials, **pilot testing** was performed for a limited number of vehicles (i.e. five vehicles) for each test site. The purpose of the pilot tests was to fine-tune the i-DREAMS technology. This includes all the processes associated with production, installation and interventions but also collection, processing and visualization of data. In addition, it offered the chance to implement changes based on user feedback before transitioning to large-scale testing.

The on-road trials focused on monitoring driving behavior and the impact of real-time interventions (i.e., in-vehicle warnings) and post-trip interventions (i.e., post-trip-feedback & gamification) on driving behavior.

The experimental design of the i-DREAMS on-road study is displayed in Table 1 and has been subdivided into **four consecutive phases**:

- **Phase 1:** monitoring (baseline measurement)
- **Phase 2:** real-time intervention
- **Phase 3:** real-time intervention and post-trip feedback
- **Phase 4:** real-time intervention and post-trip feedback and gamification

It should be noted that in Greece, data from an additional telematics experiment which took place for a 3-month timeframe were collected and analyzed in order to enhance the power of the analyses presented. The experimental design of the i-DREAMS on-road trials for Greece was subdivided into three phases (i.e. phase 1 – monitoring, phase 3 - real-time intervention and post-trip feedback and phase 4 - real-time intervention and post-trip feedback and gamification; while there was no real-time interventions provided by the app (phase 2 was not existed).

Table 1: Description and duration of each Phase

Phases	Description	Duration per participant
Phase 1	Monitoring (baseline measurement; no interventions)	4 weeks
Phase 2	In-vehicle intervention	4 weeks
Phase 3	Post-trip feedback on the smartphone	4 weeks
Phase 4	Post-trip feedback on smartphone + gamified web platform	6 weeks

Firstly, **Phase 1** of the field trials refers to a reference period after the installation of the i-DREAMS system in order to monitor driving behavior without interventions.

Secondly, **Phase 2** of the field trials refers to a monitoring period during which only in-vehicle real-time warnings were provided using adaptive Advanced Driver Assistance Systems (ADAS).

Thirdly, in **Phase 3** of the field trials, feedback via the i-DREAMS smartphone app is combined with in-vehicle warnings.

Lastly, in **Phase 4** of the field trials, gamification features are added to the app, with additional support of a web-dashboard.

In essence, the i-DREAMS project focuses on calibrating the subjective experience of coping capacity and task complexity in driving. The interaction between these concepts is best investigated by applying a combined nudging-coaching approach. This combined approach is used as the **blueprint of the on-road trials' experimental design**.

Figure 2 provides an overview of the different phases of the experimental design of the i-DREAMS on-road study.

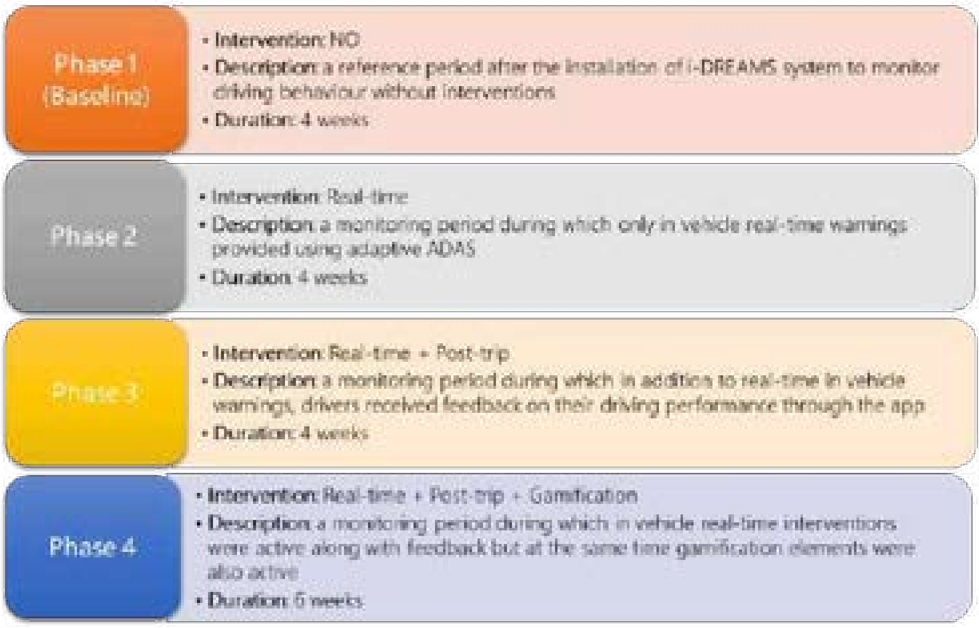

Figure 2: Overview of the different phases of the experimental design of the i-DREAMS on-road study

2.2 Overview of the variable types and technologies

As the key output of the i-DREAMS project is an **integrated set of monitoring** and communication tools for intervention and support, state-of-the-art technologies and systems were utilized in order to monitor driving performance indicators. An OBD-II device supporting all OBD-II protocols is installed in each vehicle. A modern vehicle supports hundreds of parameters, which are recorded by the OBD-II device which accommodates the proper Software Development Kit (SDK) to extract the necessary data as well as a rich set of APIs (Application Programming Interfaces) to communicate with third party systems. This OBD-II integrates 2G or 3G GSM/GPRS technology through which all data recorded from the vehicle through its sensors is transmitted to remote servers (Cloud). The mobile network is used for data transmission without any user involvement.

More specifically, data from the **Mobileye system** (Mobileye, 2022), a **dash camera** and the **Cardio gateway** (CardioID Technologies, 2022) which records driving behavior (e.g., speed, acceleration, deceleration, steering) along with GNSS signals were used. In particular, the Mobileye system is as a network sensor and a camera-based system mounted on the windshield that measures parameters, like headway monitoring, lane position monitoring, traffic sign recognition and pedestrian recognition. The system can be connected to the CAN bus and enables the integration with several ADAS ecosystem products. The Cardio gateway is a system based on sensors which is connected to the Mobileye equipment through the CAN bus of the vehicle and can transfer data through different communication technologies (BLE, CAN, I2C, SPI, WiFi). Information about the current warning stage, as defined by Mobileye, were also collected for comparison with the i-DREAMS warning stage (i.e. normal driving, danger phase, avoidable accident phase). At the same time, information about the current state of the i-DREAMS platform were collected.

The **fundamental challenge within the i-DREAMS project** is how explanatory variables (i.e. various variables performance metrics and indicators of task complexity and coping capacity) are correlated with the dependent variable “risk” in order to predict STZ.

There are three main components of the nature of variables which are used in i-DREAMS:

- **Discrete variables**: variables that are categorical (ordinal or nominal) and can only take discrete values from the real numbers. A few examples of discrete variables in i-DREAMS could be fatigue (yes, no), time of the day (daytime, night time driving) and STZ (normal phase, danger phase, avoidable accident phase).
- **Continuous variables**: variables that can take any values from the real numbers. A few examples of continuous variables in i-DREAMS could be speeding, headway and composite variables, such as weighted sum or weighted average variables.
- **Latent variables**: variables that are not observable to the analyst and so it is not known whether they are continuous or discrete. Examples of latent variables in i-DREAMS are task complexity and coping capacity which are latent explanatory variables and so observable indicators are needed to measure these latent variables. Risk is also initially conceived in i-DREAMS as a latent variable.

Explanatory variables of risk and the most **reliable indicators of coping capacity**, such as average speed, headway, illegal overtaking, harsh accelerations, harsh brakings, distance traveled, duration, forward collision warnings or pedestrian collision warnings will be assessed.

Specifically, the **main risk factors** that will be explored within the i-DREAMS project are:

- Speeding
- Headway
- Overtaking
- Fatigue
- Harsh accelerations
- Harsh brakings
- Vehicle control events (combination of harsh acceleration, braking and cornering events)

Table 2 provides an overview of the variables examined along with their corresponding description.

Table 2: Driving performance indicators of the analyzed data along with their corresponding description (Source: Mobileye, CardioID)

Source	Variable	Description	Unit	Type	Range
	grpby_seconds	Total trip duration	seconds	Integer	
	trip_uuid	Unique ID of the trip		String	
	driver_uuid	Unique driver ID		String	
	vehicle_uuid	Unique ID of the vehicle		String	
	vehicle_class	Vehicle class identifier		String	Car, Bus, Truck, Train, Tram
	trip_start	The trip start date and time in ISO8601 format		String	
	trip_end	The trip stop date and time in ISO8601 format		String	
	Phase	phase of the experiment		Integer	1 - no interventions/monitoring , 2 - real-time warnings, 3 - real-time warnings and post-trip feedback, 4 - real-time warnings and post-trip feedback along with gamification
i-Dreams STZ	iDreams_Headway_Map_level__-1	Real-time headway intervention level -1 **level -1 => no vehicle detected (Normal Driving)**		Integer	0 - intervention level unequal to -1 1 - intervention level equal to -1
	iDreams_Headway_Map_level__0	Real-time headway intervention level 0 **level 0 => vehicle detected, but headway >= 2.5 (Normal Driving)**		Integer	0 - intervention level unequal to 0 1 - intervention level equal to 0
	iDreams_Headway_Map_level__1	Real-time headway intervention level 1 **level 1 => vehicle detected, headway < 2.5, but above warning threshold (Normal Driving)**		Integer	0 - intervention level unequal to 1 1 - intervention level equal to 1
	iDreams_Headway_Map_level__2	Real-time headway intervention level 2 **level 2 => first warning stage (Dangerous Driving)**		Integer	0 - intervention level unequal to 2 1 - intervention level equal to 2
	iDreams_Headway_Map_level__3	Real-time headway intervention level 3 **level 3 => second warning stage (Avoidable Accident)**		Integer	0 - intervention level unequal to 3 1 - intervention level equal to 3
	iDreams_Overtaking_Map_level__0	Real-time overtaking intervention level 0 **level 0 => no warning (Normal Driving)**		Integer	0 - intervention level unequal to 0 1 - intervention level equal to 0
	iDreams_Overtaking_Map_level__1	Real-time overtaking intervention level 1 **level 1 => visual warning (Normal Driving)**		Integer	0 - intervention level unequal to 1 1 - intervention level equal to 1

Source	Variable	Description	Unit	Type	Range
	iDreams_Overtaking_Map_level__2	Real-time overtaking intervention level 2 **level 2 => visual and auditory warning (Dangerous Driving)**		Integer	0 - intervention level unequal to 2 1 - intervention level equal to 2
	iDreams_Overtaking_Map_level__3	Real-time overtaking intervention level 3 **level 3 => frequent warning (Avoidable Accident)**		Integer	0 - intervention level unequal to 3 1 - intervention level equal to 3
	iDreams_Speeding_Map_level__0	Real-time speeding intervention level 0 **level 0 => no warning (Normal Driving)**		Integer	0 - intervention level unequal to 0 1 - intervention level equal to 0
	iDreams_Speeding_Map_level__1	Real-time speeding intervention level 1 **level 1 => visual indication (Normal Driving)**		Integer	0 - intervention level unequal to 1 1 - intervention level equal to 1
	iDreams_Speeding_Map_level__2	Real-time speeding intervention level 2 **level 2 => visual speeding warning (Dangerous Driving)**		Integer	0 - intervention level unequal to 2 1 - intervention level equal to 2
	iDreams_Speeding_Map_level__3	Real-time speeding intervention level 3 **level 3 => visual and auditory warning (Avoidable Accident)**		Integer	0 - intervention level unequal to 3 1 - intervention level equal to 3
	iDreams_Fatigue_Map_level__0	Real-time fatigue intervention level 0 **level 0 => no warning (Normal Driving)**		Integer	0 - intervention level unequal to 0 1 - intervention level equal to 0
	iDreams_Fatigue_Map_level__1	Real-time fatigue intervention level 1 **level 1 => visual warning (Dangerous Driving)**		Integer	0 - intervention level unequal to 1 1 - intervention level equal to 1
	iDreams_Fatigue_Map_level__2	Real-time fatigue intervention level 2 **level 2 => visual and auditory warning (Dangerous Driving)**		Integer	0 - intervention level unequal to 2 1 - intervention level equal to 2
	iDreams_Fatigue_Map_level__3	Real-time fatigue intervention level 3 **level 3 => frequent warnings (Dangerous Driving)**		Integer	0 - intervention level unequal to 3 1 - intervention level equal to 3
Gateway IMU	DrivingEvents_Map_lvl__H	H - High event (harsh acceleration, harsh braking, and harsh cornering) severity level		String	0 - high event severity level not detected 1 - high event severity level detected
	DrivingEvents_Map_lvl__L	L - Low event (harsh acceleration, harsh braking, and harsh cornering) severity level		String	0 - low event severity level not detected 1 - low event severity level detected
	DrivingEvents_Map_lvl__M	M - Medium event (harsh acceleration, harsh braking, and harsh cornering) severity level		String	0 - medium event severity level not detected 1 - medium event severity level detected

Source	Variable	Description	Unit	Type	Range
	DrivingEvents_Map_evt__ha	Type of event - harsh acceleration: 'ha'		String	0 - harsh acceleration not detected 1 - harsh acceleration detected
	DrivingEvents_Map_evt__hb	Type of event - harsh braking: 'hb'		String	0 - harsh braking not detected 1 - harsh braking detected
	DrivingEvents_Map_evt__hc	Type of event - harsh cornering: 'hc'		String	0 - harsh cornering not detected 1 - harsh cornering detected
	IBI_value	Time interval between successive heart beats	milliseconds	Integer	
Mobileye	ME_Car_speed	Vehicle speed	km/h	Integer	
	ME_Car_wipers	Wipers		Boolean	0 - missing values False - Wipers are off, True - Wipers are on
	ME_Car_high_beam	High-beam		Boolean	0 - missing values False - High-beam is off True - High-beam is on
	ME_AWS_hw_measurement	Headway measurement	seconds	Float	
	ME_AWS_tsr_level	Traffic sign recognition level	km/h or mp/h	Integer	0 - no warning, 1 - 0-5 units over speed limit, 2 - 5-10 units over speed limit, 3 - 10-15 units over speed limit, 4 - 15-20 units over speed limit, 5 - 20-25 units over speed limit, 6 - 25-30 units over speed limit, 7 - 30+ units over speed limit
	ME_AWS_fcw	Forward collision warning		Boolean	0 - missing values False - Forward collision warning is inactive True - Forward collision warning is active
	ME_AWS_ldw	Lane departure warning		Boolean	0 - missing values False - Lane departure warning is inactive True - Lane departure warning is active (left or right)
	ME_AWS_pcw	Pedestrian collision warning		Boolean	0 - missing values False - Pedestrian collision warning is inactive True - Pedestrian collision warning is active

Source	Variable	Description	Unit	Type	Range
	ME_AWS_pedestrian_dz	Pedestrian in danger zone		Boolean	0 - missing values False - Pedestrian not detected in danger zone True - Pedestrian detected in danger zone
	ME_AWS_time_indicator	Indicates lighting conditions		String	1 - day, 2 - dusk, 3 - night
	ME_TSR_tsr_1_speed	Display 1 speed traffic sign code		Integer	
	GPS_spd	Speed	km/h	Float	
	GPS_distances	Total trip distance	km	Float	
	ME_LDW_Map_type_L_mean	Left lane departure warning		Boolean	0 - missing values False - Left lane departure warning is inactive True - Left lane departure warning is active
	ME_LDW_Map_type_R_mean	Right lane departure warning		Boolean	0 - missing values False - Right lane departure warning is inactive True - Right lane departure warning is active

2.3 Questionnaires

In addition to the vehicle data, questionnaire data were also collected both before and after the trial. The **number of participants** who answered both for the entry and exit questionnaires and for which data was available at the time of writing this deliverable is shown below. It should be noted that at the time of writing this deliverable, the questionnaire data from Portugal were not completed yet; thus, questionnaire data for buses have not been included in the analysis.

- 45 car drivers in Belgium
- 23 truck drivers in Belgium
- 54 car drivers in UK
- 28 car drivers in Germany
- 65 car drivers in Greece

The full questionnaires are given in i-DREAMS Deliverable 7.2 in Annex 2 (Brown et al., 2023).
Information collected pre-trial included:

- **Screening questionnaire**: driver details (age, gender, driving experience, employment status, etc.), vehicle details (model, age, etc.).
- **Entry questionnaire**: current use of and opinions on different ADAS, driving style and confidence, opinions on driving and safety, self-assessment of driver's risk-taking behaviors (e.g., speeding, mobile phone use), crash and offence history, sleepiness and driving, medical conditions.

Information collected post-trial included:

- **User experience questionnaire**: opinions on the i-DREAMS system - except for Greece, in which an alternative driving experiment without the use of i-DREAMS in-vehicle system was used - (ease of use, works as described), opinions on the i-DREAMS smartphone app (ease of use, usefulness).
- **Exit questionnaire**: opinions on the i-DREAMS system (improvement of driving, usefulness, trust, clarity of warnings, etc.), experience of driving situations, driver behavior (driving and non-driving related behaviors), overall experience rating.

In particular, a set of 12 questions were asked identically at both trial entry and trial exit (respectively EQ11 and EX3 in Annex 2 of Deliverable 7.2), to allow analysis of before and after responses. These questions related to the areas of perceived knowledge, self-efficacy, attitude, personal norm, and subjective norm. The theory used in the development of these questions is described in more detail in i-DREAMS Deliverable 7.1 (Katrakazas et al., 2020).

2.4 Aggregation and cleaning

In the transportation research domain, traffic data used for behavior prediction or safety assessment are usually aggregated (Abdel-Aty et al., 2005, Franke and Krems, 2013) in order for post-trip or post-event interventions to be applied. At the same time, real-time applications (Habtemichael et al., 2012, Vlahogianni and Barmpounakis, 2017) demand the use of highly disaggregated or time-series data, in order to identify different behaviors or critical events in the future.

Highly disaggregated data which describe all the available driving performance indicators, such as average speed, headway, harsh acceleration or harsh braking were collected. A methodological framework was employed in which data were aggregated in **30-seconds or 60-seconds intervals** and the mean and standard deviations of the aforementioned kinematic characteristics were extracted. It should be noted that the aforementioned intervals have been also utilized in previous traffic safety studies (Katrakazas et al., 2019).

The most crucial step in the data aggregation and cleaning was to identify **NA values and remove validly the missing data** from the dataset. Then, a basic procedure was followed for each type of variable. There are two different types of indicators that appear in the data: level-type variables and continuous variables. “Level-type” variables include the speeding, headway measurements, overtaking, fatigue and harsh events. The latter appear in a categorization of high, medium, and low events, but also as harsh braking, harsh acceleration, and harsh cornering events.

With regards to headway, overtaking, speeding and fatigue **levels**, for the trips that had at least one value per aggregation row, the remaining levels were imputed with 0. For instance, in case there were valid values for 2 (out of 4) levels and values for the 3rd and 4th level were NAs, an imputation with 0 in the remaining levels was made. In the case where there were NA values for all levels, a replacement of NA values with -9999 value was made. Afterwards, a check per each aggregation row was implemented to ensure the accuracy and the validity of the data aggregation approach. As the aggregated variables were added in the form of mean and sum, the summary of each aggregation row should be equal to 1 in the case of the mean and equal to 30 in the case of the sum (30s aggregation level). Similarly, a check per each aggregation row was implemented in order to ensure the accuracy and the validity of the data aggregation in the case of harsh events and the summary of each aggregation row for the aforementioned variables should be equal to the corresponding variable in total (low + medium + high).

Lastly, as per “**continuous”** variables, such as speed, distance, headway, forward collision warning, pedestrian collision warning, etc, the replacement of NA values was done by the imputation with the mean or median value of the corresponding variable per trip.

2.5 Variables used

After an extensive data cleaning and preparation, the next step of the analysis involved a collinearity testing so that any highly correlated variables were excluded from the models. When two variables have an absolute value of correlation coefficient at least 0.6, then these two variables are highly correlated. The **most appropriate variables were selected** to be included in the GLM and SEM analysis, using either correlation or feature selection algorithms.

2.5.1 Definition of task complexity and coping capacity

The cornerstone of the i-DREAMS platform is the assessment of task complexity and coping capacity. **Task complexity** relates to the current status of the real world context in which a vehicle is being operated. Since this context is consistent of various individual elements which, together, determine the complexity of the task imposed on the vehicle operator, a multi-dimensional approach in further operationalizing this concept is adopted. In particular, task complexity context is monitored via registration of road layout (i.e. highway, rural, urban), time and location, traffic volumes (i.e. high, medium, low) and weather.

As for **coping capacity**, Figure 3 shows that this concept is dependent upon two underlying factors and it consists of several aspects of both vehicle and operator state. These are also multi-dimensional in nature.

More specifically, the latent variables associated to "**vehicle state**" are estimated on the basis of various metrics. The factor 'vehicle' entails three aspects, as shown below:

- **Technical specifications**, measured on the basis of average speed, braking power, acceleration performance, etc.
- **Actuators & admitted actions**, measured on the basis of accelerator, brakes, steering wheel, etc.
- **Current status**, measured on the basis of fuel efficiency, schedule maintenance), real-time information either from on board systems (OBD II, FMS, Tachometer), Telematics/GPS, or smartphone, or additional information coming from ADAS systems - (headway & collision monitoring, pedestrian warning, lane keeping monitoring, on board cameras, etc.

Additionally, the latent variables associated to "**operator state**" are estimated on the basis of various metrics. The factor 'operator' entails six aspects, as shown below:

- **Mental state**, measured on the basis of metrics on alertness, attention, emotions, etc.
- **Behavior**, measured on the basis of metrics such as speeding, harsh acceleration / braking / cornering, seat belt use etc.
- **Competencies**, measured on the basis of metrics on risk assessment, attention regulation, self-appraisal, etc.
- **Personality**, measured on the basis of metrics on adventure seeking, disinhibition, experience seeking, boredom susceptibility, etc.
- **Sociodemographic profile**, measured on the basis of age, gender, experience, socio-economic status, nationality, ethnicity, cultural identity, etc.
- **Health status**, measured on the basis of metrics on current symptoms, neurologic and cardiovascular indicators, medication, etc.

As already outlined, coping capacity is not only dependent upon the status of the operator, but of the vehicle as well. Each of these operator- and vehicle-related aspects can be further operationalized by a combination of different variables, as shown in Figure 3.

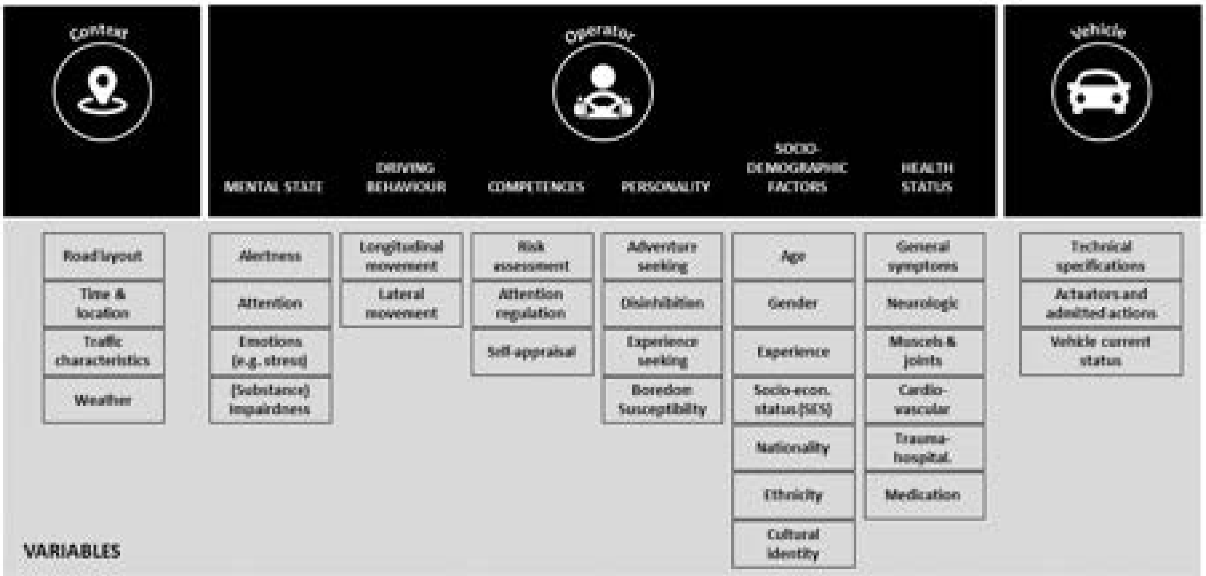

Figure 3: Monitoring context, operator & vehicle: an illustrative canvas

According to the i-DREAMS concept of a context-aware Safety Tolerance Zone, ‘risk’ results from the interaction of ‘task complexity’ and ‘coping capacity’. However, these three core aspects are **unobserved / latent variables**, which cannot be measured directly, but can be estimated on the basis of various metrics. Based on the abovementioned, task complexity as a latent variable can be measured by metrics and indicators related to the road environment. Coping capacity is also a latent variable, including two distinct aspects, each one being a latent variable itself. These are vehicle state and operator state. Risk as a latent variable can be measured by indicators such as danger phase events and avoidable accident events, as detected by the Safety Tolerance Zone monitor. Latent variables analysis will be performed on the basis of dedicated techniques such as Structural Equation Modeling.

Figure 4 illustrates the conceptual framework of the i-DREAMS platform for the prediction of risk in function of coping capacity and task complexity.

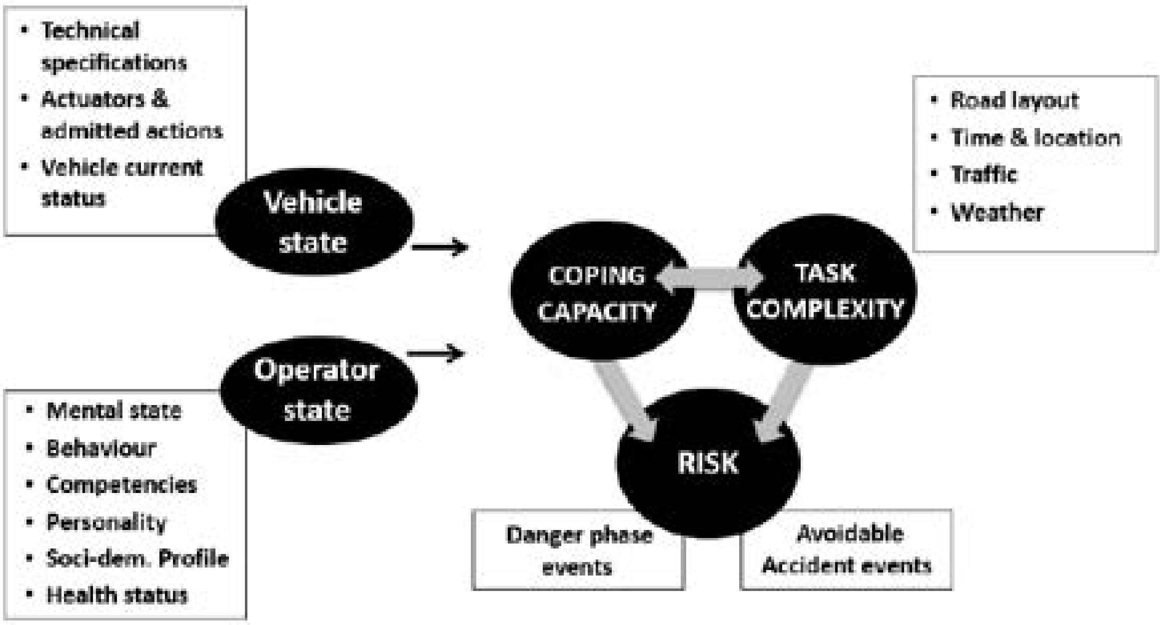

Figure 4: Post-hoc prediction of risk in function of coping capacity and task complexity

2.5.2 Variables used to define task complexity and coping capacity

The most appropriate variables which were used in order to define task complexity and coping capacity (vehicle and operator state) along with the variables that were finally utilized to represent risk are shown in Table 3.

With regards to car wipers, considered as an indicator of weather conditions, can be used to clear rain, snow, or debris from the windshield of a vehicle, which are all common weather-related hazards. The speed at which the wipers move can also indicate the intensity of the precipitation or debris. For instance, if the wipers are moving very fast, it may indicate heavy rain or snow. On the other hand, if the wipers are moving slowly, it could mean that there is only light precipitation. Overall, car wipers are an important safety feature of a vehicle and can help drivers navigate through different weather conditions.

In addition, high beam headlights are considered an indicator of lighting conditions as they are used to provide maximum illumination when driving in low light or dark conditions. The high beam headlights are designed to project a beam of light further down the road, which can help drivers to see obstacles or pedestrians that may be difficult to see with low beam headlights. Overall, high beam headlights are an important feature of a vehicle that can help drivers navigate through different lighting conditions.

Table 3: Variables for task complexity and coping capacity (vehicle and operator state) and risk

Task complexity	Coping capacity - vehicle state	Coping capacity - operator state		Risk
Car wipers	Vehicle age	Distance	Inter Beat Interval	Headway map levels
Car high beam	First vehicle registration	Duration	Headway	Speeding map levels
Time indicator	Fuel type	Average speed	Overtaking	Overtaking map levels
Distance	Engine Cubic Centimetres	Harsh acceleration/braking	Fatigue	Fatigue map levels
Duration	Engine Horsepower (HP)	Forward collision warning (FCW)	Gender	Harsh acceleration
Month	Gearbox	Pedestrian collision warning (PCW)	Age	Harsh braking
Day of the week	Vehicle brand	Lane departure warning (LDW)	Educational level	Vehicle control events

2.5.3 Descriptive statistics

Descriptive statistics for the available parameters in database used for the different countries (i.e. Belgium, UK, Germany, Greece and Portugal) and transport modes (i.e. cars, trucks and buses) per each phase are presented in Annex 1.

3 Analysis

3.1 Purpose of the analysis

There are two main purposes for data analysis in i-DREAMS, prediction and explanatory analysis, and the type of analytical methods to be used depends on these purposes:

- **Prediction** is mostly done to identify (in real-time) the level of the STZ at which the driver is, and in order to trigger real-time in-vehicle interventions.
- **Explanatory analysis** is mostly done to identify the relationship between risk and factors contributing to risk. This relationship may help better understand the underlying reasons of driving behavior and ultimately help improve interventions (both in-vehicle and post trip). In addition, understanding the effects of explanatory variables on risk may also help evaluating the effectiveness of interventions.

Another dimension of data analysis in i-DREAMS is the temporal element of data analysis:

- **Real-time**: The collected data may be analysed in real-time (e.g. large amount of data, time series nature of real-time data)
- **Post-trip**: The collected data may be aggregated and analysed after the trip has been completed.

Proper analytical methods have been used to capture the unique properties of data in both cases. However, it is noted that, while it seems intuitive that real-time data analysis corresponds to the **prediction purpose**, and post-trip data analysis corresponds to the explanatory analysis purpose, it may be worth investigating whether there are additional combinations applicable within the scope of i-DREAMS.

It should be mentioned that the **analytical models for STZ identification** have already been described in previous project Deliverables 3.2 (Katrakazas et al., 2020) and 4.2 (Yang et al., 2020). In summary, Dynamic Bayesian Networks (DBNs), Long-Short-Term-Memory networks (LSTMs), as well as Discrete Choice Models (DCM) and Structural Equation Models (SEM) can be used for STZ identification and explanation of measurement impacts. Furthermore, a plethora of analytical tools have been already documented in order to be able to predict or explain safety risk and the impact of interventions.

A schematic overview of the proposed mathematical models (DBN, LSTM, DCM and SEM) to be considered for the analysis is given in Figure 5.

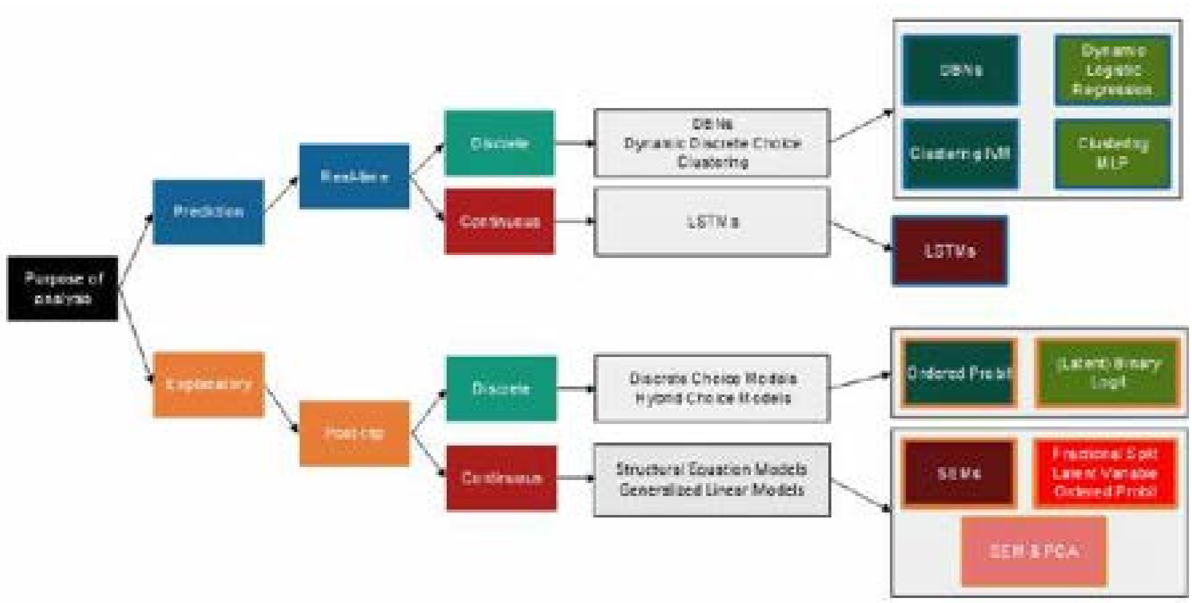

Figure 5: Schematic overview of modeling approaches considered for the analysis of risk factors

Following the **Big Data analysis** and processing carried out in previous Deliverables, the processed data analysis methods include two families of techniques:

- **Multivariate regression analysis** (e.g. Generalized Linear Models) for exploratory analysis in order to identify the key correlations between observed metrics while controlling for the differences between the sample groups.
- **Latent variables analysis** (e.g. Structural Equation Models) for latent analysis in order to quantify the effects between latent and observable variables of task complexity and coping capacity with complex relationships.

3.2 Generalized Linear Models (GLMs)

In statistics, the **Generalized Linear Model** (GLM) is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution. The GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value (Hastie and Pregibon, 2017).

In a generalized linear model (GLM), each outcome Y of the dependent variables is assumed to be generated from a particular distribution in an exponential family, a large class of probability distributions that includes the **normal, binomial, Poisson and gamma distributions**, among others. The mean, μ, of the distribution depends on the independent variables, X, through:

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

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 this framework, the variance is typically a function, V, of the mean:

$$Var(Y|X) = V(g^{-1}(X\beta)) \quad (2)$$

It is convenient if V follows from an exponential family of distributions, but it may simply be that the variance is a function of the predicted value.

The unknown parameters, β, are typically estimated with maximum likelihood, maximum quasi-likelihood, or Bayesian techniques.

GLMs were formulated as a **way of unifying various other statistical models**, including linear regression, logistic regression and Poisson regression. In particular, Hastie and Tibshirani (1990) proposed an iteratively reweighted least squares method for maximum likelihood estimation of the model parameters. Maximum-likelihood estimation remains popular and is the default method on many statistical computing packages. Other approaches, including Bayesian approaches and least squares fits to variance stabilized responses, have been developed.

A key point in the development of GLM was the **generalization of the normal distribution** (on which the linear regression model relies) to the exponential family of distributions. This idea was developed by Collins et al. (2001). Consider a single random variable y whose probability (mass) function (if it is discrete) or probability density function (if it is continuous) depends on a single parameter θ. The distribution belongs to the exponential family if it can be written as follows:

$$f(y;\theta) = s(y)t(\theta)e^{a(y)b(\theta)} \quad (3)$$

where: a, b, s, and t are known functions. The symmetry between y and θ becomes more evident if the equation above is rewritten as follows:

$$f(y;\theta) = \exp[\alpha(y)b(\theta) + c(\theta) + d(y)] \quad (4)$$

where: s(y)=exp[d(y)] and t(θ)=exp[c(θ)]

If a(y) =y then the distribution is said to be in the canonical form. Furthermore, any additional parameters (besides the parameter of interest θ) are regarded as nuisance parameters forming parts of the functions a, b, c, and d, and they are treated as though they were known. Many well-known distributions belong to the **exponential family**, including Poisson, normal or binomial distributions. On the other hand, examples of well-known and widely used distributions that cannot be expressed in this form are the student’s t-distribution and the uniform distribution.

It should be mentioned that the **Variance Inflation Factor (VIF)** is a measure of the amount of multicollinearity in regression analysis. Multicollinearity exists when there is a correlation between multiple independent variables in a multiple regression model. The default VIF cutoff value is 5; only variables with a VIF less than 5 will be included in the model (VIF<5). However, in certain cases, even if VIF is less than 10, then it can be accepted.

3.3 Structural Equation Models (SEMs)

Structural Equation Modelling (SEM) is widely used for modelling complex and multi-layered relationships between observed and unobserved variables, such as ‘task complexity’, ‘coping capacity’ etc. Observed variables are measurable, whereas unobserved variables are latent constructs – analogous to factors / components in a factor / principal component analysis.

Structural equation models have two components: a measurement model and a structural model. The measurement model is used to determine how well various observable exogenous variables can measure (i.e. load on) the latent variables, as well as the related measurement errors. The structural model is used to explore how the model variables are inter-related, allowing for both direct and indirect relationships to be modelled. In this sense, SEMs differ from ordinary regression techniques in which relationships between variables are direct.

The general formulation of SEM is as follows (Washington et al., 2011; 2020):

$$\eta = \beta\eta + \gamma\xi + \varepsilon \qquad (5)$$

where η is a vector of endogenous variables, ξ is a vector of exogenous variables, β and γ are vectors of coefficients to be estimated, and ε is a vector of regression errors.

The measurement models are then as follows (Chen, 2007):

$x = \Lambda_x \xi + \delta$, for the exogenous variables (6)

$y = \Lambda_y \eta + \zeta$, for the endogenous variables (7)

where x and δ are vectors related to the observed exogenous variables and their errors, y and ζ are vectors related to the observed endogenous variables and their errors, and Λ_x, Λ_y are structural coefficient matrices for the effects of the latent exogenous and endogenous variables on the observed variables.

The structural model is often represented by a path analysis, showing how a set of ‘explanatory’ variables can influence a ‘dependent’ variable. The paths can be drawn so as to reflect whether the explanatory variables are correlated causes, mediated causes, or independent causes to the dependent variable.

3.4 Neural Networks (NNs)

An Artificial Neural Network (ANN) is a highly complex, non-linear, parallel processor with a natural propensity for storing experimental knowledge and making it available afterward. A multi-layer perceptron ANN is typically made up of three kinds of layers: an input layer, an output layer, and one or more hidden layers. The input layer receives the values of the explanatory variables, i.e., the input data. The hidden layer, made up of m neurons, adds up the weights of the input values of the various explanatory variables, and calculates the complex association patterns. With regards to the hidden layer, **activation function** applies a non-linear map to the linear transformation of input values, introducing nonlinearity into the model. A single hidden layer is usually enough for crash analysis applications, but the definition of the number of neurons in it is generally the object of experimentation;. For the output layer, the values of the various hidden neurons are summed and the network's output values are presented (Garefalakis et al., 2022; Silva et al., 2020).

3.5 Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory Models (LSTMs) are a special kind of RNN, capable of learning long-term dependencies (Girma et al., 2019). They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior and not something they struggle to learn. All recurrent LSTMs have the form of a chain of repeating modules of neural network.

LSTMs use ”memory block” in the hidden unit to capture the long-term dependencies that may exist in the data (Girma et al., 2019). This memorizing capability of LSTM has shown the best performance across many time-series tasks, such as activity recognition, video captioning, language translation. The cell state (memory block) of LSTM has one or more memory cells that are regulated by structures called gates, which control the addition of new sequential information and the removal of useless ones to and from memory, respectively. Gates are a combination of sigmoid activation functions and an element-wise multiplication or Hadamard product and they are used to control information that passes through the network. An LSTM is often composed by three gates, namely forget, input, and output gates, which are described below:

- **Forget gate:** Forget gate decides what information to keep or remove from the cell state. The first step in LSTM is to decide what information are going to throw away from the cell state. This decision is made by a sigmoid layer called the “forget gate layer.”
- **Input gate**: Input gate decides what new information to add and how to update the old cell state, $C_{t-1,}$ to the new cell state C_t for the next memory block. This has two parts. First, a sigmoid layer called the “input gate layer” decides which values we’ll update. Next, a *tanh* layer creates a vector of new candidate values, $C_{t'}$, that could be added to the state. Then the old cell state C_{t-1} updates into the new cell state C_t and the old state is multiplied by f_t.
- **Output gate:** Output gate filters out and decides which information to produce as an output from a memory block at a given time step t. This output will be based on cell state, but will be a filtered version. First, a sigmoid layer, which decides what parts of the cell state are going to output, is run. Then, the cell state, used as tanh (to push the values to be between −1 and 1) and multiply it by the output of the sigmoid gate, in order to take and output the parts needed.

3.6 Grouped Random Parameters Binary Logit (GRPL) Models

Binary Logit Discrete Choice Models have been widely used to correlate a binary dependent variable with explanatory variables (Hensher et al., 2005). These models assume that effects of explanatory variables are fixed across the sample. However, this assumption may not always hold and the effects of explanatory variables may vary across individuals due to unobserved heterogeneity (Hensher and Greene, 2003). In addition, the empirical data in this study contain multiple observations for each driver (multiple near-misses per trip for each driver) creating several panels in the data. The grouped random parameters logit model has been used in the literature to address the above limitations of the simple binary logit model (Afghari et al, 2022) and thus is used in this study to model the binary near-miss indicator. The specification of this model is briefly presented in the following.

Let Y_{it} be a binary dependent variable representing a near-miss (Y_{it}=1: near-miss, Y_{it}=0: no event) of the i^{th} driver at time *t*. Assuming a random utility theory (Hensher et al., 2005), the utility of near-miss for this driver (U_{it}) is stated as:

$$U_{it} = \beta_i X_{it} + \varepsilon_{it} \quad (8)$$

where β_i are estimable parameters (including the intercept), X_{it} are explanatory variables and ε_{it} is the random error term assumed to be identically and independently distributed across observations and describing the random part of the utility. Assuming that ε_{it} is generalized extreme value distributed (Mcfadden, 1980), the probability of a near-miss can be presented as:

$$P(Y_{it} = 1) = \frac{1}{1+e^{-(\beta_i X_{it})}} \qquad (9)$$

Note that the estimable parameters are allowed to vary across individuals to account for unobserved heterogeneity in the data. However, the parameters are fixed across multiple observations of the same individual, accounting for the panel nature of the data. This model is referred to as the *grouped random parameters model* in the literature (Afghari et al, 2022). The likelihood of having a near-miss across all individuals can then be determined by the product of the above equation over the entire observations.

3.7 Ordered Probit Fractional Split (OPFS) Models

Ordered discrete choice models are proper analytical models for this type of risk indicator (Washington et al., 2020). However, these models have the implicit assumption that one outcome category may be selected at a time (Hensher et al., 2005). Such an assumption does not hold for modelling speeding behavior of drivers over a defined period of time (e.g. 1 minute) because multiple speeding categories may occur during this time. For example, whilst 35% of this 1-minute window may consist of 1st STZ, another 45% may consist of the 2nd and the remaining 20% may consist of the 3rd STZ level for speeding. Thus, the discrete outcome during this time window is not binary anymore, and the conventional discrete outcome models are not suitable. An alternative modelling approach in such circumstances is fractional response modelling where the outcome variable is a fraction (proportion) summing to unity across all categories (Afghari et al., 2018).

Let Y_{it} be the actual proportion of speeding STZ levels that driver i commits during time interval t (e.g. 1-minute intervals); and let s (s = 1, 2, 3) represent speeding STZ categories (i.e. STZ1, STZ2, and STZ3) during 1-minute intervals. In ordered models, the actual proportion of STZ levels (Y_{it}) is associated with an underlying latent variable (y^*_{it}). This latent variable is then mapped to the actual STZ proportions by thresholds (τ) and using the following linear function:

$$Y^*_{it} = \kappa X_{it} + \delta_i \quad \text{and} \quad Y_{sit} = S \quad \text{if} \quad \tau_{s-1} < Y^*_{it} < \tau_s \qquad (10)$$

where κ is the vector of parameters, X_{it} is the vector of covariates and δ_i is the random error term. To estimate the latent propensity of STZ proportions, it is assumed that:

$$E(Y_{sit} \mid X_{it}) = H_{sit}(.) \,, \quad 0 \le H_{sit}(.) \le 1 \,, \quad \sum_{s=1}^{S} H_{sit} = 1 \qquad (11)$$

where $H_{sit}(.)$ is the probability density function for the STZ category s. Depending on the distributional assumption for the probability of error terms, $H_{sit}(.)$ can take standard normal or standard logistic probability density functions for the ordered probit or ordered logit models, respectively. The former functional form is used in this section to construct an ordered probit model for speeding STZ. The probability of each STZ category is then presented as:

$$P(Y_{sit} = s) = \varphi\{\tau_s - (\kappa X_{it})\} - \varphi\{\tau_{s-1} - (\kappa X_{it})\} \tag{12}$$

where $\varphi(.)$ is the standard normal cumulative probability density function. The corresponding quasi log-likelihood function is then expressed as:

$$LL = \sum_{i=1}^{N} Log(\int_{\kappa} \prod_{s=1}^{S} P(Y_{sit} = s)^{w_{sit}} d\kappa) \tag{13}$$

where w_{sit} is the fraction (proportion between 0 and 1) of STZ category s for driver i and time period t, and the rest of notations are as previously stated. These fractions sum to unity over the categories ($\sum_{s=1}^{S} w_{sit} = 1$). This model is referred to as ordered fractional split (Afghari et al., 2020). Note that w_{sit} takes binary values (0 or 1) in conventional choice models; one for the chosen alternative and zero for the non-chosen alternative. Maximum likelihood approach is used to estimate this log-likelihood function.

3.8 Model goodness-of-fit measures

In the context of model selection, **model Goodness-of-Fit measures** consist an important part of any statistical model assessment. Several goodness-of-fit metrics are commonly used, including the goodness-of-fit index (GFI), the (standardized) Root Mean Square Error Approximation (RMSEA), the comparative fit index (CFI) and the Tucker-Lewis Index (TLI). Such criteria are based on differences between the observed and modelled variance-covariance matrices. A detailed description of the aforementioned metrics is presented below:

The **Akaike Information Criterion (AIC),** which accounts for the number of included independent variables, is used for the process of model selection between models with different combination of explanatory variables (Vrieze, 2012).

$$AIC = -2L(\theta) + q \tag{14}$$

where: q is the number of parameters and L(θ) is the log-likelihood at convergence. Lower values of AIC are preferred to higher values because higher values of -2L(θ) correspond to greater lack of fit.

The **Bayesian Information Criterion (BIC)** is used for model selection among a finite set of models; models with lower BIC are generally preferred.

$$BIC = -2L(\theta) + q\ln(N) \tag{15}$$

The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) provide measures of model performance that account for model complexity. AIC and BIC combine a term reflecting how well the model fits the data with a term that penalizes the model in proportion to its number of parameters.

The **Comparative Fit Index (CFI)** is based on a noncentral x^2 distribution. It evaluates the model fit by comparing the fit of a hypothesized model with that of an independence model. The values of CFI range from 0 to 1, indicating a good fit for the model when the value exceeds 0.95 (Lee and Sohn, 2022). In general, values more than 0.90 for CFI are generally accepted as indications of very good overall model fit (CFI>0.90). The formula is represented as follows:

$$CFI = 1 - \frac{\max({x_H}^2 - df_H, 0)}{\max({x_H}^2 - df_H, {x_I}^2 - df_I)} \tag{16}$$

where: x^2_H is the value of x^2 and df_H is degrees of freedom in the hypothesized model, and x^2_I is the value of x^2 and df_I is the degrees of freedom in the independence model.

The **Tucker Lewis Index (TLI)** considers the parsimony of the model. Therefore, if the fit indices of two models are similar, a simpler model (i.e. greater degrees of freedom) is chosen. TLISI is an unstandardized value, so it can have a value less than 0 or greater than 1. It indicates a good fit for the model when the value exceeds 0.95 (Lee and Sohn, 2022). In general, values more than 0.90 for TLI are generally accepted as indications of very good overall model fit (TLI>0.90). The formula is represented as follows:

$$TLI = \frac{\frac{{x_I}^2}{df_I} - \frac{{x_H}^2}{df_H}}{\frac{{x_I}^2}{df_I} - 1} \tag{17}$$

where: x^2_H is the value of x^2 and df_H is the degrees of freedom in the hypothesized model, and x^2_I is the value of x^2 and df_I is the degrees of freedom in the independence model.

Currently, one of the most widely used goodness-of-fit indices is the **Root Mean Square Error Approximation (RMSEA)**. RMSEA measures the unstandardized discrepancy between the population and the fitted model, adjusted by its degrees of freedom (df). Different proposals have been made as to the correct use of RMSEA. The most common approach is to calculate and interpret the sample's RMSEA (McDonald and Ho, 2002). RMSEA is considered a "badness-of-fit measure," meaning that lower index values represent a better-fitting model. RMSEA index ranges between 0 and 1. Its value 0.05 or lower is indicative of model fit with observed data. P close value tests the null hypothesis that RMSEA is no greater than 0.05. If P close value is more than 0.05, the null hypothesis is accepted that RMSEA is no greater than 0.05 and it indicates the model is closely fitting the observed data (RMSEA<0.05). The formula is represented as follows:

$$RMSEA = \sqrt{\frac{{x_H}^2 - df_H}{df_H(n-1)}} \tag{18}$$

where: x^2_H is the value of x^2 and df_H is the degrees of freedom in the hypothesized model; n is the sample size.

The **Root Mean Squared Error (RMSE)** is one of the most commonly used measures for evaluating the quality of predictions. It shows how far predictions fall from measured true values using Euclidean distance.

The formula of RMSE, which is the square root of the average squared error, is represented as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum e_t^2} \tag{19}$$

where: N is the number of forecasted points, and e_t is the error (i.e. $observed_t - forecasted_t$)

The **Goodness of Fit Index (GFI)** is a measure of fit between the hypothesized model and the observed covariance matrix. The adjusted goodness of fit index (AGFI) corrects the GFI, which is affected by the number of indicators of each latent variable (Baumgartner and Hombur, 1996). The GFI and AGFI range between 0 and 1, with a value of over 0.9 generally indicating acceptable model fit. In general, values more than 0.90 for GFI are generally accepted as indications of very good overall model fit (GFI>0.90).

Lastly, the **Hoelter** index is calculated to find if chi-square is insignificant or not. If its value is more then 200 for the model then model is considered to be good fit with observed data (Hoelter>200). Values of less than 75 indicate very poor model fit. The Hoelter only makes sense to interpret if N > 200 and the chi square is statistically significant.

For the classification models the confusion matrix and the corresponding metrics will be utilized. In order to compare the classification performance of the several configurations (hyperparameters and mix of considered inputs), well-established machine learning error metrics were calculated. The following metrics were utilized, based on the **confusion matrix**, which provides True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) metrics. The classification algorithms are evaluated using the **accuracy, precision, recall, f1-score, and false alarm rate** as defined below.

Accuracy, which represents the proportion of correctly classified observations, is defined as:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{20}$$

Precision, which quantifies the number of positive class predictions that actually belong to the positive class, is defined as follows:

$$Precision = \frac{TP}{TP+FP} \tag{21}$$

Recall, also known as True Positive Rate, is defined as follows:

$$Recall = \frac{TP}{TP+FN} \tag{22}$$

F1score, which combines precision and recall into a single measure, is defined as follows:

$$f1 - score = \frac{2x\ (Precision)x\ (Recall)}{(Precision)+(Recall)} \tag{23}$$

False alarm rate is defined as follows:

$$False\ Alarm\ Rate = \frac{FP}{FP+TN} \tag{24}$$

4 Synthesis of risk factors

4.1 Generalized Linear Models

A high number of regression model tests were conducted for **different combinations of variables**. An attempt was made to use the same independent variables in the model applied. For each configuration, various alternatives were tested through the respective log-likelihood test comparisons. The optimal combination of variables was the one that had a sufficient number of statistically significant independent variables at a 95% confidence level (p-values ≤ 0.05).

In order to ensure that the results are reliable, accurate, and not biased by chance, it is important to account for **chance capitalization**, which refers to the possibility of obtaining significant results simply by chance, especially when testing a large number of variables. This can be done by adjusting the significance level or using methods such as Bonferroni or False Discovery Rate (FDR) correction to account for multiple comparisons. In this analysis, the Bonferroni correction was used that involves dividing the desired level of significance by the number of tests being conducted. This approach can be conservative, as it reduces the chance of false positives but also decreases the power of the test.

Moreover, the independent variables were also checked for multicollinearity through the Variance Inflation Factor (VIF). A standard guideline is that VIF values higher than 10 indicate high multicollinearity (Kutner et al., 2004). However, a threshold equal to 5 is also commonly used (Sheather, 2009). Subsequently, the final models were selected as the ones with the independent variable configuration with the **lowest AIC and BIC values** for each developed model.

4.1.1 Belgium

GLMs were employed to investigate the relationship of key performance indicators (i.e. speeding, headway, overtaking and fatigue) for Belgian car drivers.

4.1.1.1 Speeding

The relationship between speeding and risk is widely recognized in the road safety community and as such, speeding is a commonly used dependent variable in transportation human factors research. The first Generalized Linear Regression model investigated the relationship between the **speeding and several explanatory variables of task complexity and coping capacity** (operator state). In particular, the dependent variable of the developed model is the dummy variable “speeding”, which is coded with 1 if there is a speeding event and with 0 if not. For task complexity, the variables used are time indicator, wipers and high beam, while for coping capacity - operator state, the variables used are distance traveled and harsh acceleration. It should be mentioned that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 4.

Table 4: Parameter estimates and multicollinearity diagnostics of the GLM for speeding

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	3.668	0.043	85.768	< .001	-
Time indicator	0.908	0.078	11.683	< .001	1.882
Weather	0.009	4.217×10^{-4}	20.952	< .001	1.228
High beam - Off	-0.018	7.062×10^{-4}	-25.286	< .001	1.470
Harsh acceleration	2.661	0.181	14.689	< .001	1.013
Distance	-6.128×10^{-4}	7.273×10^{-5}	-8.426	< .001	1.678
Summary statistics					
AIC	17404.428				
BIC	17413.817				
Degrees of freedom	88377				

Based on Table 4, it can be observed that all explanatory variables are statistically significant at a 95% confidence level; there is no issue of multicollinearity as the VIF values are much lower than 5. With regard to the coefficients, it was revealed that the indicators of task complexity, such as time indicator and wipers were positively correlated with speeding. The former refers to the time of the day (day coded as 1, dusk coded as 2, night coded as 3) which means that higher speeding events occur at night compared to during the day. This may be due to fewer cars on the road, lower visibility, and a false sense of security that comes with driving in the dark. Interestingly, wipers (wipers off coded as 0, wipers on coded as 1) were also found to have a positive correlation with speeding which means that there are more speeding events during adverse (e.g. rainy) weather conditions. This may be due to the fact that wet and slippery roads can make it more difficult to maintain control of the vehicle. Additionally, rain can reduce visibility and make it harder to see other cars or obstacles on the road. Taking into account the indicator of high beam (indicating lighting conditions; no high beam detected), a negative correlation was identified which means that when high beam was off - and, therefore, it was daytime - there were less speeding events. This finding comes in agreement with the previous argument with the indicator of time of the day that higher speeding events occur at night compared to the rest of the day.

Regarding the indicators of coping capacity - operator state, harsh accelerations had a positive relationship with the dependent variable (i.e. speeding), indicating that as the number of harsh acceleration increases, speeding also increases. This is a noteworthy finding of the current research as it confirms that harsh driving behavior events present a statistically significant positive correlation with speeding. Lastly, total distance traveled was negatively correlated with speeding which may be due to the fact that the longer a person drives, the more fatigued they may become, causing them to drive slower and more cautiously.

4.1.1.2 Headway

One of the major contributors to road crashes is the headway between two vehicles; when it is too short to allow the following driver to react appropriately to harsh braking by the leading vehicle. The headway between two vehicles can be expressed in terms of time and space. Within this framework, the second GLM investigated the relationship between the **headway and several explanatory variables of task complexity and coping capacity** (operator state). More specifically, the dependent variable of the developed model is the dummy variable "headway", which is coded with 1 if there is a headway event and with 0 if not. For task complexity, the variables used are time indicator, wipers and high beam, while for coping capacity - operator state, the variables used are exposure indicators of distance traveled and duration. It is worth noting that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 5.

Table 5: Parameter estimates and multicollinearity diagnostics of the GLM for headway

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	4.057	0.059	68.260	< .001	-
Duration	0.001	6.144×10^{-5}	17.806	< .001	1.005
Distance	0.001	8.553×10^{-5}	12.561	< .001	1.458
Weather	-0.002	5.417×10^{-4}	-3.463	< .001	1.650
High beam - Off	0.014	0.002	6.710	< .001	1.675
Time indicator	-1.059	0.035	-30.005	< .001	1.574
Summary statistics					
AIC	13569.585				
BIC	13579.111				
Degrees of freedom	101275				

Findings derived from Table 5 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 respect to the coefficients, it was found that time of the day (indicator of task complexity) was negatively correlated with headway, which means that drivers tend to keep safer distances from the vehicle in front of them during the night. This may probably be due to the fact that there is no heavy traffic during night hours; thus, headway events are avoided. Interestingly, high beam (indicating lighting conditions; no high beam detected) was positively correlated with headway which means that when high beam was off - and, therefore, it was daytime - there were more highway events. This finding comes in agreement with the previous argument with the indicator of time of the day that lower headway events occur at night compared to the rest of the day. In addition, wipers were also found to have a negative correlation with headway which means that there are less headway events during adverse (e.g. rainy) weather conditions. Furthermore, exposure indicators of distance and duration appeared to have a positive relationship with the dependent variable (i.e. headway).

4.1.1.3 Overtaking

The third GLM investigated the relationship between the **overtaking and several explanatory variables of task complexity and coping capacity** (operator state). For instance, the dependent variable of the developed model is the dummy variable "overtaking", which is coded with 1 if there is a overtaking event and with 0 if not. With regards to task complexity, the variables used are time indicator and wipers, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration, drowsiness. It should be noted that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 6.

Table 6: Parameter estimates and multicollinearity diagnostics of the GLM for overtaking

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-1.357	0.014	-94.380	< .001	-
Duration	4.017×10^{-4}	6.735×10^{-6}	59.641	< .001	1.010
Distance	8.217×10^{-4}	2.268×10^{-5}	36.233	< .001	1.509
Harsh acceleration	0.009	2.470×10^{-4}	36.319	< .001	1.565
Time indicator	-0.002	0.015	-0.121	0.904	1.684

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
Weather	0.001	9.121×10^{-5}	14.161	< .001	1.454
Drowsiness	1.201×10^{-5}	3.850×10^{-7}	31.193	< .001	1.047
Summary statistics					
AIC	123393.241				
BIC	123402.672				
Degrees of freedom	92129				

Taking into account the aforementioned Table 6, a series of interesting findings can be provided. First of all, the majority of the explanatory variables (expect for time indicator) were statistically significant at a 95% confidence level and there was no issue of multicollinearity as the VIF values were much lower than 5. It is worth noting that a similar pattern as the previous GLM for headway was identified. In particular, the indicator of time of the day was negatively correlated with overtaking, which means that drivers were not willing to perform an illegal overtaking during night, probably due to low traffic volumes occurred. On the other hand, wipers (indicating weather condition) was positively correlated with overtaking. With regards to the indicators of coping capacity – operator state, such as harsh accelerations, distance, duration and drowsiness appeared to have a positive relationship with the dependent variable (i.e. overtaking), indicating that as the values of the aforementioned independent variables increases, overtaking also increases.

4.1.1.4 Fatigue

The fourth GLM investigated the relationship between the **fatigue and several explanatory variables of task complexity and coping capacity** (operator state). In particular, the dependent variable of the developed model is the dummy variable “fatigue”, which is coded with 1 if there is a fatigue event and with 0 if not. For task complexity, the variables used are time indicator, wipers and high beam, while for coping capacity - operator state, the variables used are distance traveled, duration and harsh braking. It should be mentioned that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 7.

Table 7: Parameter estimates and multicollinearity diagnostics of the GLM for fatigue

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	0.046	0.014	3.350	< .001	-
Duration	1.942×10^{-5}	5.944×10^{-6}	3.267	0.001	1.026
Distance	-0.003	4.858×10^{-5}	-54.333	< .001	1.170
Time indicator	0.498	0.018	27.067	< .001	1.316
Weather	0.003	1.076×10^{-4}	29.736	< .001	1.278
High beam - Off	-0.015	3.067×10^{-4}	-49.304	< .001	1.367
Harsh braking	-1.103	0.029	-38.047	< .001	1.022
Summary statistics					
AIC	136914.741				
BIC	136924.247				
Degrees of freedom	99256				

All the explanatory variables were statistically significant at a 95% confidence level, as shown in Table 7. With regards to multicollinearity diagnostics, VIF values for all independent variables were much lower than 5. It was observed that the indicators of task complexity such as wipers and time indicator were positively correlated with fatigue. For instance, it was revealed that during the night, drivers are more prone to becoming fatigued due to the body's natural circadian rhythm. It can be more challenging for drivers to stay alert and focused when driving at night, especially during the early morning hours when the body is naturally in a state of rest. At the same time, high beam (indicating lighting conditions; high beam no detected) was negatively correlated with fatigue, which implies that when high beam was off - and, therefore, it was daytime - there were less fatigue events. Furthermore, indicators of coping capacity – operator state, such as duration had a positive relationship with the dependent variable (i.e. fatigue), indicating that the longer the duration is, the higher the probability of driver being fatigue becomes. This is a noteworthy finding of the current research as it confirms that exposure indicators present a statistically significant positive correlation with fatigue levels. Lastly, harsh braking and distance had a negative relationship with fatigue.

4.1.2 UK

GLMs were employed to investigate the relationship of key performance indicators (i.e. speeding and headway) for UK car drivers. It should be noted that results for overtaking and fatigue were not statistically significant; thus, they were not included.

4.1.2.1 Speeding

The first Generalized Linear Regression model investigated the relationship between the **speeding and several explanatory variables of task complexity and coping capacity** (vehicle and operator state). In particular, the dependent variable of the developed model is the dummy variable "speeding", which is coded with 1 if there is a speeding event and with 0 if not. For task complexity, the variables used are wipers on and high beam, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration events, gender, forward collision warning and right lane departure warning. It should be noted that for vehicle state, variables such as fuel type, vehicle age and gearbox were not statistically significant; and thus, these independent variables were not included in the analysis. The model parameter estimates are summarized in Table 8.

Table 8: Parameter estimates and multicollinearity diagnostics of the GLM for speeding

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-3.824	0.014	-274.620	< .001	-
Duration	4.672×10-5	7.877×10^{-7}	59.317	< .001	1.058
Harsh acceleration	-0.187	0.012	-15.377	< .001	1.014
Weather	-0.273	0.023	-11.713	< .001	1.008
High beam	0.128	0.078	1.635	0.102	1.002
Forward collision warning	10.603	2.479	4.276	< .001	1.001
Right lane departure warning	0.357	0.014	25.348	< .001	1.026
Distance	0.002	1.876×10^{-5}	117.628	< .001	1.072
Gender - Male	0.373	0.012	31.757	< .001	1.056
Summary statistics					
AIC	263599.548				
BIC	263610.743				
Degrees of freedom	537681				

Based on Table 8, it can be observed that all explanatory variables are statistically significant at a 95% confidence level and there is no issue of multicollinearity as the VIF values are much lower than 5. With regard to the coefficients, it was revealed that the indicators of coping capacity are all positively correlated with speeding except for harsh acceleration events that appear to be fewer when speeding occurs. The opposite happens with FCW and LDW events that appear to be higher in case of speeding. An increase in the trip duration and the distance travelled is associated with an increase in speeding events, as well. The use of wipers though is, as expected, negatively associated with speeding events. Gender was a significant variable in this model showing that male drivers (males coded as 0, females as 1), are possibly prone to speeding while the use of high beams also was connected with higher speeding events possibly due to lighter night hours traffic.

4.1.2.2 Headway

The second GLM investigated the relationship between the **headway and several explanatory variables of task complexity and coping capacity** (vehicle and operator state). More specifically, the dependent variable of the developed model is the dummy variable “headway”, which is coded with 1 if there is a headway event and with 0 if not. For task complexity, the variables used are weather and high beam, while for coping capacity - operator state, the variables used are distance traveled, duration, gender, right lane departure warning and harsh acceleration. The model parameter estimates are summarized in Table 9.

Table 9: Parameter estimates and multicollinearity diagnostics of the GLM for headway

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-2.675	0.009	-309.038	< .001	-
Duration	4.599×10-5	6.055×10^{-7}	75.950	< .001	1.037
Harsh acceleration	0.156	0.008	20.065	< .001	1.009
Weather	-0.133	0.014	-9.438	< .001	1.009
High beam	-1.575	0.085	-18.505	< .001	1.001
Right lane departure warning	0.106	0.010	10.737	< .001	1.019
Distance	0.003	1.263×10^{-5}	215.943	< .001	1.050
Gender - Male	0.052	0.008	6.733	< .001	1.040
Summary statistics					
AIC	549886.488				
BIC	549897.683				
Degrees of freedom	537681				

Findings derived from Table 9 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. For the model for headway variable, the FCW variable is not statistically significant, while harsh acceleration events are positively correlated with headway showing that higher number of harsh acceleration events are associated with shorter headways. High beam use and wipers use are negatively correlated with the headway showing that drivers in nighttime and during rainy weather conditions keep safer distances. It should be noted that both speeding, and headway variables are binary with 0 translating to no events and 1 to the occurrence of speeding or headway events accordingly.

4.1.3 Germany

GLMs were employed to investigate the relationship of key performance indicators (i.e. speeding, overtaking and fatigue) for German car drivers.

4.1.3.1 Speeding

The first Generalized Linear Regression model investigated the relationship between the **speeding and several explanatory variables of task complexity and coping capacity** (vehicle and operator state). In particular, the dependent variable of the developed model is the dummy variable “speeding”, which is coded with 1 if there is a speeding event and with 0 if not. For task complexity, the variables used are time indicator and high beam, for coping capacity - vehicle state, the variables used are type of fuel and vehicle age, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration, drowsiness, gender and age. The model parameter estimates are summarized in Table 10.

Table 10: Parameter estimates and multicollinearity diagnostics of the GLM for speeding

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	1.105	0.057	19.549	< .001	-
Duration	0.003	3.414×10^{-5}	73.366	< .001	1.262
Distance	5.735×10^{-4}	3.723×10^{-5}	15.404	< .001	1.029
Harsh acceleration	1.282×10^{-4}	1.974×10^{-6}	64.951	< .001	1.222
Fuel type - Petrol	0.219	0.010	21.446	< .001	1.328
Vehicle Age	3.162×10^{-5}	3.340×10^{-6}	9.469	< .001	1.277
Gender - Female	-0.275	0.021	-13.025	< .001	1.256
Age	-0.003	0.001	-2.289	0.022	1.076
Drowsiness	1.009×10^{-5}	2.656×10^{-6}	3.800	< .001	1.113
Time indicator	8.547×10^{-5}	1.925×10^{-6}	44.405	< .001	1.080
High beam - On	0.817	0.059	13.963	< .001	1.073
Summary statistics					
AIC	127971.813				
BIC	127981.881				
Degrees of freedom	174299				

Based on Table 10, it can be observed that all explanatory variables are statistically significant at a 95% confidence level; there is no issue of multicollinearity as the VIF values are much lower than 5. With regard to the coefficients, it was revealed that the indicators of task complexity, such as time and high beam (indicating lighting conditions; no high beam detected) were positively correlated with speeding. Regarding the indicators of coping capacity – vehicle state such as fuel type and vehicle age were positively correlated with speeding. Furthermore, it was demonstrated that indicators of coping capacity – operator state, such as harsh accelerations, distance, duration and drowsiness had a positive relationship with the dependent variable (i.e. speeding), indicating that as the values of the aforementioned independent variables increases, speeding also increases. This is a noteworthy finding of the current research as it confirms that harsh driving behavior events present a statistically significant positive correlation with speeding.

Taking into consideration socio-demographic characteristics, gender and age were negatively correlated with speeding. In particular, the negative value of the “Gender” coefficient implied that as the value of the variable was equal to 1 (males coded as 0, females as 1), the speeding percentage was lower. Results revealed that the vast majority of male drivers displayed less cautious behavior during their trips and exceeded more often the speed limits than female drivers. It is also remarkable that the negative value of the “Age” coefficient implied that as the value of the variable increased (higher value indicates increased age and, therefore, increased years of participant’s experience), the speeding percentage was lower. Young drivers appeared to have a riskier driving behavior than older drivers and were more prone to exceed the speed limits.

4.1.3.2 Overtaking

The second GLM investigated the relationship between the **overtaking and several explanatory variables of task complexity and coping capacity** (vehicle and operator state). For instance, the dependent variable of the developed model is the dummy variable “overtaking”, which is coded with 1 if there is a overtaking event and with 0 if not. With regards to task complexity, the variables used are time indicator and high beam, for coping capacity - vehicle state, the variables used are type of fuel and vehicle age, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration, drowsiness, gender and age. The model parameter estimates are summarized in Table 11.

Table 11: Parameter estimates and multicollinearity diagnostics of the GLM for overtaking

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-6.177	0.147	-41.985	< .001	-
Duration	1.082×10^{-4}	1.308×10^{-5}	8.274	< .001	1.384
Distance	-6.167×10^{-4}	7.495×10^{-5}	-8.227	< .001	1.200
Harsh acceleration	5.157×10^{-5}	6.526×10^{-6}	7.901	< .001	1.122
Fuel type - Petrol	0.218	0.028	7.869	< .001	1.599
Vehicle Age	6.051×10^{-5}	8.820×10^{-6}	6.860	< .001	1.320
Gender - Female	-0.437	0.049	-8.865	< .001	1.201
Age	-0.014	0.003	-5.416	< .001	1.394
Drowsiness	8.631×10^{-6}	4.970×10^{-6}	1.737	0.082	1.293
Time indicator	-1.125×10^{-4}	9.554×10^{-6}	-11.777	< .001	3.102
High beam - Off	7.737	0.088	87.972	< .001	3.291
Summary statistics					
AIC	61147.387				
BIC	61157.455				
Degrees of freedom	174299				

Taking into account the aforementioned Table 11, a series of interesting findings can be provided. First of all, all the explanatory variables were statistically significant at a 95% confidence level and there was no issue of multicollinearity as the VIF values were much lower than 5. It is worth noting that a similar pattern as the previous GLM for headway was identified. In particular, the indicator of time of the day was negatively correlated with overtaking, which means that drivers were not willing to perform an illegal overtaking during night, probably due to low traffic volumes occurred. On the other hand, high beam (indicating lighting conditions; no high beam detected) was positively correlated with overtaking.

With regards to the indicators of coping capacity – vehicle state, such as fuel type and vehicle age were positively correlated with overtaking, which means that drivers of older vehicle fleet were not willing to perform an illegal overtaking. Similarly, the indicators of coping capacity – operator state, such as harsh accelerations, duration and drowsiness appeared to have a positive relationship with the dependent variable (i.e. overtaking), indicating that as the values of the aforementioned independent variables increases, overtaking also increases. Interestingly, distance traveled was negatively correlated with overtaking. Lastly, gender and age had a negative relationship with the dependent variable (i.e. overtaking).

4.1.3.3 Fatigue

The third GLM investigated the relationship between the **fatigue and several explanatory variables of task complexity and coping capacity**. In particular, the dependent variable of the developed model is the dummy variable “fatigue”, which is coded with 1 if there is a fatigue event and with 0 if not. For task complexity, the variables used are time indicator and high beam, for coping capacity - vehicle state, the variables used are type of fuel and vehicle age, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration, gender and age. The model parameter estimates are summarized in Table 12.

Table 12: Parameter estimates and multicollinearity diagnostics of the GLM for fatigue

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-3.608	0.056	-64.056	< .001	-
Duration	$8.322{\times}10^{-4}$	$6.632{\times}10^{-6}$	125.488	< .001	1.172
Distance	0.001	$3.138{\times}10^{-5}$	36.108	< .001	1.124
Harsh acceleration	$-3.967{\times}10^{-5}$	$3.720{\times}10^{-6}$	-10.665	< .001	1.052
Fuel type - Diesel	-0.528	0.013	-40.328	< .001	1.421
Vehicle Age	$1.496{\times}10^{-4}$	$4.105{\times}10^{-6}$	36.437	< .001	1.794
Gender - Female	-0.930	0.029	-31.665	< .001	1.280
Age	0.012	0.001	8.306	< .001	1.139
Time indicator	$1.317{\times}10^{-4}$	$3.089{\times}10^{-6}$	42.645	< .001	1.075
High beam - Off	4.576	0.035	129.661	< .001	1.337
Summary statistics					
AIC	134848.401				
BIC	134858.470				
Degrees of freedom	174299				

All the explanatory variables were statistically significant at a 95% confidence level, as shown in Table 12. With regards to multicollinearity diagnostics, VIF values for all independent variables were much lower than 5. It was observed that the indicators of task complexity such as time and high beam (indicating lighting conditions; no high beam detected) were positively correlated with fatigue. In addition, the indicator of coping capacity – vehicle state such as vehicle age was positively correlated with fatigue. On the other hand, fuel type had a negative impact on the dependent variable “fatigue”. Furthermore, indicators of coping capacity – operator state, such as distance traveled and duration had a positive relationship with the dependent variable (i.e. fatigue), indicating that the longer the distance and duration is, the higher the probability of driver being fatigue becomes. This is a noteworthy finding of the current research as it confirms that exposure indicators present a statistically significant positive correlation with fatigue levels. Finally, harsh accelerations had a negative relationship with fatigue. Lastly, the negative value of the “gender” coefficient implied that female drivers were less fatigued as compared to male drivers.

4.1.4 Greece

GLMs were employed to investigate the relationship of key performance indicators (i.e. speeding and headway) for Greek car drivers. It should be noted that variables for headway, overtaking and fatigue were not available; thus, results for the aforementioned indicators were not included.

4.1.4.1 Speeding

The GLM applied investigated the relationship between the **speeding and several explanatory variables of coping capacity** (vehicle and operator state). In particular, the dependent variable of the developed model is the dummy variable “speeding”, which is coded with 1 if there is a speeding event and with 0 if not. For task complexity, the variables used are time indicator, for coping capacity - vehicle state, the variables used are type of fuel, gearbox and vehicle age, while for coping capacity - operator state, the variables used are distance traveled, duration, harsh acceleration, harsh braking, gender and age. The model parameter estimates are summarized in Table 13.

Table 13: Parameter estimates and multicollinearity diagnostics of the GLM for speeding

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	66.123	16.472	4.014	< .001	-
Duration	6.326×10^{-4}	2.547×10^{-5}	24.839	< .001	1.110
Distance	0.002	8.698×10^{-5}	21.915	< .001	1.146
Harsh acceleration	-0.433	0.051	-8.508	< .001	1.346
Harsh braking	0.113	0.067	1.696	0.090	1.447
Age	-0.044	0.002	-27.215	< .001	1.331
Gender1	0.397	0.059	6.698	< .001	1.732
Fuel_type - Petrol	0.297	0.046	6.389	< .001	1.368
VehicleAge	0.032	0.008	3.919	< .001	1.456
Gearbox - Automatic	-0.518	0.056	9.251	< .001	1.353
Time indicator	0.201	0.021	9.544	< .001	1.057
Summary statistics					
AIC	19378.588				
BIC	19386.426				
Degrees of freedom	18736				

Based on Table 13, it can be observed that all explanatory variables are statistically significant at a 95% confidence level; there is no issue of multicollinearity as the VIF values are much lower than 5. With regard to the coefficients, it was revealed that the indicators of task complexity, such as time indicator was positively correlated with speeding. Time indicator refers to the time of the day (day coded as 1, dusk coded as 2, night coded as 3) which means that higher speeding events occur at night compared to during the day. This may be due to fewer cars on the road, lower visibility, and a false sense of security that comes with driving in the dark.

With regard to the coefficients, it was revealed that the indicators of coping capacity – vehicle state, such as fuel type and gearbox were negatively correlated with speeding and vehicle age was positively correlated with speeding. More specifically, the positive value of the variable “Fuel type” coefficient implied that when the fuel type was petrol (diesel coded as 1, hybrid electric coded as 2 and petrol coded as 3), the speeding percentage became higher. This indicated that vehicles with gasoline-powered engines provided higher speeding events compared to other types of vehicles, such as electric cars and hybrid cars. Additionally, the positive value of the “Vehicle Age” coefficient revealed that the higher the value of this variable, the higher the speeding percentage. This means that the increased proportion of older vehicles increases the risk to exceed the speed limits. This finding was also confirmed by Torok (2020) who found that by reducing the number of older vehicles on the roads, especially vehicles older than 15 years, road safety can be improved. This was probably due to the fact that in the current years, with the permanent development and safety improvements of the automotive sector, more and more vehicles are equipped with advanced driver assistance systems which include the ability of the vehicle to stop, the stability control of the vehicle, the passive safety systems (e.g. frontal and side airbags) or the ability of the vehicle to perceive its environment (e.g. frontal and backward sensors) in order to comply with the speed limits.

Furthermore, it was demonstrated that indicators of coping capacity – operator state, such as harsh braking, distance and duration had a positive relationship with the dependent variable (i.e. speeding), indicating that as the values of the aforementioned independent variables increases, speeding also increases. This is a noteworthy finding of the current research as it confirms that harsh driving behavior events present a statistically significant positive correlation with speeding.

Taking into consideration socio-demographic characteristics, gender and age were negatively correlated with speeding. In particular, the positive value of the "Gender" coefficient implied that as the value of the variable was equal to 0 (males coded as 0, females as 1), the speeding percentage was higher. Results revealed that the vast majority of male drivers displayed less cautious behavior during their trips and exceeded more often the speed limits than female drivers. It is also remarkable that the negative value of the "Age" coefficient implied that as the value of the variable increased (higher value indicates increased age and, therefore, increased years of participant's experience), the speeding percentage was lower. Young drivers appeared to have a riskier driving behavior than older drivers and were more prone to exceed the speed limits.

4.1.5 Portugal

GLMs were employed to investigate the relationship of key performance indicators (i.e. speeding, headway, overtaking and fatigue) for Portuguese bus drivers.

4.1.5.1 Speeding

The first GLM investigated the relationship between the **speeding and several explanatory variables of task complexity and coping capacity** (operator state). In particular, the dependent variable of the developed model is the dummy variable "speeding", which is coded with 1 if there is a speeding event and with 0 if not. For task complexity, the variable used is time indicator while for coping capacity - operator state, the variables used are distance traveled, harsh acceleration, harsh braking and fatigue. It should be mentioned that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 14.

Table 14: Parameter estimates and multicollinearity diagnostics of the GLM for speeding

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	3.441	0.020	168.858	< .001	-
Time indicator	0.164	0.008	21.306	< .001	1.002
Harsh braking	0.294	0.082	3.594	< .001	1.051
Harsh acceleration	0.490	0.112	4.371	< .001	1.052
Fatigue	-0.095	0.008	-12.527	< .001	1.378
Distance	0.010	1.038×10-4	99.797	< .001	1.379
Summary statistics					
AIC	153657.374				
BIC	153668.223				
Degrees of freedom	380656				

Based on Table 14, it can be observed that all explanatory variables are statistically significant at a 95% confidence level; there is no issue of multicollinearity as the VIF values are much lower than 5. With regard to the coefficients, it was revealed that the indicators of task complexity, such as time indicator was positively correlated with speeding. Time indicator refers to the time of the day (day coded as 1, dusk coded as 2, night coded as 3) which means that higher speeding events occur at night compared to during the day. This may be due to fewer cars on the road, lower visibility, and a false sense of security that comes with driving in the dark. Regarding the indicators of coping capacity - operator state, distance and harsh events (i.e. harsh acceleration and harsh braking) had a positive relationship with the dependent variable (i.e. speeding), indicating that as the total distance traveled and the number of harsh events increases, speeding also increases. Lastly, fatigue was negatively correlated with speeding which implies that the more fatigued the driver is, the slower and more cautiously they drive.

4.1.5.2 Headway

The second GLM investigated the relationship between the **headway and several explanatory variables of task complexity and coping capacity** (operator state). More specifically, the dependent variable of the developed model is the dummy variable “headway”, which is coded with 1 if there is a headway event and with 0 if not. For task complexity, the variable used is time indicator, while for coping capacity - operator state, the variables used are exposure indicators of distance traveled, harsh acceleration, harsh braking. It is worth noting that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models. The model parameter estimates are summarized in Table 15.

Table 15: Parameter estimates and multicollinearity diagnostics of the GLM for headway

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-5.526	0.055	-100.579	< .001	-
Time indicator	-0.192	0.022	-8.781	< .001	1.001
Harsh braking	0.897	0.242	3.708	< .001	1.045
Harsh acceleration	0.147	0.318	0.462	0.644	1.045
Distance	0.009	2.425×10^{-4}	35.162	< .001	1.000
Summary statistics					
AIC	27567.794				
BIC	27567.794				
Degrees of freedom	380657				

Findings derived from Table 15 demonstrated that the majority of 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 respect to the coefficients, it was found that time of the day (indicator of task complexity) was negatively correlated with headway, which means that drivers tend to keep safer distances from the vehicle in front of them during the night. This may probably be due to the fact that there is no heavy traffic during night hours; thus, headway events are avoided. Furthermore, exposure indicator of distance as well as harsh events (i.e. harsh acceleration and harsh braking) appeared to have a positive relationship with the dependent variable (i.e. headway).

4.1.5.3 Overtaking

The third GLM investigated the relationship between the **overtaking and several explanatory variables of task complexity and coping capacity** (operator state). For instance, the dependent variable of the developed model is the dummy variable “overtaking”, which is coded with 1 if there is a overtaking event and with 0 if not. With regards to task complexity, the variable used is time indicator, while for coping capacity - operator state, the variables used are distance traveled, harsh acceleration, harsh braking and average speed. It should be noted that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models.The model parameter estimates are summarized in Table 16.

Table 16: Parameter estimates and multicollinearity diagnostics of the GLM for overtaking

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	-7.928	0.171	-46.241	< .001	-
Time indicator	-0.120	0.065	-1.855	0.064	1.004
Average speed	1.229	0.074	16.522	< .001	1.018
Distance	0.007	8.365×10^{-4}	7.835	< .001	1.016
Harsh braking	-0.316	0.690	-0.459	0.646	1.044
Harsh acceleration	0.568	0.996	0.570	0.568	1.045
Summary statistics					
AIC	4195.226				
BIC	4206.076				
Degrees of freedom	380656				

Taking into account the aforementioned Table 16, a series of interesting findings can be provided. First of all, the majority of the explanatory variables (expect for harsh events) were statistically significant at a 95% confidence level and there was no issue of multicollinearity as the VIF values were much lower than 5. It is worth noting that a similar pattern as the previous GLM for headway was identified. In particular, the indicator of time of the day was negatively correlated with overtaking, which means that drivers were not willing to perform an illegal overtaking during night, probably due to low traffic volumes occurred. With regards to the indicators of coping capacity – operator state, such as harsh accelerations, distance and average speed appeared to have a positive relationship with the dependent variable (i.e. overtaking), indicating that as the values of the aforementioned independent variables increases, overtaking also increases. For instance, this means that the longer the distance of the trip is, the higher the number of the overtaking events occur. In addition, increased number of total harsh acceleration can be an indicator of overtaking which requires drivers to accelerate quickly to pass another vehicle. On the other hand, harsh braking had a negative correlation with overtaking which means that drivers tend to avoid overtaking when they perform harsh braking. Harsh braking can be a sign of aggressive driving, and drivers who exhibit this behavior may be less likely to take risks or make sudden maneuvers, such as overtaking.

4.1.5.4 Fatigue

The fourth GLM investigated the relationship between the **fatigue and several explanatory variables of task complexity and coping capacity** (operator state). In particular, the dependent variable of the developed model is the dummy variable “fatigue”, which is coded with 1 if there is a fatigue event and with 0 if not. For task complexity, the variable used is time indicator, while for coping capacity - operator state, the variables used are distance traveled, harsh acceleration, harsh braking and average speed. It should be mentioned that the explanatory variables of vehicle state, such as fuel type, vehicle age or gearbox, or socio-demographic characteristics, such as gender, age or educational level are not statistically significant at a 95% confidence level; thus, these variables are not included in the models.The model parameter estimates are summarized in Table 17.

Table 17: Parameter estimates and multicollinearity diagnostics of the GLM for fatigue

Variables	Estimate	Standard Error	z-value	Pr(|z|)	VIF
(Intercept)	0.556	0.010	58.498	< .001	-
Time indicator	0.101	0.004	27.277	< .001	1.001
Average speed	-0.045	0.014	-3.180	0.001	1.075
Distance	0.009	7.428×10^{-5}	123.989	< .001	1.074
Harsh braking	0.224	0.039	5.758	< .001	1.050
Harsh acceleration	0.334	0.057	5.862	< .001	1.051
Summary statistics					
AIC	455426.929				
BIC	455437.779				
Degrees of freedom	380656				

All the explanatory variables were statistically significant at a 95% confidence level, as shown in Table 17. With regards to multicollinearity diagnostics, VIF values for all independent variables were much lower than 5. It was observed that time indicator was positively correlated with fatigue. This may be due to the fact that during the night, drivers are more prone to becoming fatigued due to the body's natural circadian rhythm. It can be more challenging for drivers to stay alert and focused when driving at night, especially during the early morning hours when the body is naturally in a state of rest. Moreover, indicators of coping capacity – operator state, such as distance and harsh events had a positive relationship with the dependent variable (i.e. fatigue), indicating that the longer the distance is, the higher the probability of driver being fatigue becomes. This is a noteworthy finding of the current research as it confirms that exposure indicators present a statistically significant positive correlation with fatigue levels. Lastly, average speed had a negative relationship with fatigue, which implies that the higher the average speed is, the lower the fatigue events are. This finding may be due to the fact that driving at a higher average speed makes drivers be alert and can help reduce fatigue.

4.2 Structural Equation Models

Following exploratory analysis, the latent variable (or variables) associated to the latent variable "**task complexity**" and "**coping capacity**" were estimated from the various indicators. This way, the effect of different personal factors on 'operator state' was defined, and further analyzed for different countries (i.e. Belgium, UK, Germany, Greece, Portugal) and different travel modes (i.e. cars, trucks, buses). Several SEM were applied in order to identify the impact of task complexity and coping capacity on the STZ level, controlling for the above exogenous factors.

4.2.1 Belgium (Cars)

4.2.1.1 Speeding

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three stages of the STZ) of speeding. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 39 drivers, 1,173 trips (23,725 minutes)
- Phase 2: real-time interventions - 43 Belgian car drivers, 1,549 trips (31,414 minutes)
- Phase 3: real-time & post-trip interventions - 51 Belgian car drivers, 1,973 trips (40,121 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 49 Belgian car drivers, 2,468 trips (52,077 minutes)

The results for phase 1 are shown in Figure 6 below. It is shown that several operator state indicators load on the latent variable coping capacity, as follows:

- Driver's age, with a negative correlation indicating that older drivers have lower coping capacity.
- Better general driving skills are associated with higher coping capacity.
- Higher exposure on rural roads per week is associated with lower coping capacity, possibly because those drivers have lower exposure in complex (urban) environments and cannot sustain sufficient skills to cope with them.
- A sportive and ambitious driving style is associated with higher coping capacity, possibly indicating a younger age and a higher alertness of these drivers. It is worth mentioning that a sporty and ambitious driver is someone who enjoys driving and wants to push themselves and their vehicle to the limits. They may enjoy taking their car on winding roads, racing, or participating in other high-performance driving activities. They may also be interested in upgrading their vehicle with performance modifications to enhance its capabilities. Overall, a sporty and ambitious driver is someone who is passionate about driving and wants to get the most out of their car.
- Driver's confidence to their own driving skills is associated with higher coping capacity.
- Drivers reporting of always driving higher than the speed limit is associated with higher coping capacity.

These results are in line with the dedicated exploratory analysis of Deliverable 6.2.

At the same time, in line with Deliverable 6.1 on Task Complexity investigation, there are two indicators loading on the latent variable:

- 'wipers on' (indicating rainy weather conditions)
- 'high-beam on' (indicating night-time or poor visibility conditions)

The latent variable Risk is measured by means of the STZ levels for speeding (level 1 'normal driving' used as the reference case), with positive correlations of Risk with the STZ indicators.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation – albeit the magnitude of this correlation is very small. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers' coping capacity increases as the complexity of driving task increases. This finding may be a sign of risk compensating behavior of drivers when the complexity of driving task is high, and is in line with the theoretical model of i-DREAMS, validating the assumption that risk is an outcome of the interaction between the two variables in addition to their separate effects. The more complex the situation becomes as a result of speeding, the better the driver's coping capacity will become, for example because of increased alertness.

Task Complexity increase is associated with lower risk, which is not intuitive. Although the initial assumption was that task complexity would increase risk, once its effect is moderated by that of coping capacity the opposite is the case. It is noted however that the task complexity latent variable is measured by environmental indicators (i.e. rainy weather, night-time) which are known to induce compensatory behaviors by drivers, in particular expressed as reduced speed during these more demanding conditions. Variables on road type, traffic conditions etc. would need to be included for a complete picture of the role of task complexity on the risk expressed in terms of speeding STZ.

At the same time, coping capacity is associated with higher risk, again an interesting finding. It could be assumed that higher coping capacity might reduce risk; however, the coping capacity indicators in our sample include static demographic and self-reported behavior indicators and therefore are more representative of driver personality and general driving styles, and less so of the real-time operator state during the experiment. For instance, indicators related to the level of sleepiness, fatigue or distraction were either not available or not significant in this model. Therefore, it can be concluded that younger, more confident and less compliant drivers exhibited lower risk in this experiment, in terms of exceeding the STZ speeding boundaries – a finding which can be attributed to higher alertness and exposure in complex environments, without however taking into account the variations of their state during these trips.

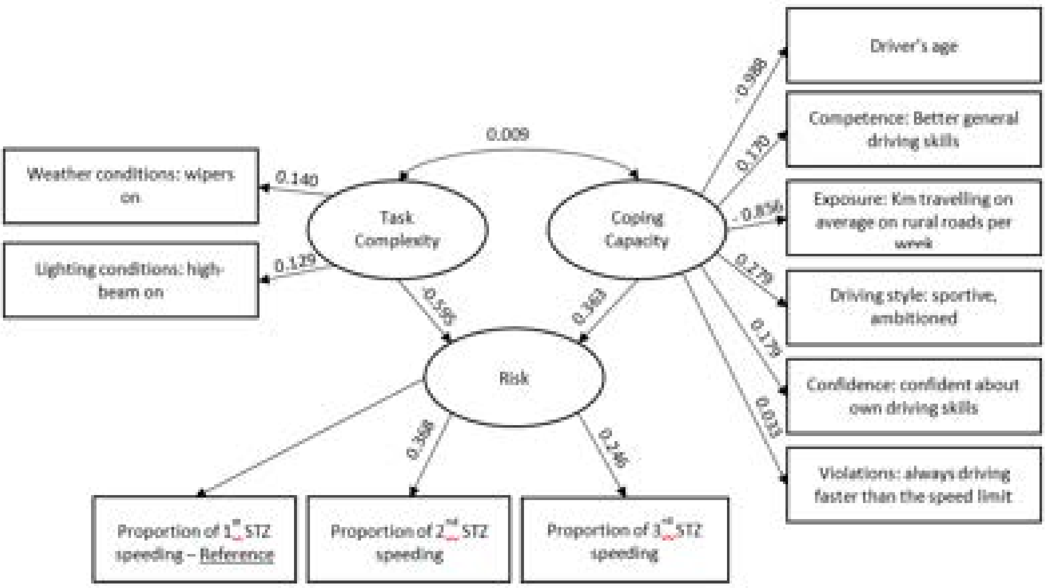

Figure 6: Results of SEM on Risk (speeding STZ) – Belgian car drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.661; TLI is 0.560 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.121. Table 18 summarizes the model fit of SEM applied for speeding.

Table 18: Model Fit Summary for speeding – Belgian car drivers – experiment Phase 1

Model Fit measures	Value
AIC	273200.6
BIC	273402.4
CFI	0.661
TLI	0.560
RMSEA	0.121

Residual variances details are presented in Table 19 that follows.

Table 19: Residual variances for speeding – Belgian car drivers – experiment Phase 1

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Wiper	0.042	0.000	96.426	0.000
.Night	0.185	0.007	25.853	0.000
.Age	0.024	0.010	2.415	0.016
.COMPT1	0.971	0.009	108.795	0.000
.Rural	0.268	0.008	33.666	0.000
.Style	0.218	0.002	108.336	0.000
.CONF	0.215	0.002	108.775	0.000
.VIO2	0.999	0.009	108.912	0.000
.iSP2	0.010	0.000	31.670	0.000
.iSP3	0.047	0.000	97.501	0.000
TC	-0.001	0.000	-5.668	0.000
CC	0.976	0.014	71.608	0.000
.RISK	0.003	0.000	9.341	0.000

Figure 7, Figure 8 andFigure 9 show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the relationships among risk, task complexity and coping capacity are fairly consistent between the different phases, it is interesting to note however:

- The impact of exposure on rural roads disappears during the interventions phases, possibly indicating that the implementation of the i-DREAMS interventions helped drivers to counterbalance this effect.
- On phase 2, the indicators of numbers of Forward Collision Warnings (FCW) and Pedestrian Collision Warnings (PCW) are loading on task complexity, introducing the impact of real-time events recorded expressing demanding and risky situations as indicators of task complexity. It is noted that the overall impact of task complexity on Risk is only slightly reduced. Given that these FCW and PCW events may or may not be directly associated with exceeding the speed limit, as is the representation of risk in this case. Furthermore, these indicators were not found to be significant in the 3rd and 4th phase of the experiment, but the number of such events was also lower during these phases of the experiment.
- On phase 4, the structural relationship between task complexity and coping capacity changes to a negative coefficient, and the relationship between task complexity and risk changes to a positive coefficient. This finding may not be directly interpreted, but it is possible that the presence of all i-DREAMS interventions on phase 4 lead to a different interaction between the three latent variables, which would need additional indicators available in order to draw conclusions.
- The loadings of the observed proportions of the STZ of speeding are consistent between the different phases, it is noted though that the loading of the 3rd STZ level

becomes notably higher in the 4th phase of the experiment. This may indicate that the increased risk in these conditions is determined by those drivers who do not respond to the interventions and reach the 3rd level – their proportion however is smaller in the 3rd and 4th phase of the experiment.

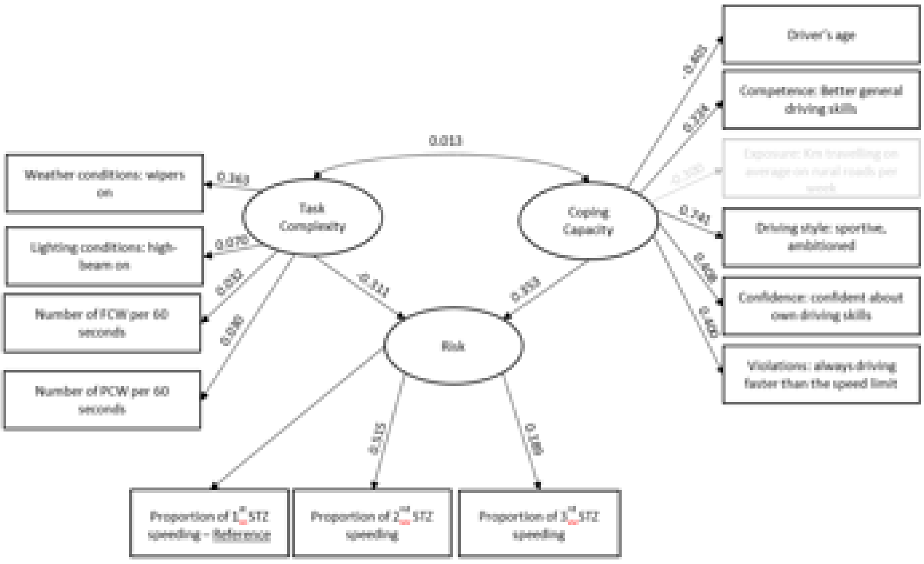

Figure 7: Results of SEM on Risk (speeding STZ)– Belgian car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.473; TLI is 0.335 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.082. Table 20 summarizes the model fit of SEM applied for speeding.

Table 20: Model Fit Summary for speeding – Belgian car drivers – experiment Phase 2

Model Fit measures	Value
AIC	57294.26
BIC	57518.77
CFI	0.473
TLI	0.335
RMSEA	0.082

Residual variances details are presented in Table 21 that follows.

Table 21: Residual variances for speeding – Belgian car drivers – experiment Phase 2

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Wiper	0.050	0.002	25.373	0.000
.HBeam	0.006	0.000	120.714	0.000
.FCW	0.013	0.000	122.607	0.000
.PCW	0.003	0.000	121.884	0.000
.Age	0.839	0.008	107.963	0.000
.COMPT1	0.950	0.008	119.233	0.000
.Style	0.110	0.003	40.771	0.000
.CONF	0.157	0.001	107.216	0.000
.VIO2	0.840	0.008	108.022	0.000

Variable	Estimate	Std.Err	z-value	P(>|z|)
.iSP2	0.006	0.000	17.758	0.000
.iSP3	0.049	0.000	100.601	0.000
TC	0.007	0.002	3.741	0.000
CC	0.161	0.006	27.813	0.000
.RISK	0.002	0.000	5.543	0.000

The results for phase 3 are shown in Figure 8 below.

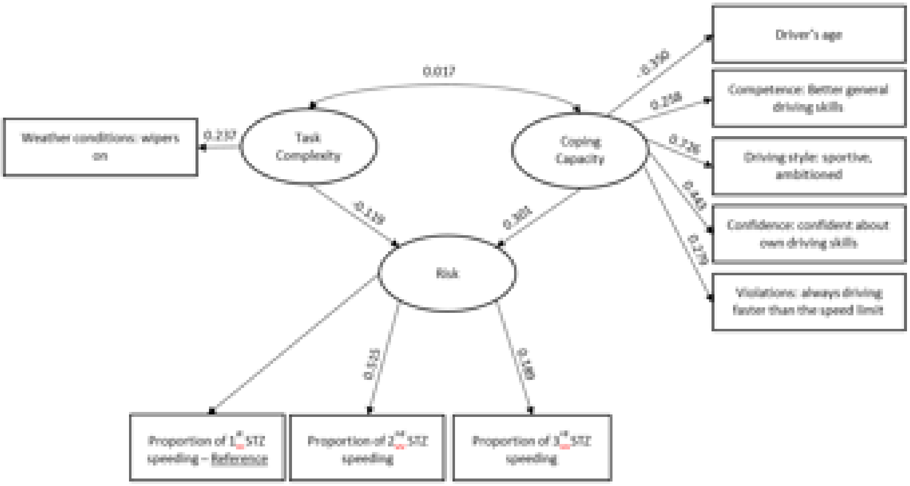

Figure 8: Results of SEM on Risk (speeding STZ)– Belgian car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.484; TLI is 0.291 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.103. Table 22 summarizes the model fit of SEM applied for speeding.

Table 22: Model Fit Summary for speeding – Belgian car drivers – experiment Phase 3

Model Fit measures	Value
AIC	338636.6
BIC	338808.6
CFI	0.484
TLI	0.291
RMSEA	0.103

Residual variances details are presented in Table 23 that follows.

Table 23: Residual variances for speeding – Belgian car drivers – experiment Phase 3

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Age	0.877	0.007	128.226	0.000
.COMPT1	0.934	0.007	135.295	0.000
.Style	0.116	0.003	46.070	0.000
.CONF	0.148	0.001	115.393	0.000
.VIO2	0.922	0.007	133.988	0.000

Variable	Estimate	Std.Err	z-value	P(>|z|)
.iSP2	0.005	0.000	58.011	0.000
.iSP3	0.053	0.001	93.492	0.000
TC	0.056	0.000	141.635	0.000
CC	0.123	0.004	27.475	0.000
.RISK	0.001	0.000	9.576	0.000

The results for phase 4 are shown in Figure 9 below.

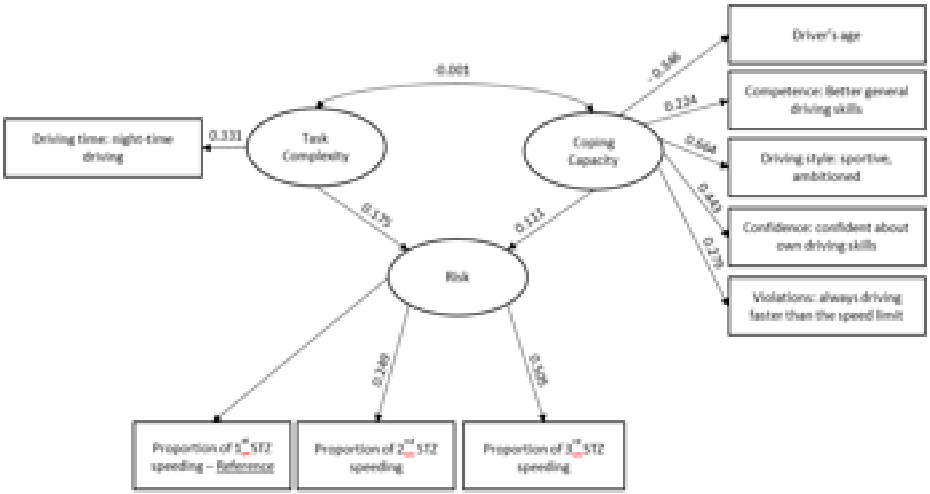

Figure 9: Results of SEM on Risk (speeding STZ)– Belgian car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.817; TLI is 0.709 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.037. Table 24 summarizes the model fit of SEM applied for speeding.

Table 24: Model Fit Summary for speeding – Belgian car drivers – experiment Phase 4

Model Fit measures	Value
AIC	271111.2
BIC	271253.0
CFI	0.817
TLI	0.709
RMSEA	0.037

Residual variances details are presented in Table 25 that follows.

Table 25: Residual variances for speeding – Belgian car drivers – experiment Phase 4

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Age	0.880	0.009	102.258	0.000
.COMPT1	0.950	0.007	145.629	0.000
.Style	0.139	0.006	22.614	0.000
.iSP2	0.006	0.000	110.846	0.000
.iSP3	0.036	0.001	28.890	0.000
TC	0.109	0.001	161.364	0.000

Variable	Estimate	Std.Err	z-value	P(>|z|)
CC	0.120	0.007	16.474	0.000
.RISK	0.000	0.000	9.695	0.000

4.2.1.2 Headway

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three phases of the STZ) based on headway measurement. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 38 Belgian car drivers, 633 trips (16,393 minutes)
- Phase 2: real-time interventions - 42 Belgian car drivers, 813 trips (21,412 minutes)
- Phase 3: real-time & post-trip interventions - 50 Belgian car drivers, 990 trips (27,691 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 49 Belgian car drivers, 1,222 trips (35,284 minutes)

The results for phase 1 are shown in Figure 10 below. It is shown that the latent variable coping capacity is measured by means of the operator state indicators that were significant in the speeding-based SEM Risk model (see previous section), with the addition of the IBI (Inter-Beat-Interval), which was also observed in the dedicated exploratory investigation of coping capacity alone (see Deliverable 6.2). Task complexity is measured by the same indicators as in the previous model, and in line with the exploratory findings of Deliverable 6.1 (Papazikou et al., 2023).

Risk is measured by means of the STZ levels for headway (level 1 ‘normal driving’ used as the reference case), with positive correlation of Risk with the 2nd and 3rd level of the STZ headway indicators – which are here grouped together due to lack of sufficient data for the 3rd level.

The structural model between task complexity, coping capacity and risk shows great consistency with that of the previous section.

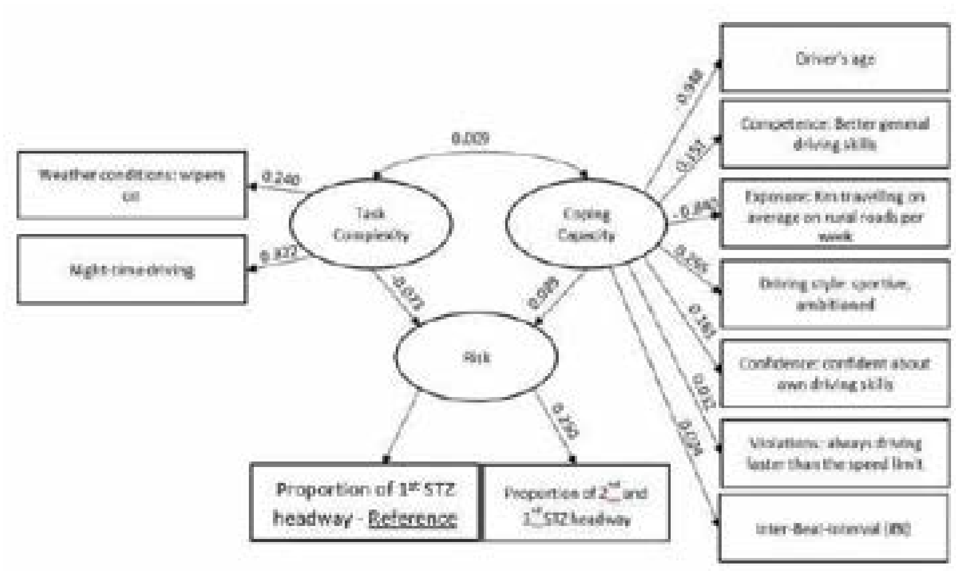

Figure 10: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.526; TLI is 0.395 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.142. Table 26 summarizes the model fit of SEM applied for headway.

Table 26: Model Fit Summary for headway – Belgian car drivers – experiment Phase 1

Model Fit measures	Value
AIC	248443
BIC	248626
CFI	0.526
TLI	0.395
RMSEA	0.142

Residual variances details are presented in Table 27 that follows.

Table 27: Residual variances for headway – Belgian car drivers – experiment Phase 1

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Wiper	0.047	0.001	55.717	0.000
.Night	0.147	0.004	34.992	0.000
.Age	0.089	0.013	6.979	0.000
.COMPT1	0.959	0.011	86.881	0.000
.Rural	0.237	0.009	25.156	0.000
.Style	0.231	0.003	86.344	0.000
.CONF	0.224	0.003	86.851	0.000
.VIO2	0.957	0.011	87.064	0.000
.IBI	0.999	0.011	87.068	0.000
TC	0.003	0.001	4.240	0.000
CC	0.795	0.016	49.074	0.000
.RISK	0.061	0.001	84.318	0.000

Figure 11, 12 and 13 show the respective results of the 2nd, 3rd and 4th phase of the experiment. Overall, there are fluctuations between both the structural and the measurement equations of the model in the different phases. For instance, IBI is a significant indicator of coping capacity only in phases 1 & 3, and the signs of the regression coefficients between the latent variables change in different phases. These findings may be due to the differences in the samples, as well as the higher sensitivity of headway measurements as STZ determinants.

It may be interesting to emphasise on the model of phase 2. In that phase, the introduction of the real-time interventions reveals a significant indicator loading on task complexity, which is the number of PCW recorded per minute. The negative sign of this loading indicates that higher number of PCW per minute is associated with lower task complexity, which may imply that warning the drivers about the presence of pedestrians removes a burden from the drivers shoulders and decreases the complexity of driving task for them. At the same time, the correlation of task complexity with risk becomes non-significant, and the correlation between task complexity and coping capacity becomes negative. This suggests that higher task complexity, measured by night-time driving and PCWs, results in lower coping capacity, which in turn results in higher risk of exceeding the headway thresholds of safe driving. Although this sounds intuitive, especially at the beginning of the implementation of interventions, the different patterns shown by the model in different phases does not allow to conclude on the nature of the relationships.

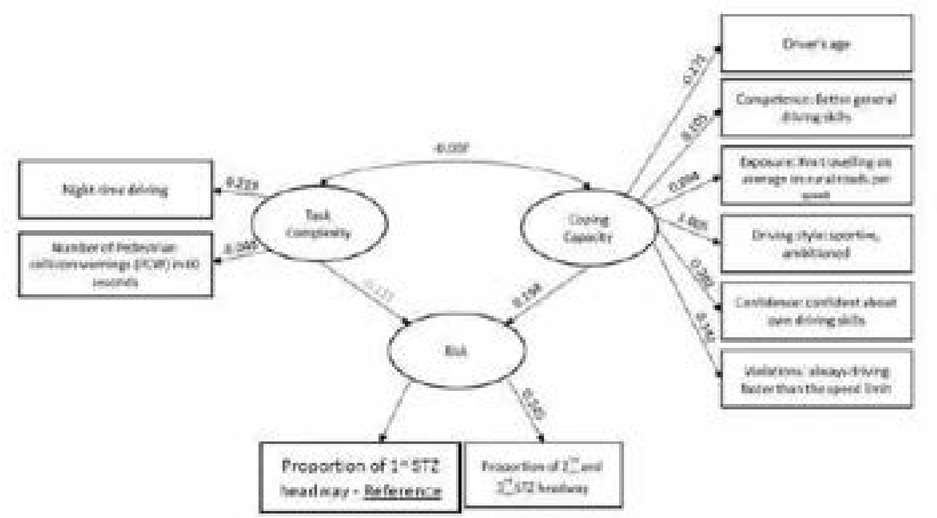

Figure 11: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.361; TLI is 0.158 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.127. Table 28 summarizes the model fit of SEM applied for headway.

Table 28: Model Fit Summary for headway – Belgian car drivers – experiment Phase 2

Model Fit measures	Value
AIC	241196.0
BIC	241371.4
CFI	0.361
TLI	0.158
RMSEA	0.127

Residual variances details are presented in Table 29 that follows.

Table 29: Residual variances for headway – Belgian car drivers – experiment Phase 2

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Night	0.090	0.003	26.356	0.000
.PCW	0.003	0.000	101.950	0.000
.Age	0.927	0.009	100.706	0.000
.COMPT1	0.989	0.010	103.423	0.000
.Rural	0.991	0.010	103.441	0.000
.Style	-0.002	0.007	-0.341	0.733
.CONF	0.178	0.002	100.200	0.000
.VIO2	0.883	0.009	96.207	0.000
TC	0.005	0.003	1.372	0.170
CC	0.073	0.004	17.462	0.000
.RISK	0.056	0.001	66.161	0.000

The results for phase 3 are shown in Figure 12 below.

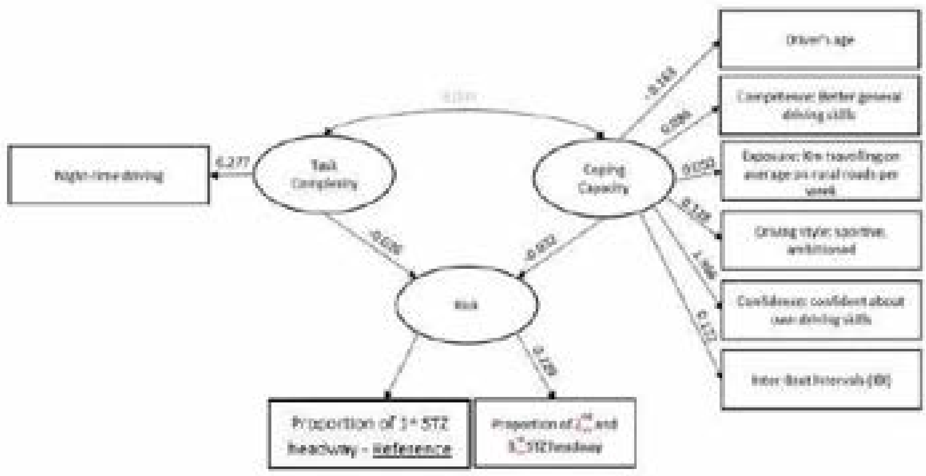

Figure 12: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.446; TLI is 0.261 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.109. Table 30 summarizes the model fit of SEM applied for headway.

Table 30: Model Fit Summary for headway – Belgian car drivers – experiment Phase 3

Model Fit measures	Value
AIC	242126.3
BIC	242275.4
CFI	0.446
TLI	0.261
RMSEA	0.109

Residual variances details are presented in Table 31 that follows.

Table 31: Residual variances for headway – Belgian car drivers – experiment Phase 3

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Age	0.800	0.009	89.701	0.000
.COMPT1	1.207	0.012	97.045	0.000
.Rural	0.333	0.003	97.411	0.000
.Style	0.246	0.003	94.588	0.000
.CONF	-0.578	0.128	-4.516	0.000
.IBI	0.985	0.010	95.180	0.000
TC	0.077	0.001	97.316	0.000
CC	0.022	0.004	5.537	0.000
.RISK	0.051	0.001	97.380	0.000

The results for phase 4 are shown in Figure 13 below.

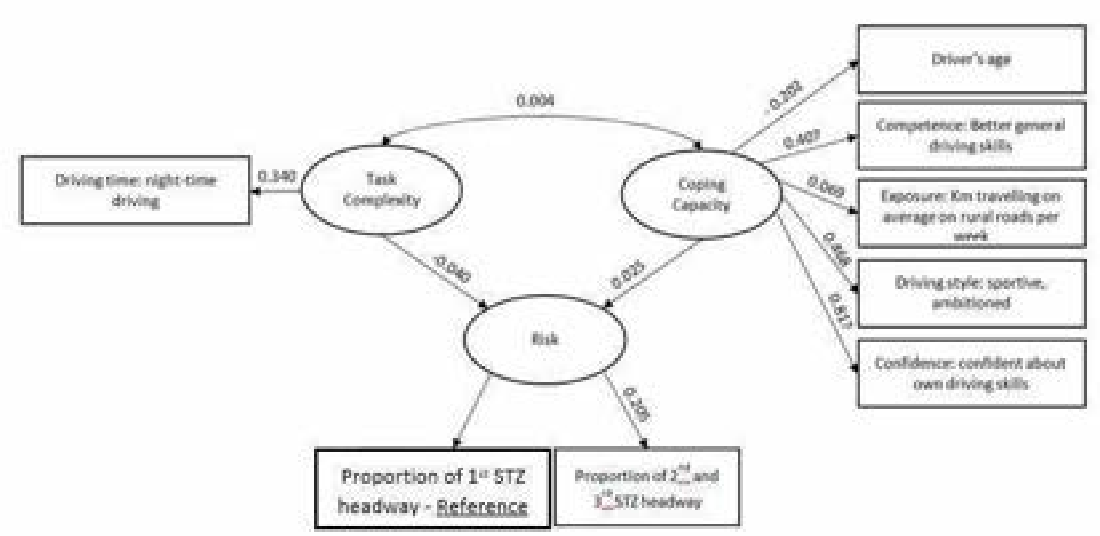

Figure 13: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.454; TLI is 0.236 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.125. Table 32 summarizes the model fit of SEM applied for headway.

Table 32: Model Fit Summary for headway – Belgian car drivers – experiment Phase 4

Model Fit measures	Value
AIC	396860.4
BIC	397004.4
CFI	0.454
TLI	0.236
RMSEA	0.125

Residual variances details are presented in Table 33 that follows.

Table 33: Residual variances for headway – Belgian car drivers – experiment Phase 4

Variable	Estimate	Std.Err	z-value	P(>|z|)
.Age	0.959	0.007	130.202	0.000
.COMPT1	0.834	0.007	114.274	0.000
.Rural	0.995	0.008	132.565	0.000
.Style	0.195	0.002	102.767	0.000
.CONF	0.069	0.003	22.845	0.000
TC	0.116	0.001	132.823	0.000
CC	0.041	0.003	15.860	0.000
.RISK	0.042	0.000	132.791	0.000

4.2.2 Belgium (Trucks)

4.2.2.1 Speeding

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three stages of the STZ) of speeding. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 23 Belgian truck drivers, 1,148 trips (117,160 minutes)
- Phase 2: real-time interventions - 22 Belgian truck drivers, 1,691 trips (146,315 minutes)
- Phase 3: real-time & post-trip interventions - 22 Belgian truck drivers, 1,440 trips (139,245 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 23 Belgian truck drivers, 1,767 trips (187,636 minutes)

The results for phase 1 are shown in Figure 14 below. It is shown that one trip related and several operator state indicators load on the latent variable coping capacity, as follows:

- Trip duration, with a negative correlation indicating that higher trip duration is associated with lower coping capacity.
- Driver's age, with a negative correlation indicating that older drivers have lower coping capacity.
- A sportive and ambitious driving style is associated with higher coping capacity, possibly indicating a younger age and a higher alertness of these drivers.
- Driver's confidence to their own driving skills is associated with higher coping capacity.

These results are in line with the dedicated exploratory analysis of Deliverable 6.2 (Michelaraki et al., 2023).

At the same time, in line with Deliverable 6.1 (Papazikou et al., 2023) on task complexity investigation, there are two indicators loading on the latent variable:

- 'wipers on' (indicating rainy weather conditions)
- 'speed (indicating situational needs)

The latent variable risk is measured by means of the STZ levels for speeding (level 1 'normal driving' used as the reference case) with negative correlations of risk with the STZ indicators. The negative sign shows that the latent variable risk could in fact be representing an inverse of risk, more like a normal driving.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation –which reduces in magnitude as the driver's progress from phase 1 though phase 4. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. This finding may be a sign of risk compensating behavior of drivers when the complexity of driving task is high, and is in line with the theoretical model of i-DREAMS, validating the assumption that Risk (or its' inverse, the normal driving) is an outcome of the interaction between the two variables in addition to their separate effects.

Task complexity is negatively associated with (inverse) risk (normal driving), which is intuitive. The higher the complexity, the lower the chances to drive normally. For instances, in rainy conditions, it would make it hard for the drivers to maintain normal driving behavior and given the situation, they may feel compelled to speed and thus enters into dangerous driving phase (STZ 2).

At the same time, coping capacity is negatively associated with (inverse) risk (or what we established as normal driving), again an interesting finding (similar to the case of headway and acceleration). It could be assumed that higher coping capacity might encourage normal driving and reduce risk but it is not the case here. Furthermore, the coping capacity indicators in our sample include static demographic and self-reported behavior indicators and therefore are more representative of driver personality and general driving styles, and less so of the real-time operator state during the experiment. For instance, indicators related to the level of sleepiness, fatigue or distraction were either not available or not significant in this model. Therefore, it can be concluded that younger, and more confident drivers exhibited (lower normal driving) higher Risk in this experiment, in terms of exceeding the STZ speeding boundaries, without however taking into account the variations of their state during these trips.

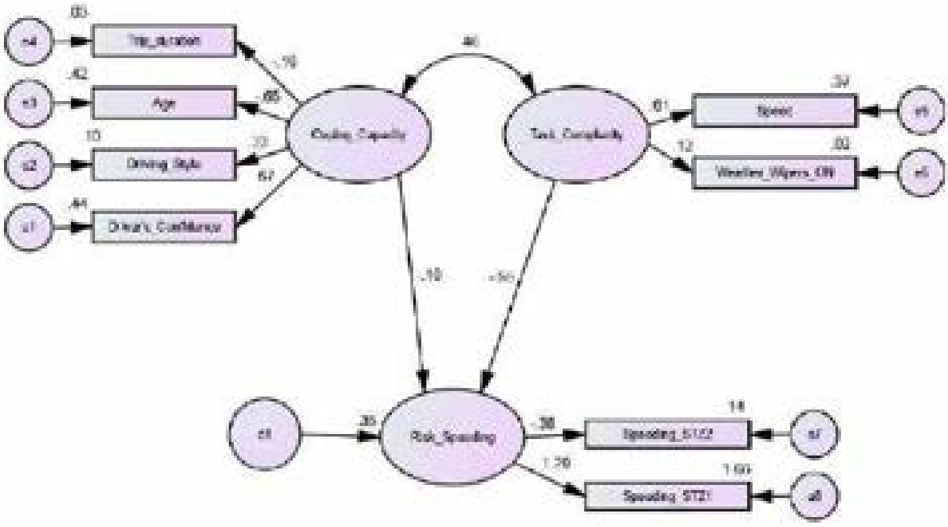

Figure 14: Results of SEM on Risk (speeding STZ) – Belgian truck drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.899; TLI is 0.834 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.080. Table 34 summarizes the model fit of SEM applied for speeding.

Table 34: Model Fit Summary for speeding – Belgian truck drivers – experiment Phase 1

Model Fit measures	Value
AIC	12877.885
BCC	12877.889
CFI	0.899
TLI	0.834
RMSEA	0.080
Hoelter's critical N (α = .05)	253
Hoelter's critical N (α = .01)	306

Residual variances details are presented in Table 35 that follows.

Table 35: Residual variances for speeding – Belgian truck drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.207	0.002	89.232	***
Task_Complexity	44.89	1.643	27.324	***
Risk_Speeding	0.108	0.003	34.537	***
Confidence	0.259	0.002	129.926	***
Style	0.154	0.001	227.281	***
Age	66.377	0.474	140.046	***
Trip_duration	4203.012	17.611	238.661	***
Speed	75.671	1.627	46.515	***
ME_Car_wipers_median	0.010	0.000	239.874	***
iDreams_Speeding_Map_level_2_mean	0.010	0.000	209.914	***
Speeding	-0.067	0.002	-27.418	***

Figure 15, 16 and 17 show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the relationships among risk, task complexity and coping capacity are fairly consistent between the different phases, it is interesting to note however:

- The structural relationship between task complexity and coping capacity remains positive across all phases, although it reduces in magnitude in phase 4. Similarly, the relationship between task complexity and risk remains the same though the magnitude increases in the negative direction. Moreover, the relationship between coping capacity and risk is also consistent across phases.
- The loadings of the observed proportions of the STZ of speeding are consistent between the different phases, it is noted though that the loading of the 2nd STZ level becomes notably higher in the 1st phase of the experiment compared to other phases. This could be attributed to i-DREAMS interventions as they were active in phase 2, 3 and 4.
- The loading of trip duration was negative in 1st phase but it changes to positive in the following phases of the experiment. This could be that with the presence of interventions, the coping capacity of the drivers increase and they can maintain normal driving for longer trips.

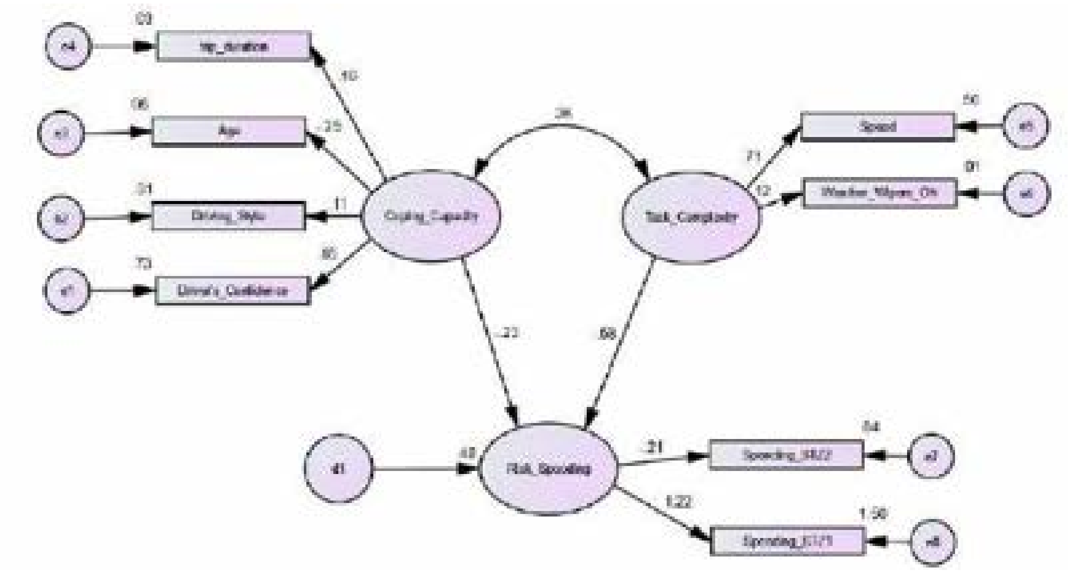

Figure 15: Results of SEM on Risk (speeding STZ) – Belgian truck drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.895; TLI is 0.827 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.074. Table 36 summarizes the model fit of SEM applied for speeding.

Table 36: Model Fit Summary for speeding – Belgian truck drivers – experiment Phase 2

Model Fit measures	Value
AIC	13650.075
BCC	13650.079
CFI	0.895
TLI	0.827
RMSEA	0.074
Hoelter's critical N (α = .05)	297
Hoelter's critical N (α = .01)	360

Residual variances details are presented in Table 37 that follows.

Table 37: Residual variances for speeding – Belgian truck drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.357	0.005	70.373	***
Task_Complexity	70.586	1.930	36.571	***
Risk_Speeding	0.093	0.004	26.119	***
Style	0.183	0.001	269.346	***
Age	98.414	0.377	261.063	***
Trip_duration	6945.027	25.953	267.604	***
Speed_mps	69.783	1.895	36.834	***
ME_Car_wipers_median	0.009	0.000	268.989	***
iDreams_Speeding_Map_level_2_mean	0.005	0.000	264.728	***
Confidence	0.132	0.005	27.523	***
Speeding_1	-0.059	0.003	-18.844	***

The results for phase 3 are shown in Figure 16 below.

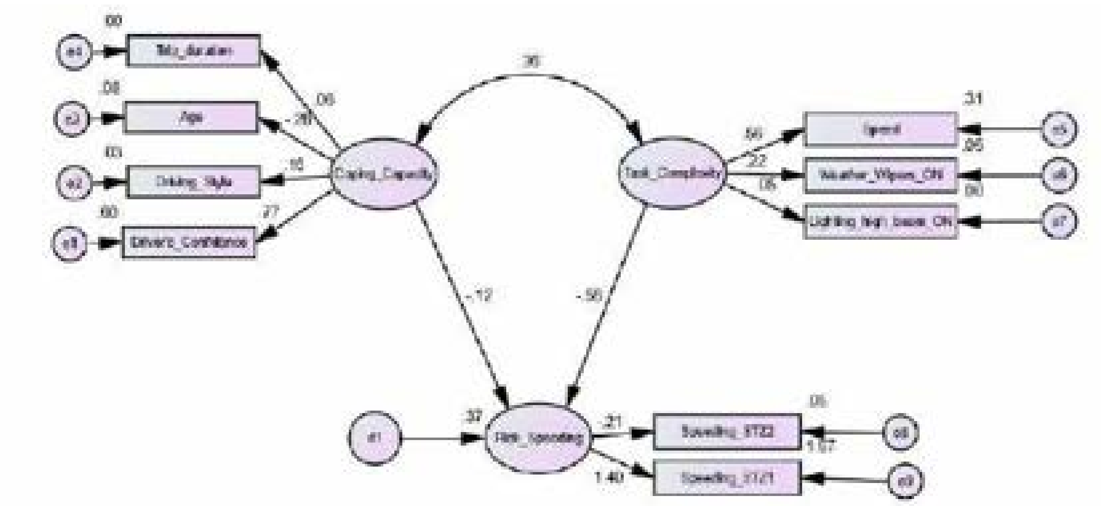

Figure 16: Results of SEM on Risk (speeding STZ) – Belgian truck drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.865; TLI is 0.747 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.062. Table 38 summarizes the model fit of SEM applied for speeding.

Table 38: Model Fit Summary for speeding – Belgian truck drivers – experiment Phase 3

Model Fit measures	Value
AIC	12903.621
BCC	12903.625
CFI	0.865
TLI	0.747
RMSEA	0.062
Hoelter's critical N (α = .05)	395
Hoelter's critical N (α = .01)	466

Residual variances details are presented in Table 39 that follows.

Table 39: Residual variances for speeding – Belgian truck drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.300	0.005	55.424	***
Task_Complexity	39.942	0.832	48.011	***
Risk_Speeding	0.107	0.005	22.822	***
Confidence	0.200	0.005	38.633	***
Style	0.183	0.001	259.391	***
Age	87.083	0.356	244.848	***
Trip_duration	10527.72	39.975	263.357	***
Speed	87.360	0.824	106.051	***
ME_Car_wipers_median	0.007	0.000	255.733	***
ME_Car_high_beam_median	0.000	0.000	263.68	***
iDreams_Speeding_Map_level_2_mean	0.004	0.000	250.436	***
Speeding_1	-0.084	0.005	-18.531	***

The results for phase 4 are shown in Figure 17 below.

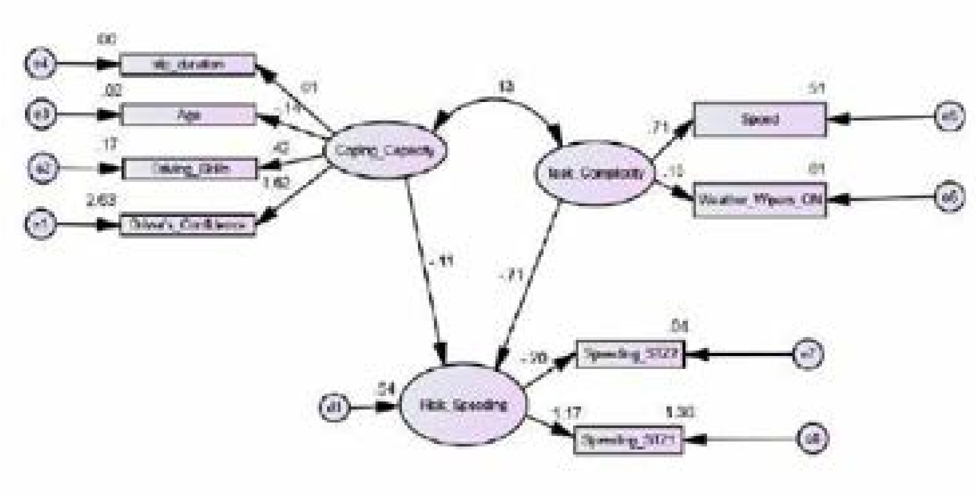

Figure 17: Results of SEM on Risk (speeding STZ) – Belgian truck drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.938; TLI is 0.898 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.075. Table 40 summarizes the model fit of SEM applied for speeding.

Table 40: Model Fit Summary for speeding – Belgian truck drivers – experiment Phase 4

Model Fit measures	Value
AIC	17806.525
BCC	17806.528
CFI	0.938
TLI	0.898
RMSEA	0.075
Hoelter's critical N (α = .05)	292
Hoelter's critical N (α = .01)	354

Residual variances details are presented in Table 41 that follows.

Table 41: Residual variances for speeding – Belgian truck drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	1.453	0.022	65.738	***
Task_Complexity	74.018	2.101	35.224	***
Risk_Speeding	0.081	0.004	23.062	***
Confidence	-0.901	0.022	-40.197	***
Skills	0.337	0.002	218.606	***
Age	121.971	0.396	308.141	***
Trip_duration	10166.125	33.187	306.33	***
Speed	72.445	2.073	34.94	***
ME_Car_wipers_median	0.005	0.000	305.372	***
iDreams_Speeding_Map_level_2_mean	-0.046	0.002	-19.116	***
Speeding_1	0.005	0.000	301.293	***

4.2.2.2 Harsh Acceleration

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three stages of the STZ) of harsh acceleration. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 23 Belgian truck drivers, 1,334 trips (28,296 minutes)
- Phase 2: real-time interventions - 22 Belgian truck drivers, 1,543 trips (34,297 minutes)
- Phase 3: real-time & post-trip interventions - 22 Belgian truck drivers, 1,346 trips (31,827 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 23 Belgian truck drivers, 1,602 trips (42,289 minutes)

The results for phase 1 are shown in Figure 18 below. It is shown that one trip related variable and several operator state indicators load on the latent variable coping capacity, as follows:

- Trip duration, with a negative correlation indicating that higher trip duration is associated with lower coping capacity.
- Driver’s age, with a negative correlation indicating that older drivers have lower coping capacity.
- A sportive and ambitious driving style is associated with higher coping capacity, possibly indicating a younger age and a higher alertness of these drivers.
- Driver’s confidence to their own driving skills is associated with higher coping capacity.

These results are mostly in line with the dedicated exploratory analysis of Deliverable 6.2 (Michelaraki et al., 2023). There is however a small exception. The ‘trip duration’ enters into the list of predictors of latent variable coping capacity while ‘driving skills’ remains insignificant.

At the same time, in line with Deliverable 6.1 on Task Complexity investigation, there are two indicators loading on the latent variable:

- ‘wipers on’ (indicating rainy weather conditions)
- ‘speed’ (indicating the situational constraints)

The latent variable Risk is measured by means of the STZ levels for acceleration (level 1 ‘normal driving’ used as the reference case), with negative correlations of Risk with the STZ indicators. The negative sign shows that the latent variable risk could in fact be representing an inverse of risk, more like a normal driving.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. This finding may be a sign of risk compensating behavior of drivers when the complexity of driving task is high, and is in line with the theoretical model of iDreams, validating the assumption that Risk (or conversely the normal driving as established above) is an outcome of the interaction between the two variables in addition to their separate effects.

Task complexity increase is associated with higher ~~(Risk)~~ normal driving (lower risk), which is not intuitive. Although the initial assumption was that task complexity would increase risk or decrease normal driving, once its effect is moderated by that of coping capacity the opposite is the case. It is noted however that the task complexity latent variable is measured by environmental indicator (i.e. rainy weather) and situational indicator (i.e. speed) which are known to induce compensatory behaviors by drivers, in particular expressed as reduced speed during the more demanding conditions. Variables on road type, traffic conditions etc. would need to be included for a complete picture of the role of task complexity on the risk (normal driving) expressed in terms of acceleration STZ.

At the same time, coping capacity is negatively associated with normal driving or inverse of risk, again an interesting finding. It could be assumed that higher coping capacity might reduce risk or improve normal driving but this is not the case here. Furthermore, the coping capacity indicators in our sample include static demographic and self-reported behavior indicators and therefore are more representative of driver personality and general driving styles, and less so of the real-time operator state during the experiment. For instance, indicators related to the level of sleepiness, fatigue or distraction were either not available or not significant in this model. Therefore, it can be concluded that younger, more confident truck drivers exhibited (higher risk) lower normal driving in this experiment, in terms of exceeding the STZ acceleration boundaries, without however taking into account the variations of their state during these trips.

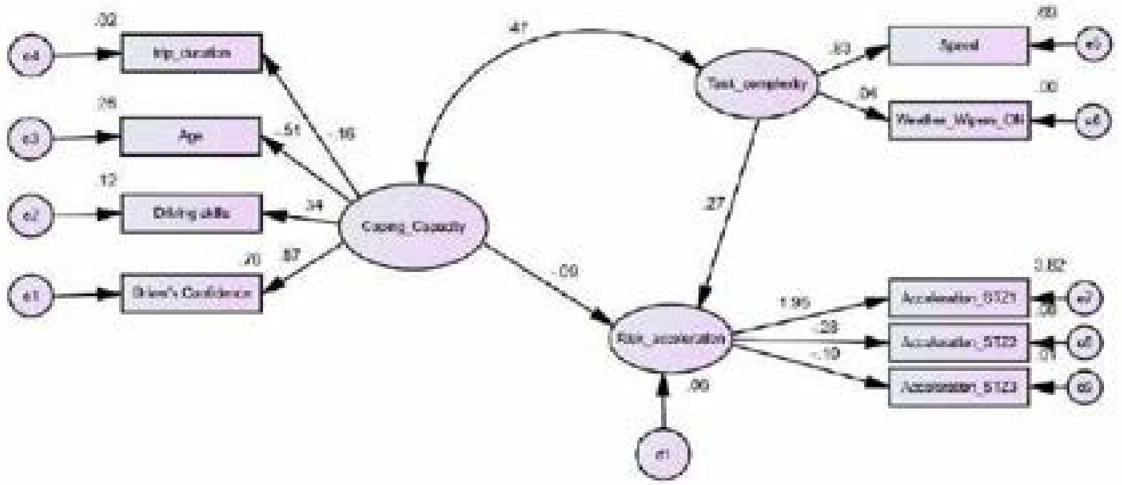

Figure 18: Results of SEM on Risk (Harsh acceleration STZ) – Belgian truck drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.921; TLI is 0.881 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.062. Table 42 summarizes the model fit of SEM applied for harsh acceleration.

Table 42: Model Fit Summary for harsh acceleration – Belgian truck drivers – experiment Phase 1

Model Fit measures	Value
AIC	2730.212
BCC	2730.234
CFI	0.921
TLI	0.881
RMSEA	0.062
Hoelter's critical N (α = .05)	386
Hoelter's critical N (α = .01)	456

Residual variances details are presented in Table 43 that follows.

Table 43: Residual variances for harsh acceleration – Belgian truck drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
Task_complexity	65.97	15.26	4.324	***
Coping_Capacity	0.326	0.007	49.42	***
Risk_acceleration	0.638	0.056	11.38	***
Speed_mps	29.82	15.24	1.957	0.05
ME_Car_wipers_median	0.006	0.000	118.9	***
Skills	0.339	0.003	113.1	***
Age	85.89	0.897	95.72	***
Trip_duration	3807	32.24	118.1	***
DEM_evt_ha_lvl_L_mean	-0.500	0.055	-9.17	***
DEM_evt_ha_lvl_M_mean	0.084	0.001	91.55	***
DEM_evt_ha_lvl_H_mean	0.017	0.000	119	***
Confidence	0.102	0.006	18.06	***

Figure 19, 20 and 21 show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the relationships among risk, task complexity and coping capacity are fairly consistent between the different phases (except phase 3 where coping capacity and risk have positive relationship), it is interesting to note however:

- In phase 3, the structural relationship between coping capacity and (inverse) risk changes to a positive coefficient. This finding may not be directly interpreted, but it is possible that the presence of real time and post trip i-DREAMS interventions in phase 3 lead to a different interaction between the latent variables coping capacity and risk, which would need additional indicators available in order to draw conclusions. Also, the magnitude of the correlation between latent variables coping capacity and task complexity reduces to extremely small value.
- The loading of 'trip duration' in phase 2 changes to positive sign which show a momentarily improvement in the coping capacity of drivers in the presence of real-time interventions. However, in the later phases 3 and 4, this trend is back as the phase 1.
- The loadings of the observed proportions of the STZ of acceleration are consistent between the different phases (The loadings of 2nd STZ level have consistently higher negative sign across all phases while the loadings of 3rd STZ level have consistently lower sign across all phases). The loading of 1st STZ level becomes notably higher in the 4th phase of the experiment. This may indicate that drivers tend to have normal driving in 4th phase in the presence of all interventions.

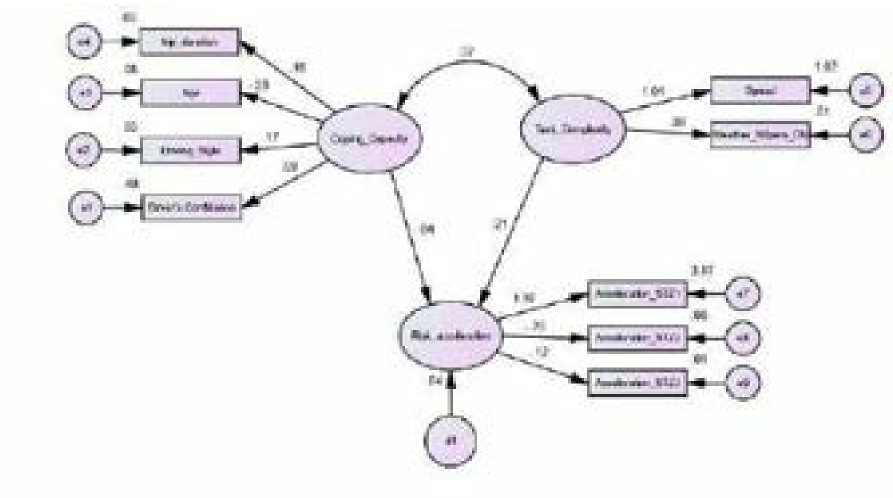

Figure 19: Results of SEM on Risk (Harsh acceleration STZ) – Belgian truck drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.813; TLI is 0.719 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.088. Table 44 summarizes the model fit of SEM applied for harsh acceleration.

Table 44: Model Fit Summary for harsh acceleration – Belgian truck drivers – experiment Phase 2

Model Fit measures	Value
AIC	6417.821
BCC	6417.839
CFI	0.813
TLI	0.719
RMSEA	0.088
Hoelter's critical N (α = .05)	197
Hoelter's critical N (α = .01)	232

Residual variances details are presented in Table 45 that follows.

Table 45: Residual variances for harsh acceleration – Belgian truck drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.207	0.008	26.008	***
Task_Complexity	109.198	12.789	8.539	***
Risk_acceleration	0.632	0.046	13.713	***
Style	0.199	0.002	127.278	***
Age	93.906	0.798	117.657	***
Trip_Duration	7056.981	55.518	127.112	***
speed_mps	-2.588	12.763	-0.203	0,839
ME_Car_wipers_median	0.010	0.000	130.306	***
DEM_evt_ha_lvl_L_mean	-0.487	0.046	-10.608	***
DEM_evt_ha_lvl_M_mean	0.071	0.001	104.147	***
DEM_evt_ha_lvl_H_mean	0.019	0.000	130.831	***
Confidence	0.227	0.008	29.59	***

The results for phase 3 are shown in Figure 20 below.

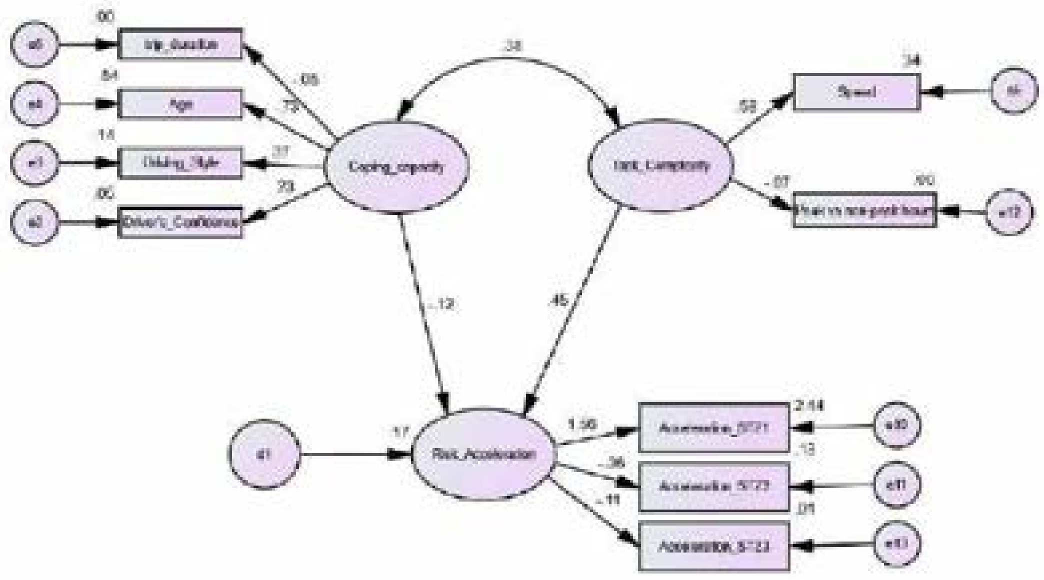

Figure 20: Results of SEM on Risk (Harsh acceleration STZ) – Belgian truck drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.882; TLI is 0.778 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.064. Table 46 summarizes the model fit of SEM applied for harsh acceleration.

Table 46: Model Fit Summary for harsh acceleration – Belgian truck drivers – experiment Phase 3

Model Fit measures	Value
AIC	3177.783
BCC	3177.802
CFI	0.882
TLI	0.778
RMSEA	0.064
Hoelter's critical N (α = .05)	372
Hoelter's critical N (α = .01)	439

Residual variances details are presented in Table 47 that follows.

Table 47: Residual variances for harsh acceleration – Belgian truck drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_capacity	0.024	0.002	14.063	***
Task_Complexity	31.965	4.574	6.989	***
Risk_Acceleration	0.338	0.024	14.296	***
Confidence	0.444	0.004	118.919	***
Style	0.162	0.002	92.562	***
Age	41.443	2.147	19.307	***
Trip_duration	9191.06	72.982	125.937	***
Speed_mps	62.137	4.566	13.609	***
DEM_evt_ha_lvl_L_mean	-0.241	0.020	-12.017	***
DEM_evt_ha_lvl_M_mean	0.069	0.001	93.597	***
time_of_day_p_np	0.198	0.002	125.743	***

Variable	Estimate	Std. Error	z-value	P(>|z|)
DEM_evt_ha_lvl_H_mean	0.011	0.000	126.573	***

The results for phase 4 are shown in Figure 21 below.

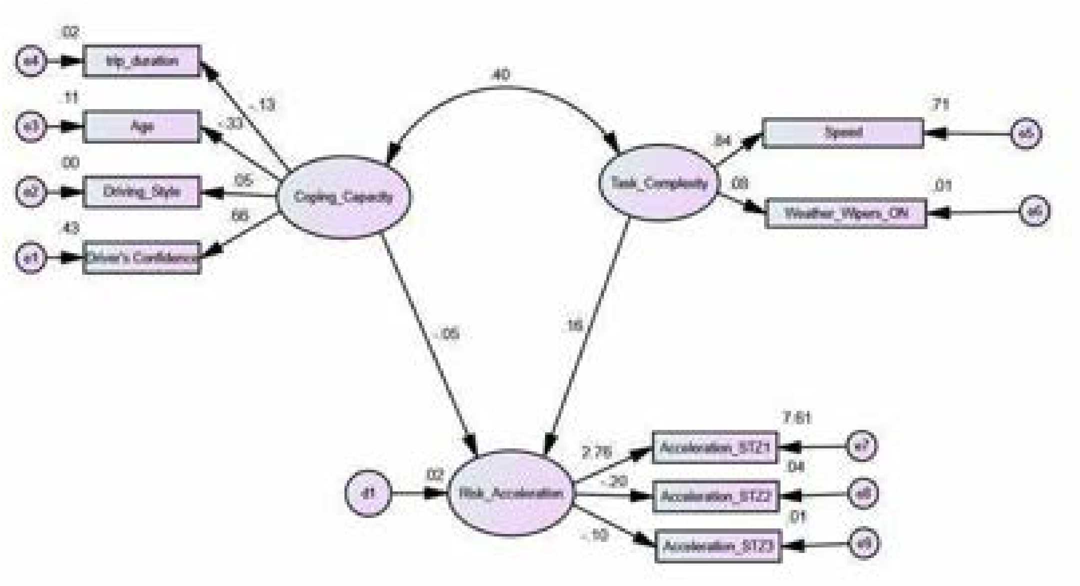

Figure 21: Results of SEM on Risk (Harsh acceleration STZ) – Belgian truck drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.843; TLI is 0.764 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.077. Table 48 summarizes the model fit of SEM applied for harsh acceleration.

Table 48: Model Fit Summary for harsh acceleration – Belgian truck drivers – experiment Phase 4

Model Fit measures	Value
AIC	6089.699
BCC	6089.713
CFI	0.843
TLI	0.764
RMSEA	0.077
Hoelter's critical N (α = .05)	256
Hoelter's critical N (α = .01)	302

Residual variances details are presented in Table 49 that follows.

Table 49: Residual variances for harsh acceleration – Belgian truck drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.226	0.009	24.724	***
Task_Complexity	81.73	8.328	9.814	***
Risk_Acceleration	1.304	0.184	7.08	***
Confidence	0.297	0.009	33.426	***
Style	0.227	0.002	145.037	***
Age	116.936	0.985	118.683	***
Trip_duration	10232.541	71.682	142.75	***

Variable	Estimate	Std. Error	z-value	P(>|z|)
Speed	33.659	8.296	4.057	***
ME_Car_wipers_median	0.005	0.000	144.637	***
DEM_evt_ha_lvl_L_mean	-1.157	0.184	-6.281	***
DEM_evt_ha_lvl_M_mean	0.083	0.001	110.975	***
DEM_evt_ha_lvl_H_mean	0.03	0.000	143.032	***

4.2.2.3 Headway

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three stages of the STZ) of headway. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 23 truck drivers, 1,148 trips (117,160 minutes)
- Phase 2: real-time interventions - 22 Belgian truck drivers, 1,691 trips (146,315 minutes)
- Phase 3: real-time & post-trip interventions - 22 Belgian truck drivers, 1,440 trips (139,245 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 23 Belgian truck drivers, 1,767 trips (187,636 minutes)

The results for phase 1 are shown in Figure 22 below. It is shown that one trip related variable and several operator state indicators load on the latent variable coping capacity, as follows:

- Trip duration, with a negative correlation indicating that higher trip duration is associated with lower coping capacity.
- Driver's age, with a negative correlation indicating that older drivers have lower coping capacity.
- A sportive and ambitious driving style is associated with higher coping capacity, possibly indicating a younger age and a higher alertness of these drivers.
- Driver's confidence to their own driving skills is associated with higher coping capacity.

These results are in line with the dedicated exploratory analysis of Deliverable 6.2.

At the same time, in line with Deliverable 6.1 on task complexity investigation, there are two indicators loading on the latent variable:

- 'wipers on' (indicating rainy weather conditions)
- 'speed' (indicating situational needs)

The latent variable risk is measured by means of the STZ levels for headway (level 1 'normal driving' used as the reference case), with negative correlations of risk with the STZ indicators. The negative sign shows that the latent variable risk could in fact be representing an inverse of risk, more like a normal driving.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. This finding may be a sign of risk compensating behavior of drivers when the complexity of driving task is high, and is in line with the theoretical model of i-DREAMS, validating the assumption that Risk (or conversely the normal driving as established above) is an outcome of the interaction between the two variables in addition to their separate effects.

Task complexity is negatively associated with the latent variable (inverse) risk, which was defined by different levels of headway. This was expected finding as task complexity would decrease normal driving. In rainy conditions, it would make it hard for the drivers to maintain normal driving behavior and given the situation, they may be forced to speed and come close to other drivers, thus enters into dangerous driving phase (STZ 2).

At the same time, coping capacity is negatively associated with (inverse) risk (or normal driving), which is counter-intuitive (similar to what we have noted for acceleration). It could be assumed that higher coping capacity might improve normal driving (reduce risk) but this is not the case here. Furthermore, the coping capacity indicators in our sample include static demographic and self-reported behavior indicators and therefore are more representative of driver personality and general driving styles, and less so of the real-time operator state during the experiment. For instance, indicators related to the level of sleepiness, fatigue or distraction were either not available or not significant in this model. Therefore, it can be concluded that younger, and more confident drivers exhibited (higher risk) lower normal driving) in this experiment, in terms of exceeding the STZ headway boundaries, without however considering the variations of their state during these trips.

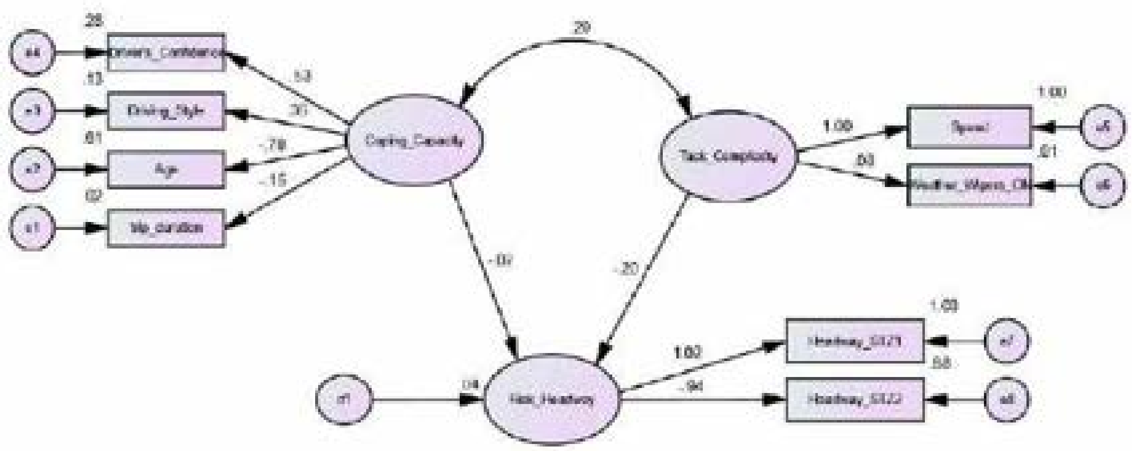

Figure 22: Results of SEM on Risk (headway STZ) – Belgian truck drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.989; TLI is 0.982 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.043. Table 50 summarizes the model fit of SEM applied for headway.

Table 50: Model Fit Summary for headway – Belgian truck drivers – experiment Phase 1

Model Fit measures	Value
AIC	3786.289
BCC	3786.293
CFI	0.989
TLI	0.982
RMSEA	0.043
Hoelter's critical N (α = .05)	866
Hoelter's critical N (α = .01)	1049

Residual variances details are presented in Table 51 that follows.

Table 51: Residual variances for headway – Belgian truck drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.132	0.002	68.053	***
Task_Complexity	120.093	14.234	8.437	***
Risk_Headway	0.024	0.000	126.04	***
trip_duration	4217.46	17.62	239.359	***
Age	44.615	0.738	60.488	***
Style	0.150	0.001	221.758	***
Confidence	0.334	0.002	169.667	***
Speed	0.468	14.226	0.033	0.974
ME_Car_wipers_median	0.01	0.000	238.774	***
Headway_1	-0.001	0.000	-7.567	***
iDreams_Headway_Map_level_2_mean	0.002	0.000	30.04	***

Figure 23, 24 and 25 show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the relationships among risk, task complexity and coping capacity are fairly consistent between the different phases, it is interesting to note however:

- The loading of age substantially reduces in phase 4. This might indicate that the impact of age on the coping capacity is compensated by the presence of all interventions and old drivers may perform normal driving.
- In phase 4, the structural relationship between coping capacity and risk changes to a positive coefficient. This is an interesting outcome. It is possible that the presence of all i-DREAMS interventions in phase 4 lead to a different interaction between the three latent variables. The combined effect of all interventions resulted in a positive relationship between coping capacity and risk (normal driving) and at the same time a negative relationship between task complexity and risk (normal driving).
- The loadings of the observed proportions of the STZ of headway are consistent between the different phases.

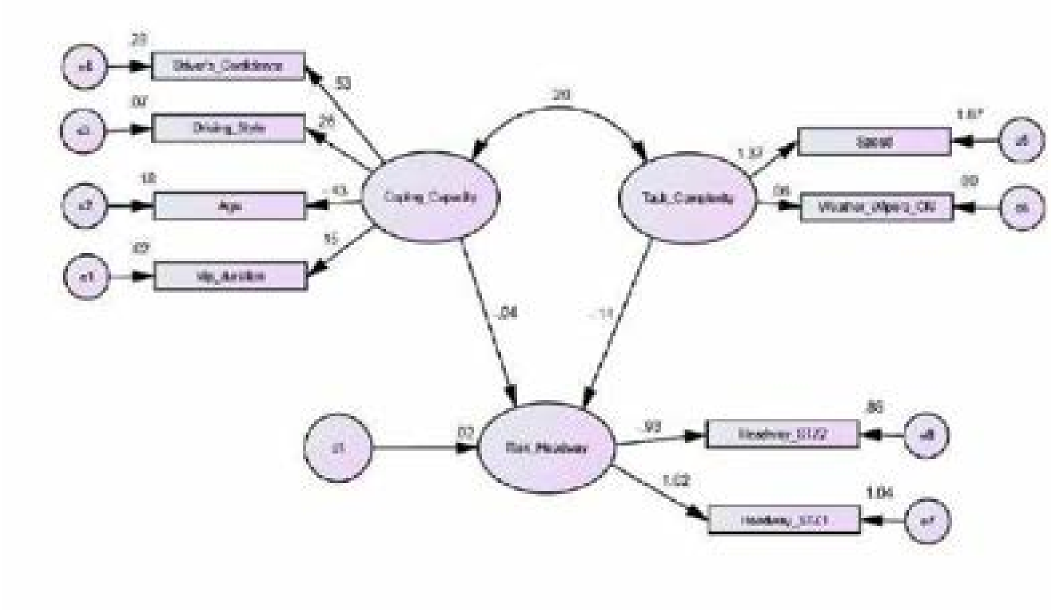

Figure 23: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.899; TLI is 0.834 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.080. Table 52 summarizes the model fit of SEM applied for headway.

Table 52: Model Fit Summary for headway – Belgian truck drivers – experiment Phase 2

Model Fit measures	Value
AIC	10742.257
BCC	10742.26
CFI	0.970
TLI	0.951
RMSEA	0.065
Hoelter's critical N (α = .05)	378
Hoelter's critical N (α = .01)	458

Residual variances details are presented in Table 53 that follows.

Table 53: Residual variances for headway – Belgian truck drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.136	0.003	53.516	***
Task_Complexity	262.388	48.028	5.463	***
Risk_headway	0.022	0.000	144.111	***
Age	85.9	0.454	189.275	***
Style	0.173	0.001	247.331	***
Confidence	0.353	0.003	137.058	***
Speed_mps	-122.019	48.03	-2.540	0.011
ME_Car_wipers_median	0.009	0.000	265.581	***
iDreams_Headway_Map_level_2_mean	0.002	0.000	36.839	***
Trip_duration	6969.862	26.51	262.91	***
Headway_1	-0.001	0.000	-7.945	***

The results for phase 3 are shown in Figure 24 below.

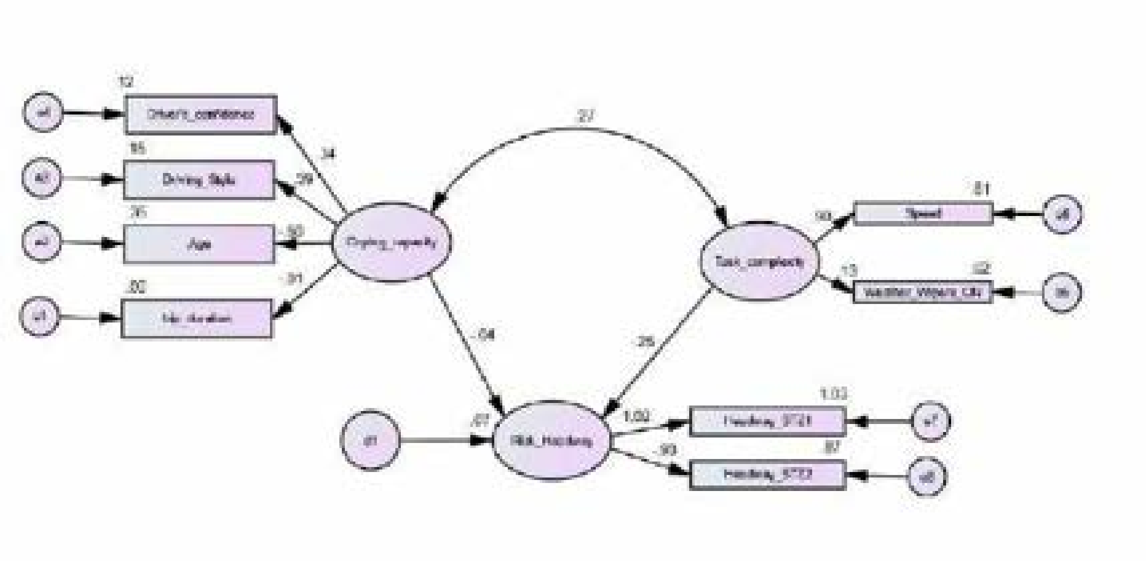

Figure 24: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.985; TLI is 0.969 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.046. Table 54 summarizes the model fit of SEM applied for headway.

Table 54: Model Fit Summary for headway – Belgian truck drivers – experiment Phase 3

Model Fit measures	Value
AIC	5150.702
BCC	5150.706
CFI	0.985
TLI	0.969
RMSEA	0.046
Hoelter's critical N (α = .05)	754
Hoelter's critical N (α = .01)	913

Residual variances details are presented in Table 55 that follows.

Table 55: Residual variances for headway – Belgian truck drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_capacity	0.059	0.001	41.073	***
Task_complexity	103.535	6.39	16.202	***
Risk_Headway	0.022	0.000	144.006	***
Trip_duration	10563.18	40.038	263.829	***
Age	60.330	0.630	95.762	***
Style	0.159	0.001	200.245	***
Confidence	0.441	0.002	219.023	***
Speed	23.766	6.373	3.729	***
ME_Car_wipers_median	0.007	0.000	252.951	***
Headway_mean_sum_1	-0.001	0.000	-8.187	***
iDreams_Headway_Map_level_2_mean	0.002	0.000	39.637	***

The results for phase 4 are shown in Figure 25 below.

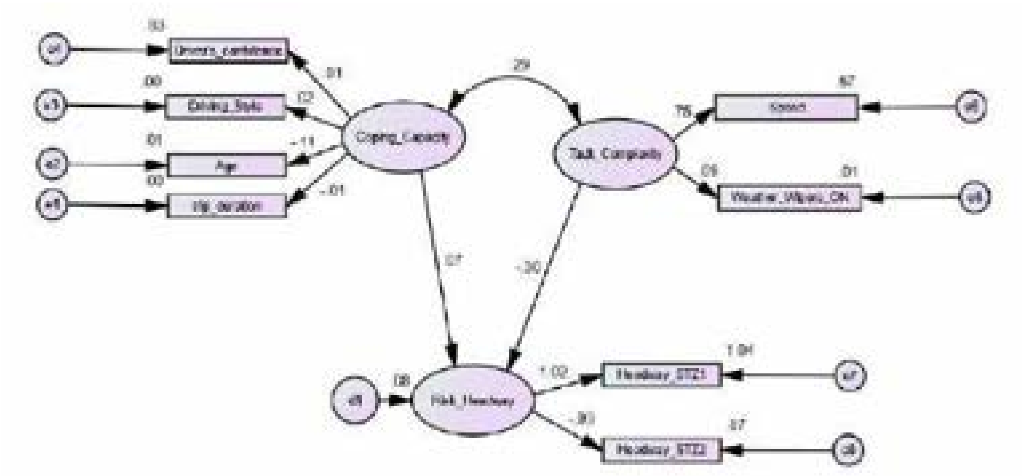

Figure 25: Results of SEM on Risk (headway STZ) – Belgian car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.943; TLI is 0.912 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.086. Table 56 summarizes the model fit of SEM applied for headway.

Table 56: Model Fit Summary for headway – Belgian truck drivers – experiment Phase 4

Model Fit measures	Value
AIC	25269.232
BCC	25269.235
CFI	0.943
TLI	0.912
RMSEA	0.086
Hoelter's critical N (α = .05)	215
Hoelter's critical N (α = .01)	259

Residual variances details are presented in Table 57 that follows.

Table 57: Residual variances for headway – Belgian truck drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
Coping_Capacity	0.457	0.041	11.26	***
Task_Complexity	83.287	5.371	15.506	***
Risk_Headway	0.026	0.000	120.696	***
trip_duration	10166.676	33.193	306.29	***
Age	122.768	0.423	290.111	***
Style	0.211	0.001	306.248	***
Confidence	0.095	0.041	2.351	0.019
Speed	63.75	5.358	11.899	***
ME_Car_wipers_median	0.005	0.000	302.361	***
Headway_1	-0.001	0.000	-9.476	***
iDreams_Headway_Map_level_2_mean	0.003	0.000	41.523	***

4.2.3 UK (Cars)

4.2.3.1 Headway

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity and coping capacity with risk where risk, expressed as the three phases of the STZ, was formed as a composite of headway. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 53 UK car drivers, 3,073 trips (56,853 minutes)
- Phase 2: real-time interventions - 54 UK car k drivers, 3,317 trips (58,458 minutes)
- Phase 3: real-time & post-trip interventions - 53 UK car drivers, 3,417 trips (59,556 minutes)
- Phase 4: real-time. post-trip interventions & gamification - 54 UK car drivers, 4,594 trips (93,974 minutes)

To begin with, a SEM analysis was performed based on data from 53 drivers and 3,073 trips, collected in phase 1 of the i-DREAMS project trials where no interventions were present. The model was developed in IBM SPSS Amos 27 Graphics software, and it is graphically described in Figure 26.

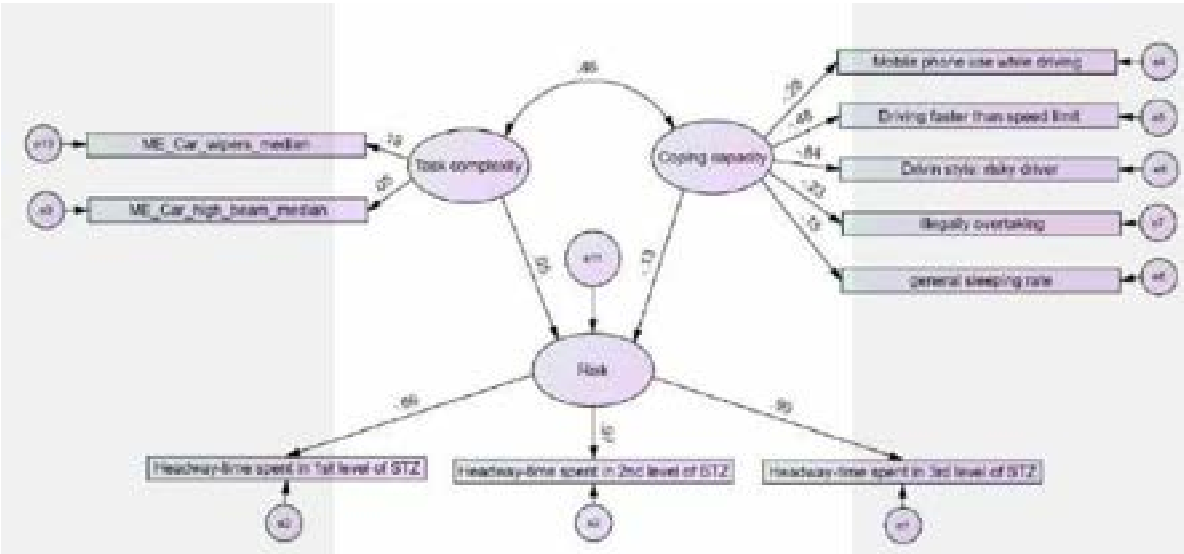

Figure 26: Results of SEM on Risk (headway STZ) – UK car drivers – experiment Phase 1

Maximum likelihood estimation method was employed. The presented model appears to be a good fit to the data. The Comparative Fit Index (CFI) is 0.984; TLI is 0.977 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.042. More details about the model fit can be found in Table 58 below.

Table 58: Model Fit Summary for headway – UK car drivers – experiment Phase 1

Model Fit measures	Value
AIC	6377.390
BIC	6599.142

Model Fit measures	Value
CFI	0.984
TLI	0.977
RMSEA	0.042
GFI	0.989
Hoelter's critical N (α = .05)	830
Hoelter's critical N (α = .01)	961

The results of SEM including residual variances details are presented in the Table 59 that follows.

Table 59: Residual variances for headway – UK car drivers – experiment Phase 1

Variable	Estimate	S.E.	C.R.	P
Coping_capacity	.004	.000	20.976	***
Task_complexity	.002	.001	3.170	.002
Risk	.035	.000	203.135	***
iDreams_Headway_Map_level_23_mean	.001	.000	20.421	***
iDreams_Headway_Map_level_1_mean	.003	.000	66.268	***
iDreams_Headway_Map_level_1_0_mean	.070	.000	234.307	***
EQ4e_Mobile_phone	.406	.002	227.243	***
EQ4b_Speed_limit	.672	.004	172.492	***
EQ5_Driving_style	.150	.005	32.465	***
EQ4g_Illegal_overtake	.266	.001	232.027	***
EQ17_General_sleep_rating	.202	.001	236.368	***
ME_Car_high_beam_median	.004	.000	232.933	***
ME_Car_wipers_median	.057	.001	80.601	***

All the observed indicators of the three latent variables task complexity, coping capacity and risk are statistically significant at 99% confidence level. The latent variables of task complexity and coping capacity have a statistically significant impact on risk that is significantly interpreted by the time spent in each of the three levels of STZ regarding the headway indicator. Coping capacity and task complexity are positively correlated (0.46).

Coping capacity seems to have a greater effect on risk than task complexity and the negative sign indicates that on cases that coping capacity increases, risk decreases. The opposite is observed for task complexity and risk as their positive relationship indicates that when driving task difficulty increases, risk also increases.

The latent construct of task complexity is represented by the indicator variables of High beam and wipers use. Wipers can be an indication of weather conditions, most specifically, they can be indicative of rain presence during the trip while high beams can indicate lighting conditions, for example, low visibility or dark. Both variables have a positive loading on the latent factor task complexity showing that an increase in the latter explains an increase in both of them accordingly.

Regarding coping capacity, all the indicator variables in the model show a negative relationship with risk except for general sleeping rate. Driver style appears to be the most important indicator (higher estimate) for coping capacity and risk development while also important indicators are the speeding (driving always above speed limit), the mobile phone usage while driving, the illegal overtaking and the general sleeping rate. The latter, as expected, has a positive relationship with coping capacity showing that better sleep habits are associated with increased levels of driver capability. Last but not least, according to the model increased level of risks are linked to increased time spent on second and third headway level of STZ.

Following the same approach, a SEM analysis was employed for driving data on phase 2 of the on-road trials (54 drivers, 3,317 trips) where intervention notifications have been introduced to the drivers. The model is graphically described in Figure 27.

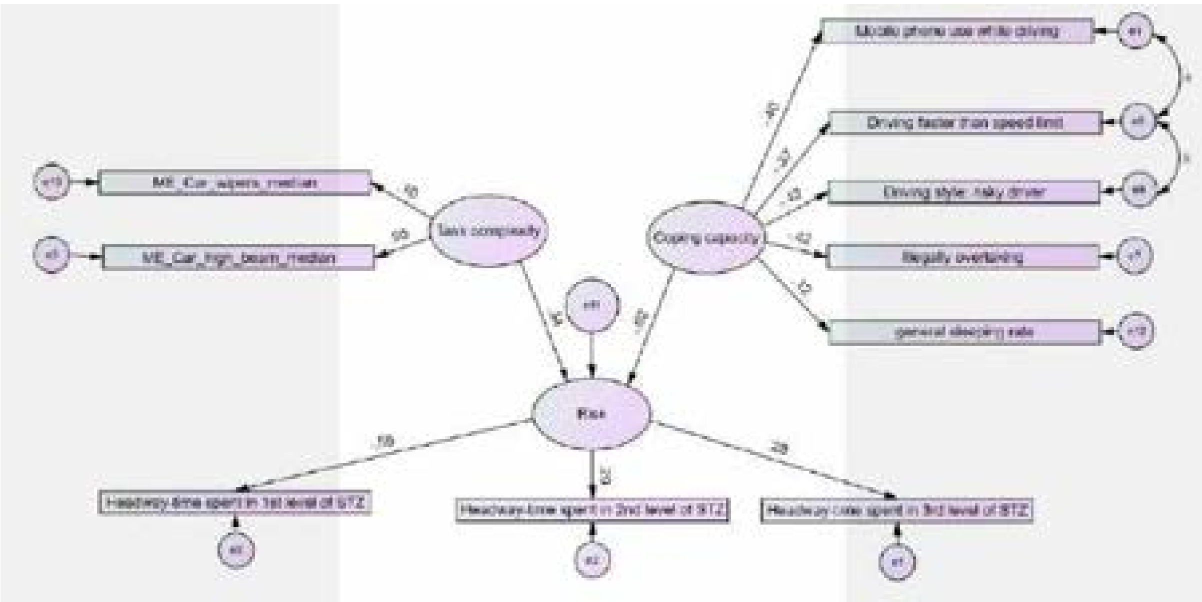

Figure 27: Results of SEM on Risk (headway STZ) – UK car drivers – experiment Phase 2

The results indicate that the model is reasonably consistent with the data as CFI is 0.885, TLI is 0.834, and RMSEA is 0.037. More details about the model fit can be found in Table 60 below.

Table 60: Model Fit Summary for headway – UK car drivers – experiment Phase 2

Model Fit measures	Value
AIC	4939.518
BIC	5171.580
CFI	0.885
TLI	0.834
RMSEA	0.037
GFI	0.992

Model Fit measures	Value
Hoelter's critical N (α = .05)	1076
Hoelter's critical N (α = .01)	1248

The results of SEM including residual variances details are presented in the Table 61 that follows.

Table 61: Residual variances for headway – UK car drivers – experiment Phase 2

Variable	Estimate	S.E.	C.R.	P
Coping_capacity	.003	.000	13.471	***
Task_complexity	.001	.000	2.732	.010
Risk	.005	.000	11.588	***
iDreams_Headway_Map_level_23_mean	.030	.000	190.514	***
iDreams_Headway_Map_level_1_mean	.064	.000	184.292	***
iDreams_Headway_Map_level_1_0_mean	.080	.002	51.613	***
EQ4e_Mobile_phone	.433	.003	164.485	***
EQ4b_Speed_limit	.693	.005	140.954	***
EQ5_Driving_style	.441	.003	150.146	***
EQ4g_Illegal_overtake	.243	.002	153.672	***
ME_Car_high_beam_median	.004	.000	222.508	***
ME_Car_wipers_median	.072	.001	133.416	***
EQ17_General_sleep_rating	.201	.001	236.729	***

Similarly with phase 1, all the observed indicators of the three latent variables task complexity, coping capacity and risk are statistically significant at 99% confidence level. The latent variables of task complexity and coping capacity have a statistically significant impact on risk that is significantly interpreted by the time spent in each of the three levels of STZ regarding the headway indicator.

The effect size of coping capacity to risk seems to be low (lower than phase 1) but statistically significant. On the contrary, the task complexity relates more strongly to risk than in phase 1 (0.34) indicating again that on cases that task complexity increases (wipers and high beam usage), risk also increases. The opposite is observed with coping capacity where when the latter increases, the risk decreases. Coping capacity and task complexity correlation is not supported in this model.

In terms of the indicators of the latent concepts, wipers appear to load stronger than high beams to task complexity as in phase 1 and again here, driver style is the stronger factor following closely by illegal overtaking, mobile phone use, speed limit and lastly general sleeping rate. Lower risk seems to be associated with higher time in the first level of STZ regarding headway.

Another SEM analysis was employed for data from 53 drivers and 3,417 trips included in phase 3 of the on-road trials where drivers can interact with i-dreams smart phone application. The model is graphically described in Figure 28.

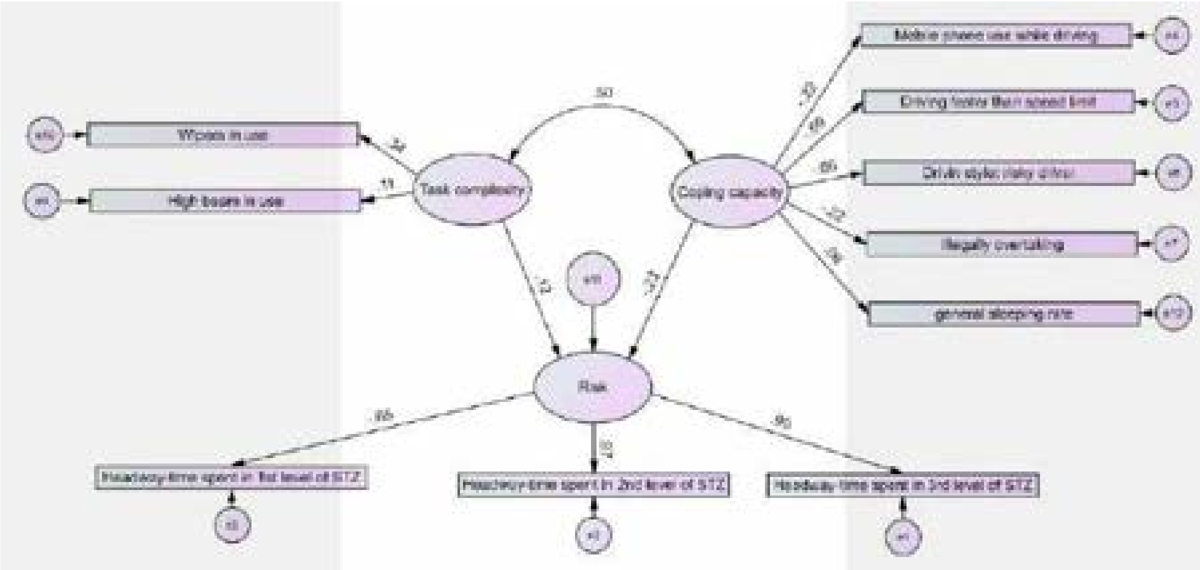

Figure 28: Results of SEM on Risk (headway STZ) – UK car drivers – experiment Phase 3

The results indicate that the model is consistent with the data as CFI is 0.988, TLI is 0.983, and RMSEA is 0.037. More details about the model fit can be found in Table 62 below.

Table 62: Model Fit Summary for headway – UK car drivers – experiment Phase 3

Model Fit measures	Value
AIC	5266.238
BIC	5489.058
CFI	0.988
TLI	0.983
RMSEA	0.037
GFI	0.991
Hoelter's critical N (α = .05)	1055
Hoelter's critical N (α = .01)	1221

The results of SEM including residual variances details are presented in the Table 63 that follows.

Table 63: Residual variances for headway – UK car drivers – experiment Phase 3

Variable	Estimate	S.E.	C.R.	P
Coping_capacity	.001	.000	10.803	***
Task_complexity	.010	.001	9.871	***

Variable	Estimate	S.E.	C.R.	P
Risk	.030	.000	195.116	***
iDreams_Headway_Map_level_23_mean	.001	.000	17.028	***
iDreams_Headway_Map_level_1_mean	.002	.000	60.390	***
iDreams_Headway_Map_level_1_0_mean	.063	.000	240.912	***
EQ4e_Mobile_phone	.419	.002	235.687	***
EQ4b_Speed_limit	.538	.004	121.442	***
EQ5_Driving_style	.320	.003	92.698	***
EQ4g_Illegal_overtake	.292	.001	235.136	***
ME_Car_high_beam_median	.004	.000	234.789	***
ME_Car_wipers_median	.078	.001	73.084	***
EQ17_General_sleep_rating	.203	.001	242.925	***

As in the two previous phases, all the observed indicators of the three latent variables task complexity, coping capacity and risk are statistically significant at 99% confidence level. The latent variables of task complexity and coping capacity have a statistically significant impact on risk that is significantly interpreted by the time spent in each of the three levels of STZ regarding the headway indicator. More specifically, higher risk is translating in more time spent in second and third level of STZ. In this model again, driving task difficulty affects positively (increases) the levels of risk while the opposite stands for coping capacity as expected.

As in phase 1, the effect of coping capacity on risk (standardised coefficient=0.23) is greater than this of task complexity (standardised coefficient=0.12) and coping capacity and task complexity are positively correlated (0.51). Wipers and high beam use show a positive relationship with task complexity and in accordance with risk while driving style, driving above speed limit, mobile phone use while driving and illegal overtaking are negatively related to coping capacity and in turn with risk.

Lastly, a SEM analysis was performed for driving data on phase 4 (54 drivers, 4,594 trips) of the on-road trials where gamification was available for the app. The model is graphically described in Figure 29.

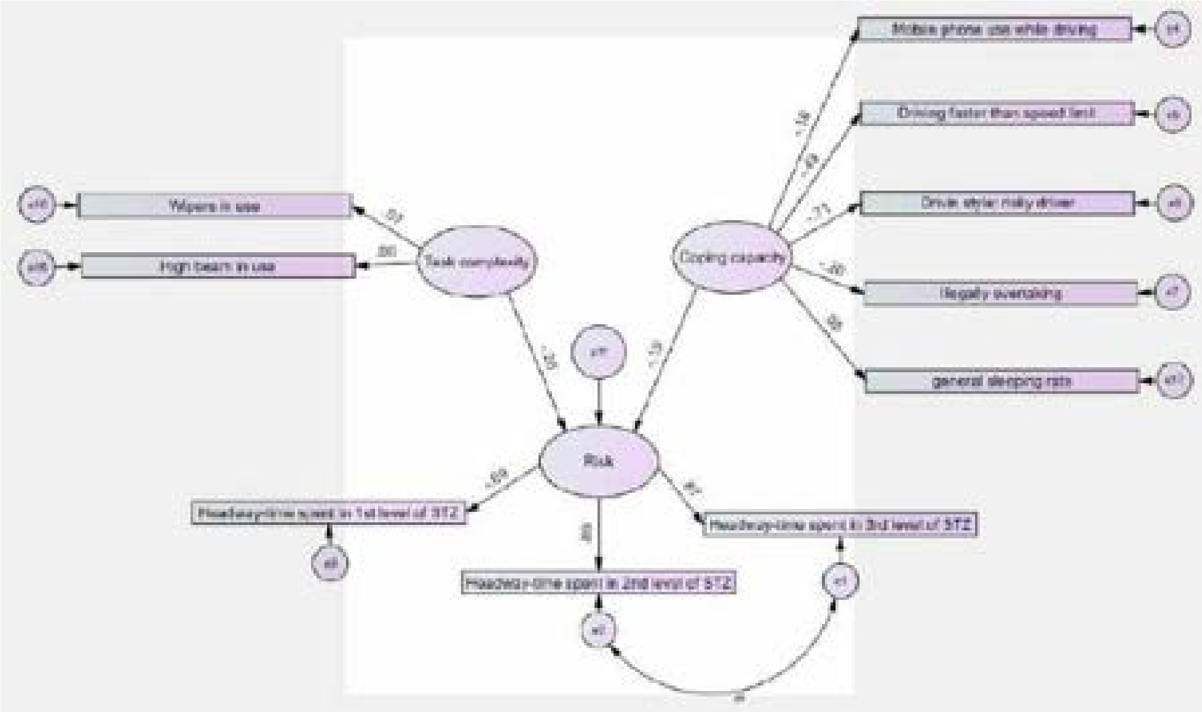

Figure 29: Results of SEM on Risk (headway STZ) – UK car drivers – experiment Phase 4

Maximum likelihood estimation method was employed. The results indicate that the model is consistent with the data as CFI is 0.989, TLI is 0.985, and RMSEA is 0.035. More details about the model fit can be found in Table 64 below.

Table 64: Model Fit Summary for headway – UK car drivers – experiment Phase 4

Model Fit measures	Value
AIC	7536.846
BIC	7770.156
CFI	0.989
TLI	0.985
RMSEA	0.035
GFI	0.992
Hoelter's critical N (α = .05)	1160
Hoelter's critical N (α = .01)	1342

The results of SEM including residual variances details are presented in the Table 65 that follows.

Table 65: Residual variances for headway – UK car drivers – experiment Phase 4

Variable	Estimate	S.E.	C.R.	P
Coping_capacity	.001	.000	8.570	***
Task_complexity	.000	.000	2.309	***
Risk	.018	.001	26.062	***
iDreams_Headway_Map_level_23_mean	.007	.000	17.030	***

Variable	Estimate	S.E.	C.R.	P
iDreams_Headway_Map_level_1_mean	.007	.000	19.519	***
iDreams_Headway_Map_level_1_0_mean	.056	.001	60.500	***
EQ4e_Mobile_phone	.401	.001	300.081	***
EQ4b_Speed_limit	.573	.003	174.136	***
EQ5_Driving_style	.255	.004	64.743	***
EQ4g_Illegal_overtake	.281	.001	296.405	***
ME_Car_high_beam_median	.053	.000	252.211	***
ME_Car_wipers_median	.004	.000	231.432	***
EQ17_General_sleep_rating	.201	.001	305.965	***

All the observed indicators presented in the model to represent the three latent concepts of task complexity, coping capacity and risk are statistically significant at 99.9% confidence level. Task complexity and coping capacity have a statistically significant impact on risk that is significantly interpreted by the time spent in each of the three levels of STZ regarding the headway indicator. As mentioned before in previous phases, lower risk relates to more time in the first level of STZ, in other words, to higher headways measurements. Similarly to phase 2, task complexity has a greater effect (standardised coefficient=-0.26) on risk than coping capacity (standardised coefficient=-0.19).

In terms of the relationship between driving task complexity and risk the picture is different than in the other three phases. The model for phase 4 indicates that increased levels of driving task difficulty, related to weather and visibility conditions, are linked to lower levels of risk. This result could be interpreted by the fact that when drivers have to face more complicated road conditions such as rain or lower visibility, they could become more alerted and cautious.

Regarding the specific indicators of the latent concept of coping capacity, the same pattern can be observed as in all other phases with the driver style to dominate in the coping capacity latent construct. The wipers and high beam use are positively related to task complexity while mobile phone use while driving, driving faster than the speed limit, driver style and illegal overtaking are all negatively related to coping capacity as it was intuitive. Furthermore, good sleeping rate is positively associated with driver capacity.

4.2.4 Germany (Cars)

4.2.4.1 Speeding

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three phases of the STZ) of speeding. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 28 German car drivers, 1,397 trips (23,617 minutes)

- Phase 2: real-time interventions - 28 German car drivers, 1,322 trips (19,469 minutes)
- Phase 3: real-time & post-trip interventions - 27 German car drivers, 1,129 trips (17,704 minutes)
- Phase 4: real-time. post-trip interventions & gamification - 28 German car drivers, 1,496 trips (23,644 minutes)

To begin with, the results for phase 1 are shown in Figure 30 below. Risk is measured by means of the STZ levels for speeding (level 1 'normal driving' used as the reference case; level 2 refers to 'dangerous driving', while no incidents with regards to level 3 'avoidable accident driving' were found). In particular, positive correlations of risk with the STZ indicators were found. It should be noted that the identified model indicated that level 3 of speeding variable does not have significant loading in the measurement model for the latent variable risk and thus, this level was not included in the final model. Level 1 and level 2 of speeding (or STZ1 and STZ 2 indicators) have positive loadings in relationship to the latent variable Risk, respectively.

To begin with, the latent variable task complexity is measured by means of the environmental indicator of "ME_AWS_time_indicator_median" (indicating time of the day). It should be noted that based on the definition of task complexity, road layout, time, location, traffic volumes and weather variables should be included in the analysis. However, road type (i.e. urban, rural, highway), location, traffic volumes (i.e. high, medium, low) and weather were not available in German dataset. Thus, only the time indicator was able to be used in the models applied. To that aim, exposure indicators, such as trip duration and distance traveled were included in the task complexity analysis. In particular, time of the day, distance and duration found to have a positive correlation with task complexity.

Furthermore, it is shown that the latent coping capacity is measured by means of both vehicle state indicators, such as "VehicleAge" (indicating the age of the vehicle), "Gearbox" (indicating the type of gearbox; automatic or manual) and "Fuel_type" (indicating the type of fuel; diesel, hybrid electric, petrol). At the same time, operator state indicators, such as "Gender" (indicating the gender of the driver; male or female) and "Age" (indicating the age of the driver) are included in the SEM applied.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation (regression coefficient=0.03) – which reduces in magnitude as the driver's progress from phases 1 and 2 though phases 3 and 4. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. Overall, the structural model between task complexity and risk shows a positive coefficient, which means that increased task complexity relates to increased risk according to the model (regression coefficient=2.19). On the other hand, 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 (regression coefficient=-0.05).

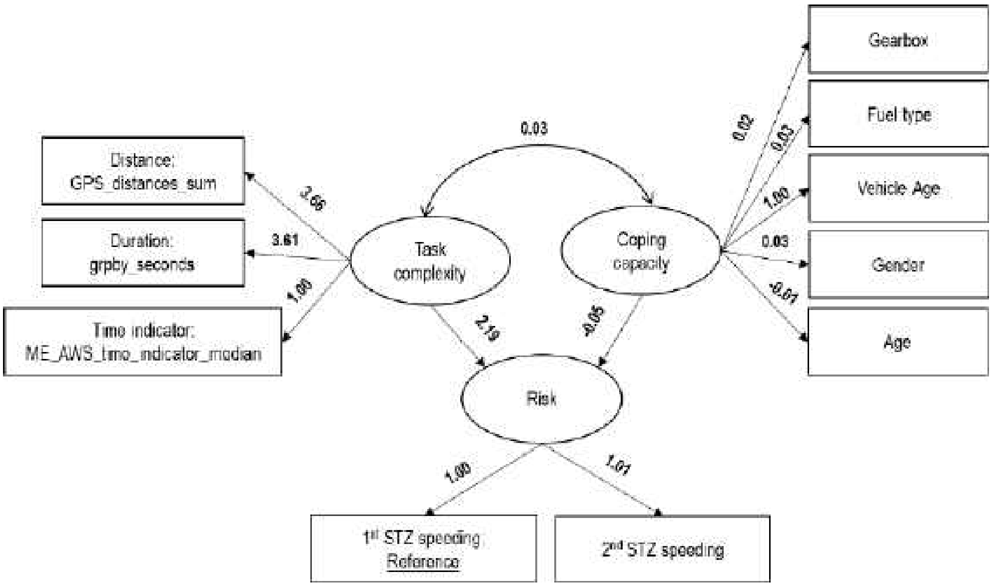

Figure 30: Results of SEM on Risk (Speeding STZ) – German car drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.981; TLI is 0.974 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.079. Table 66 summarizes the model fit of SEM applied for speeding.

Table 66: Model Fit Summary for speeding – German car drivers – experiment Phase 1

Model Fit measures	Value
AIC	813827.574
BIC	814118.257
CFI	0.981
TLI	0.974
RMSEA	0.079
GFI	0.961
Hoelter's critical N (α = .05)	234.136
Hoelter's critical N (α = .01)	270.935

Residual variances details are presented in Table 67 that follows.

Table 67: Residual variances for speeding – German car drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.977	0.008	117.885	< .001
grpby_seconds	0.706	0.007	94.961	< .001
GPS_distances_sum	0.715	0.008	90.985	< .001
VehicleAge	-11.131	3.379	-3.294	< .001
Age	0.998	0.006	156.926	< .001
Fuel_type	0.990	0.007	143.881	< .001
Gearbox	0.998	0.006	156.668	< .001
Gender	0.989	0.007	142.970	< .001
iDreams_Speeding_Map_level_0_sum	0.008	2.111×10^{-4}	35.672	< .001
iDreams_Speeding_Map_level_1_sum	-0.008	2.141×10^{-4}	-35.154	< .001

The following Figures show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the measurement equations of task complexity and coping capacity are fairly consistent between the different phases. At the same time, the loadings of the observed proportions of the STZ of speeding are consistent between the different phases. The structural model between task complexity and inverse risk (normal driving) are positively correlated among the four phases, while coping capacity and risk found to have a negative relationship in all phases of the experiment. The results for phase 2 are shown in Figure 31 below.

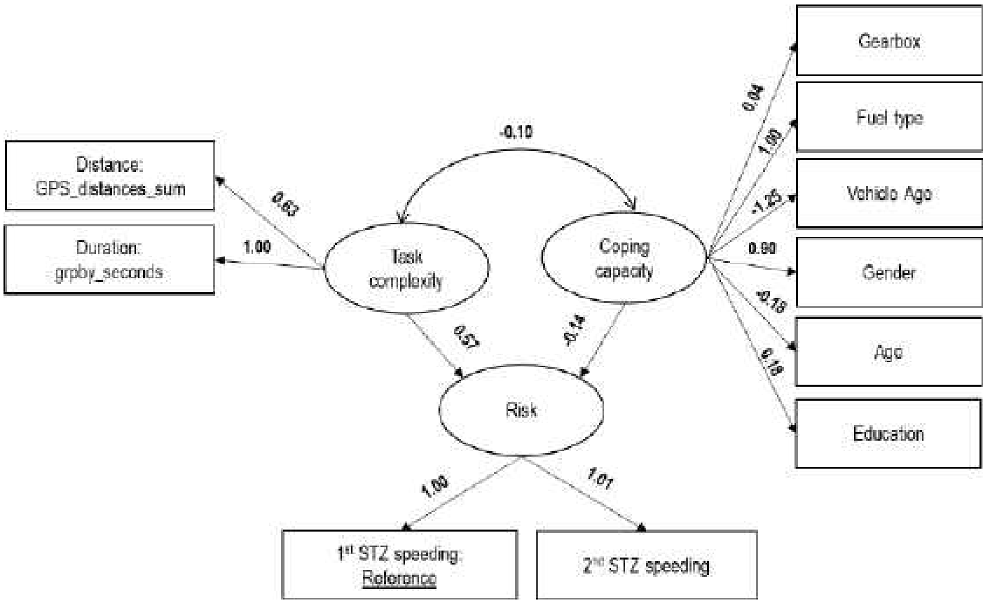

Figure 31: Results of SEM on Risk (Speeding STZ) – German car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.960; TLI is 0.944 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.117. Table 68 summarizes the model fit of SEM applied for speeding.

Table 68: Model Fit Summary for speeding – German car drivers – experiment Phase 2

Model Fit measures	Value
AIC	676463.527
BIC	676746.197
CFI	0.960
TLI	0.944
RMSEA	0.117
GFI	0.920
Hoelter's critical N (α = .05)	106.728
Hoelter's critical N (α = .01)	123.417

Residual variances details are presented in Table 69 that follows.

Table 69: Residual variances for speeding – German car drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
grpby_seconds	0.646	0.009	69.334	< .001
GPS_distances_sum	0.861	0.007	116.677	< .001
Fuel_type	1.258	0.013	96.027	< .001

Variable	Estimate	Std. Error	z-value	P(>|z|)
VehicleAge	1.401	0.018	79.814	< .001
Gearbox	1.000	0.007	139.155	< .001
Age	1.008	0.007	137.721	< .001
Gender	1.208	0.012	103.490	< .001
Education	1.008	0.007	137.647	< .001
iDreams_Speeding_Map_level_0_sum	0.009	2.543×10^{-4}	37.215	< .001
iDreams_Speeding_Map_level_1_sum	-0.010	2.589×10^{-4}	-36.706	< .001

The results for phase 3 are shown in Figure 32 below.

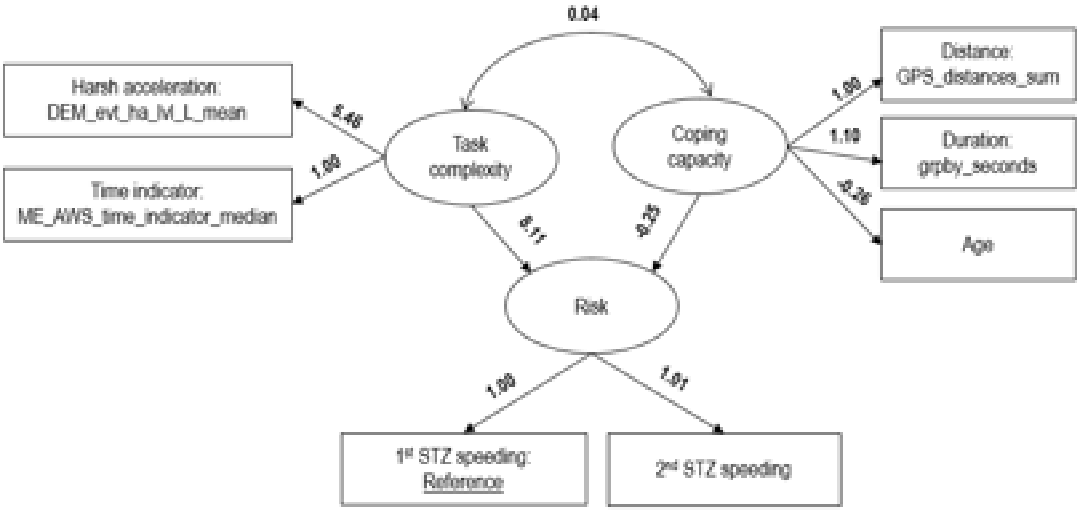

Figure 32: Results of SEM on Risk (Speeding STZ) – German car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.996; TLI is 0.993 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.059. Table 70 summarizes the model fit of SEM applied for speeding.

Table 70: Model Fit Summary for speeding – German car drivers – experiment Phase 3

Model Fit measures	Value
AIC	282420.347
BIC	282625.175
CFI	0.996
TLI	0.993
RMSEA	0.059
GFI	0.983
Hoelter's critical N (α = .05)	507.651
Hoelter's critical N (α = .01)	637.688

Residual variances details are presented in Table 71 that follows.

Table 71: Residual variances for speeding – German car drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.993	0.010	94.720	< .001
DEM_evt_ha_lvl_L_mean	0.736	0.036	20.194	< .001

Variable	Estimate	Std. Error	z-value	P(>|z|)
GPS_distances_sum	0.720	0.008	94.682	< .001
Age	0.980	0.008	129.041	< .001
Grpby_seconds	0.645	0.008	81.986	< .001
iDreams_Speeding_Map_level_0_sum	0.007	1.169×10^{-4}	56.001	< .001
iDreams_Speeding_Map_level_1_sum	-0.007	1.182×10^{-4}	-55.315	< .001

The results for phase 4 are shown in Figure 33 below.

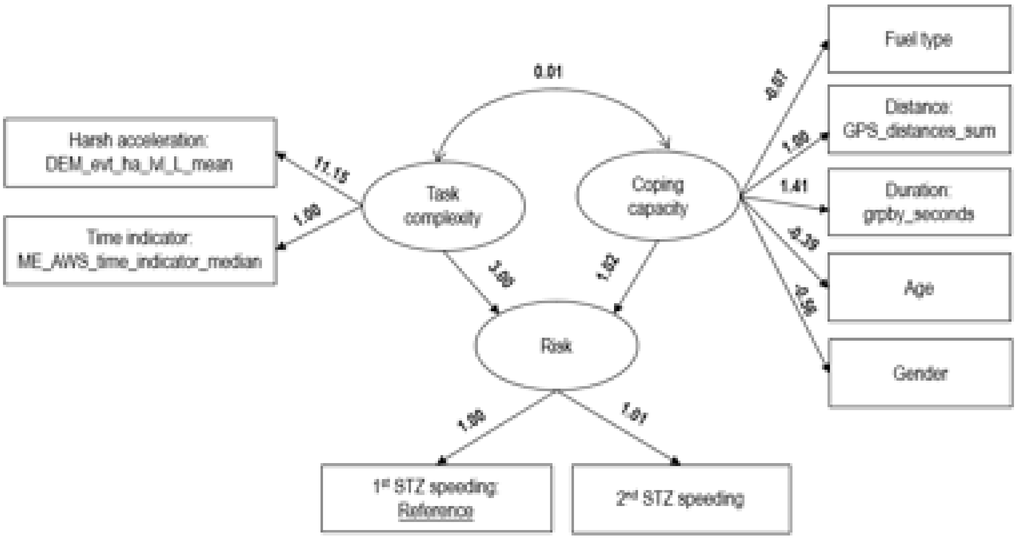

Figure 33: Results of SEM on Risk (Speeding STZ) – German car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.978; TLI is 0.966 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.100. Table 72 summarizes the model fit of SEM applied for speeding.

Table 72: Model Fit Summary for speeding – German car drivers – experiment Phase 4

Model Fit measures	Value
AIC	525983.888
BIC	526243.996
CFI	0.978
TLI	0.966
RMSEA	0.100
GFI	0.943
Hoelter's critical N (α = .05)	153.470
Hoelter's critical N (α = .01)	180.957

Residual variances details are presented in Table 73 that follows.

Table 73: Residual variances for speeding – German car drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.995	0.009	108.205	< .001
DEM_evt_ha_lvl_L_mean	0.224	0.115	1.940	0.052

Variable	Estimate	Std. Error	z-value	P(>|z|)
GPS_distances_sum	0.829	0.007	123.000	< .001
Grpby_seconds	0.639	0.008	84.763	< .001
Gender	0.944	0.007	141.467	< .001
Age	0.974	0.008	123.311	< .001
Fuel_type	0.999	0.008	125.390	< .001
iDreams_Speeding_Map_level_0_sum	0.006	9.887×10^{-5}	63.609	< .001
iDreams_Speeding_Map_level_1_sum	-0.006	9.984×10^{-5}	-62.838	< .001

4.2.5 Greece (Cars)

4.2.5.1 Speeding

Three separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three phases of the STZ) of speeding. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 65 Greek car drivers, 2,937 trips (51,786 minutes)
- Phase 3: real-time & post-trip interventions - 65 Greek car drivers, 3,935 trips (69,962 minutes)
- Phase 4: real-time. post-trip interventions & gamification - 65 Greek car drivers, 2,194 trips (39,695 minutes)

The results for phase 1 are shown in Figure 34 below. Risk is measured by means of the STZ 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’), with positive correlations of Risk with the STZ indicators.

To begin with, the latent variable task complexity is measured by means of the environmental indicators “ME_AWS_time_indicator_median” (indicating time of the day). The exposure indicator of trip duration was also included in the task complexity analysis. In particular, time of the day and duration had a positive correlation with task complexity. Moreover, the latent coping capacity is measured by means of operator state indicators, such as distance, harsh acceleration, harsh braking, age and gender. At the same time, the indicators of coping capacity - vehicle state, such as vehicle age, gearbox or fuel type are included in the SEM applied.

The structural model between the latent variables shows some interesting findings. First of all, task complexity and coping capacity are inter-related with a positive correlation (regression coefficient=0.56). This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases.

Overall, the structural model between task complexity and risk shows a positive coefficient, which means that increased task complexity relates to increased risk according to the model (regression coefficient=0.69). On the other hand, 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 (regression coefficient=-0.35).

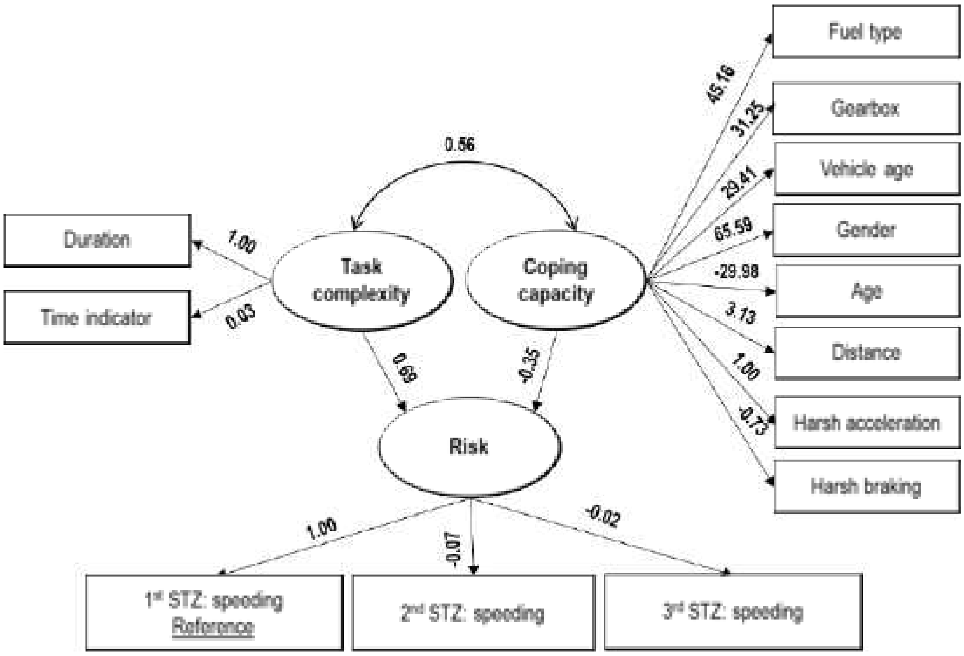

Figure 34: Results of SEM on Risk (Speeding STZ) – Greek car drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.840; TLI is 0.798 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.089. Table 74 summarizes the model fit of SEM applied for speeding.

Table 74: Model Fit Summary for speeding – Greek car drivers – experiment Phase 1

Model Fit measures	Value
AIC	692252.677
BIC	692590.360
CFI	0.840
TLI	0.798
RMSEA	0.089
GFI	0.925
Hoelter's critical N (α = .05)	164.309
Hoelter's critical N (α = .01)	183.214

Residual variances details are presented in Table 75 that follows.

Table 75: Residual variances for speeding – Greek car drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
grpby_seconds	-13.639	52.128	-0.262	0.794
GPS_distances_sum	0.998	0.011	89.383	< .001
ME_AWS_time_indicator_median	1.000	0.009	106.909	< .001
DEM_evt_ha_lvl_H_mean	1.000	0.013	76.550	< .001
Age	0.862	0.009	101.232	< .001
Gender	0.299	0.010	29.136	< .001
Fuel_type	0.674	0.008	84.400	< .001
VehicleAge	0.864	0.009	101.297	< .001
Gearbox	0.849	0.008	100.479	< .001
DEM_evt_hb_lvl_H_mean	1.000	0.013	76.556	< .001
iDreams_Speeding_Map_level_0_mean	-9.548	3.697	-2.583	0.010
iDreams_Speeding_Map_level_1_mean	0.920	0.030	31.114	< .001
iDreams_Speeding_Map_level_2_mean	0.964	0.010	94.063	< .001

The following Figures show the results of the 3rd and 4th phase of the experiment. It is observed that the measurement equations of task complexity and coping capacity are fairly consistent between the different phases. At the same time, the loadings of the observed proportions of the STZ of speeding are consistent between the different phases. The structural model between task complexity and inverse risk (normal driving) are positively correlated among the three phases, while coping capacity and risk found to have a negative relationship in all phases of the experiment. The results for phase 3 are shown in Figure 35 below.

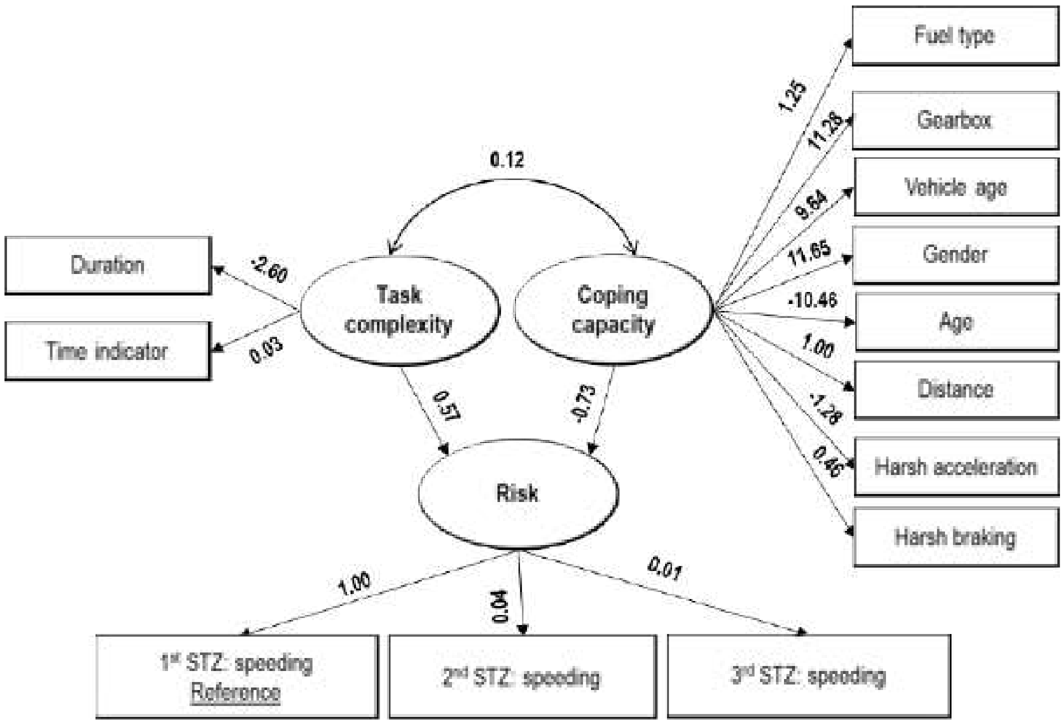

Figure 35: Results of SEM on Risk (Speeding STZ) – Greek car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.811; TLI is 0.762 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.092. Table 76 summarizes the model fit of SEM applied for speeding.

Table 76: Model Fit Summary for speeding – Greek car drivers – experiment Phase 3

Model Fit measures	Value
AIC	$2.268\times10^{+6}$
BIC	$2.268\times10^{+6}$
CFI	0.811
TLI	0.762
RMSEA	0.092
GFI	0.908
Hoelter's critical N (α = .05)	154.927
Hoelter's critical N (α = .01)	172.746

Residual variances details are presented in Table 77 that follows.

Table 77: Residual variances for speeding – Greek car drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.951	0.005	176.006	< .001
grpby_seconds	0.667	0.016	42.411	< .001
GPS_distances_sum	0.997	0.005	191.252	< .001
Age	0.629	0.004	151.007	< .001
Gender	0.540	0.004	130.780	< .001
Fuel_type	0.995	0.005	194.984	< .001
VehicleAge	0.685	0.004	161.191	< .001
Gearbox	0.569	0.004	137.849	< .001
DEM_evt_hb_lvl_H_mean	0.999	0.008	129.155	< .001
DEM_evt_ha_lvl_H_mean	0.995	0.008	129.341	< .001
iDreams_Speeding_Map_level_0_mean	21.018	8.341	2.520	0.012
iDreams_Speeding_Map_level_1_mean	1.038	0.017	61.946	< .001
iDreams_Speeding_Map_level_2_mean	0.957	0.006	160.887	< .001

The results for phase 4 are shown Figure 36 below.

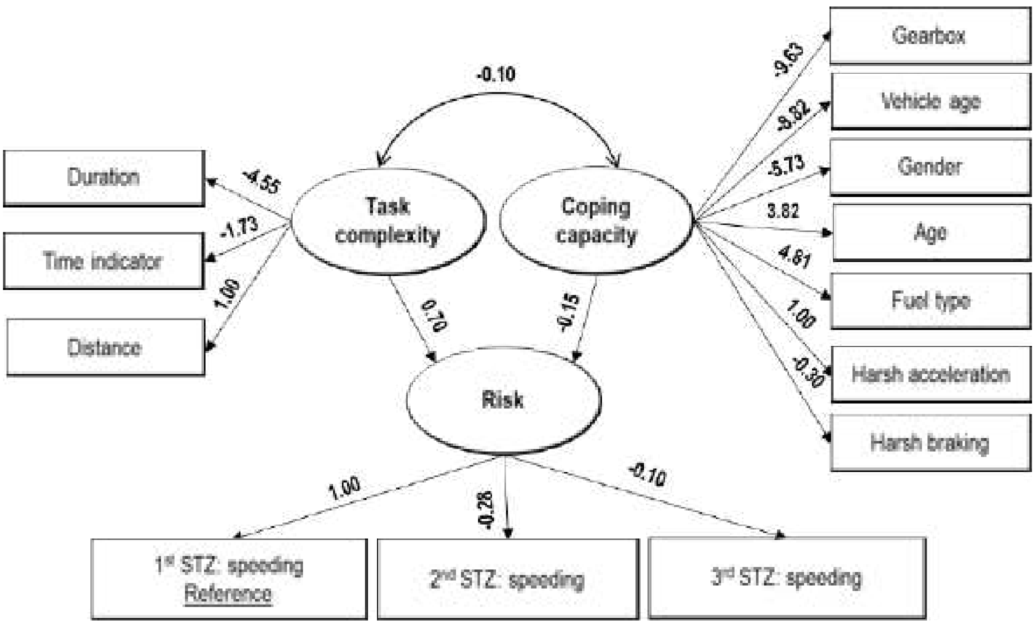

Figure 36: Results of SEM on Risk (Speeding STZ) – Greek car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.809, TLI is 0.759 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.111. Table 78 summarizes the model fit of SEM applied for speeding.

Table 78: Model Fit Summary for speeding – Greek car drivers – experiment Phase 4

Model Fit measures	Value
AIC	$4.326\times10^{+6}$
BIC	$4.326\times10^{+6}$
CFI	0.809
TLI	0.759
RMSEA	0.111
GFI	0.872
Hoelter's critical N (α = .05)	107.037
Hoelter's critical N (α = .01)	119.311

Residual variances details are presented in Table 79 that follows.

Table 79: Residual variances for speeding – Greek car drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
GPS_distances_sum	0.952	0.004	268.224	< .001
grpby_seconds	0.058	0.007	8.939	< .001
ME_AWS_time_indicator_median	0.863	0.003	267.069	< .001
DEM_evt_ha_lvl_H_mean	0.995	0.006	168.036	< .001
Age	0.881	0.003	274.527	< .001
Gender	0.731	0.003	263.853	< .001
Fuel_type	0.811	0.003	270.086	< .001
VehicleAge	0.363	0.002	188.467	< .001
Gearbox	0.240	0.002	129.755	< .001
DEM_evt_hb_lvl_H_mean	1.000	0.006	167.717	< .001
iDreams_Speeding_Map_level_0_mean	-2.192	0.073	-30.049	< .001
iDreams_Speeding_Map_level_1_mean	0.758	0.006	120.858	< .001
iDreams_Speeding_Map_level_2_mean	0.925	0.004	219.243	< .001

4.2.6 Portugal (Buses)

4.2.6.1 Headway

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three phases of the STZ) of headway. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 29 Portuguese bus drivers, 2,459 trips (202,532 minutes)
- Phase 2: real-time interventions - 29 Portuguese bus drivers, 1,363 trips (123,132 minutes)
- Phase 3: real-time & post-trip interventions - 26 Portuguese bus drivers, 1,411 trips (145,934 minutes)

- Phase 4: real-time. post-trip interventions & gamification - 22 Portuguese bus drivers, 2,098 trips (23,2323 minutes)

To begin with, the results for phase 1 are shown in Figure 37 below. Risk is measured by means of the STZ levels for headway (level 1 'normal driving' used as the reference case; level 2 refers to 'dangerous driving', while level 3 refers to 'avoidable accident driving'. In particular, negative correlations of risk with the STZ indicators were found.

The latent variable task complexity is measured by means of the environmental indicator of "ME_AWS_time_indicator_median" (indicating time of the day) and total duration. It should be noted that based on the definition of task complexity, road layout, time, location, traffic volumes and weather variables should be included in the analysis. However, road type (i.e. urban, rural, highway), location, traffic volumes (i.e. high, medium, low) and weather were not available in Portuguese dataset. Thus, only the time indicator was able to be used in the models applied. To that aim, exposure indicators, such as trip duration was included in the task complexity analysis. In particular, time of the day and duration found to have a positive correlation with task complexity.

Moreover, it is shown that the latent coping capacity is measured by means of operator state indicators, such as average speed, distance, harsh acceleration and harsh braking. It should be noted that vehicle state indicators, such as vehicle age, gearbox, type of fuel or socio-demographic characteristics were not provided.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation (regression coefficient=0.96) – which reduces in magnitude as the driver's progress from phases 1 and 2 though phases 3 and 4. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. Overall, the structural model between task complexity and risk shows a positive coefficient, which means that increased task complexity relates to increased risk according to the model (regression coefficient=5.36). On the other hand, 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 (regression coefficient=-5.02).

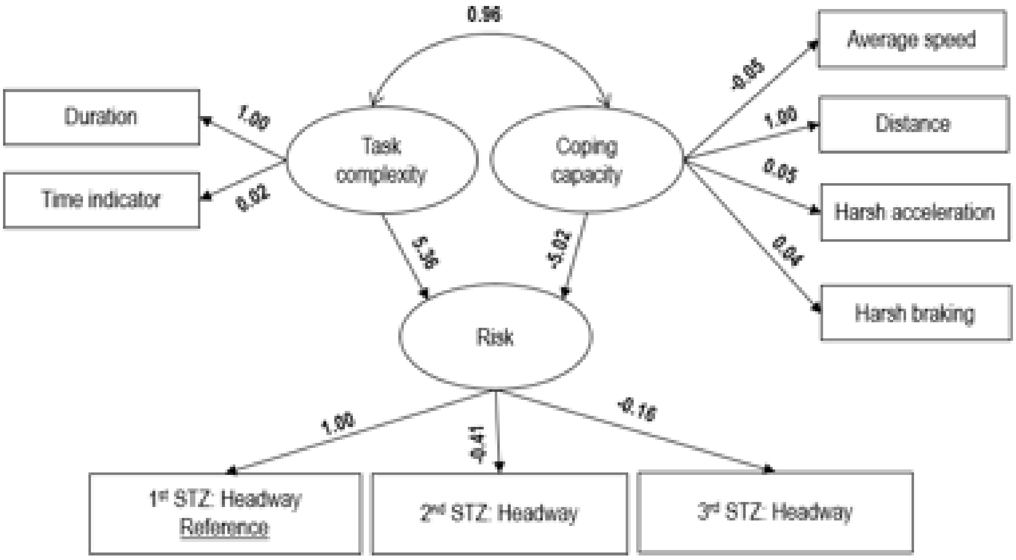

Figure 37: Results of SEM on Risk (Headway STZ) – Portuguese bus drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.983; TLI is 0.974 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.053. Table 80 summarizes the model fit of SEM applied for headway.

Table 80: Model Fit Summary for headway – Portuguese bus drivers – experiment Phase 1

Model Fit measures	Value
AIC	3.328×10+6
BIC	3.328×10+6
CFI	0.983
TLI	0.974
RMSEA	0.053
GFI	0.985
Hoelter's critical N (α = .05)	533.123
Hoelter's critical N (α = .01)	629.053

Residual variances details are presented in Table 81 that follows.

Table 81: Residual variances for headway – Portuguese bus drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
Duration	0.007	0.026	0.292	0.771
ME_AWS_time_indicator	1.000	0.004	277.573	< .001
Distance	0.095	0.007	14.439	< .001
GPS_spd	0.998	0.004	277.623	< .001
DrivingEvents_Map_evt_ha_mean	0.998	0.004	277.623	< .001
DrivingEvents_Map_evt_hb_mean	0.999	0.004	277.610	< .001
iDreams_Headway_Map_level_0_mean	-0.283	0.005	-52.424	< .001
iDreams_Headway_Map_level_1_mean	0.785	0.003	266.843	< .001
iDreams_Headway_Map_level_2_mean	0.967	0.003	279.715	< .001

The following Figures show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the measurement equations of task complexity and coping capacity are fairly consistent between the different phases. The structural model between task complexity and inverse risk (normal driving) are positively correlated in phases 1, 3 and 4, while a negative corellataion of phase 2 was identified. At the same time, coping capacity and risk found to have a negative relationship in all phases of the experiment. The results for phase 2 are shown in Figure 38 below.

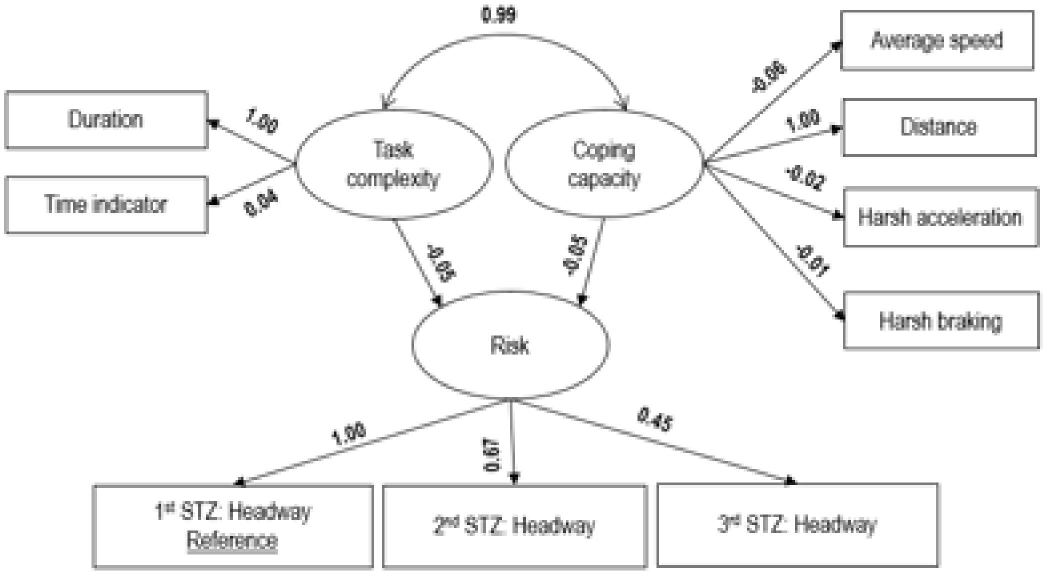

Figure 38: Results of SEM on Risk (Headway STZ) – German car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.985; TLI is 0.978 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.052. Table 82 summarizes the model fit of SEM applied for headway.

Table 82: Model Fit Summary for headway – Portuguese bus drivers – experiment Phase 2

Model Fit measures	Value
AIC	1.699×10+6
BIC	1.699×10+6
CFI	0.985
TLI	0.978
RMSEA	0.052
GFI	0.986
Hoelter's critical N (α = .05)	556.489
Hoelter's critical N (α = .01)	656.631

Residual variances details are presented in Table 83 that follows.

Table 83: Residual variances for headway – Portuguese bus drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
Duration	0.185	0.017	10.852	< .001
ME_AWS_time_indicator	0.998	0.005	199.902	< .001

Variable	Estimate	Std. Error	z-value	P(>|z|)
Distance	0.020	0.007	2.997	0.003
GPS_spd	0.997	0.005	199.905	< .001
DrivingEvents_Map_evt_ha_mean	0.999	0.005	199.907	< .001
DrivingEvents_Map_evt_hb_mean	1.000	0.005	199.906	< .001
iDreams_Headway_Map_level_0_mean	1.005	0.007	154.294	< .001
iDreams_Headway_Map_level_1_mean	1.002	0.005	187.247	< .001
iDreams_Headway_Map_level_2_mean	1.001	0.005	196.999	< .001

The results for phase 3 are shown in Figure 39 below.

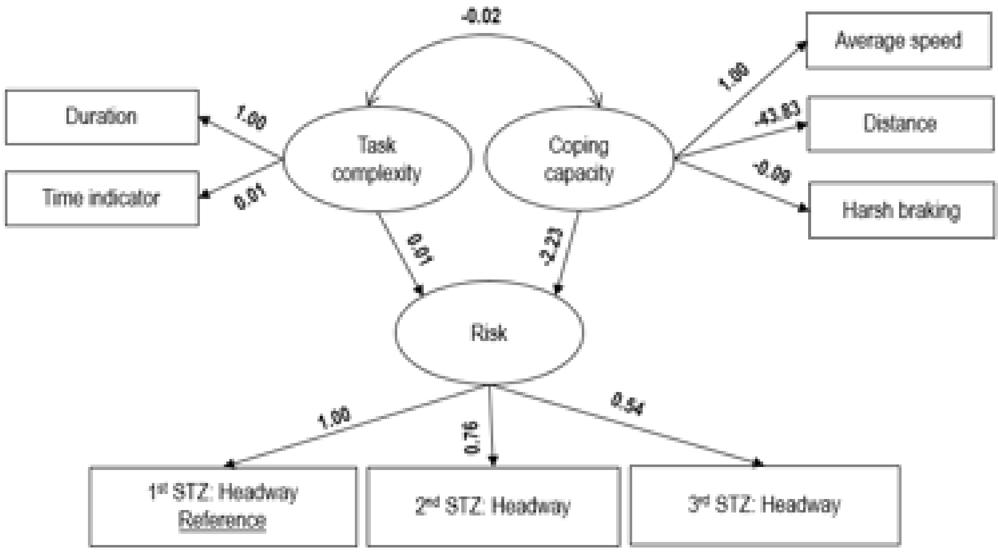

Figure 39: Results of SEM on Risk (Headway STZ) – Portuguese bus drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.998; TLI is 0.997 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.019. Table 84 summarizes the model fit of SEM applied for headway.

Table 84: Model Fit Summary for headway – Portuguese bus drivers – experiment Phase 3

Model Fit measures	Value
AIC	1.511×10+6
BIC	1.511×10+6
CFI	0.998
TLI	0.997
RMSEA	0.019
GFI	0.998
Hoelter's critical N (α = .05)	4284.444
Hoelter's critical N (α = .01)	5188.355

Residual variances details are presented in Table 85 that follows.

Table 85: Residual variances for headway – Portuguese bus drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
Duration	387.221	0.015	13.451	< .001
ME_AWS_time_indicator	1.000	0.005	199.902	< .001
Distance	1.000	0.007	1.345	< .001
GPS_spd	0.505	0.005	199.905	< .001
DrivingEvents_Map_evt_ha_mean	1.000	0.005	199.907	< .001
DrivingEvents_Map_evt_hb_mean	1.017	0.016	0.726	< .001
iDreams_Headway_Map_level_0_mean	1.010	0.005	182.677	< .001
iDreams_Headway_Map_level_1_mean	1.005	0.004	277.610	< .001
iDreams_Headway_Map_level_2_mean	387.221	0.005	-52.424	< .001

The results for phase 4 are shown in Figure 40 below.

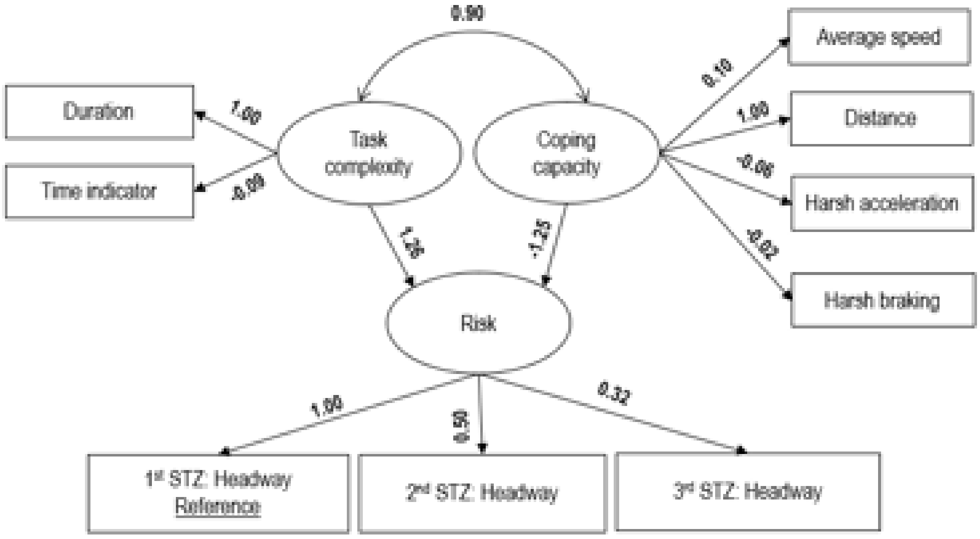

Figure 40: Results of SEM on Risk (Headway STZ) – Portuguese bus drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.964; TLI is 0.946 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.051. Table 86 summarizes the model fit of SEM applied for headway.

Table 86: Model Fit Summary for headway – Portuguese bus drivers – experiment Phase 4

Model Fit measures	Value
AIC	1.594×10+6
BIC	1.595×10+6
CFI	0.964
TLI	0.946
RMSEA	0.051
GFI	0.986
Hoelter's critical N (α = .05)	582.268
Hoelter's critical N (α = .01)	687.057

Residual variances details are presented in Table 87 that follows.

Table 87: Residual variances for headway – Portuguese bus drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
Duration	0.200	0.019	10.246	< .001
ME_AWS_time_indicator	0.994	0.005	182.697	< .001
Distance	0.012	0.016	0.726	0.468
GPS_spd	0.990	0.005	182.677	< .001
DrivingEvents_Map_evt_ha_mean	0.997	0.005	182.752	< .001
DrivingEvents_Map_evt_hb_mean	1.000	0.005	182.764	< .001
iDreams_Headway_Map_level_0_mean	1.063	0.010	106.653	< .001
iDreams_Headway_Map_level_1_mean	1.016	0.006	169.131	< .001
iDreams_Headway_Map_level_2_mean	1.006	0.006	177.870	< .001

4.2.7 Overall model (Cars)

Four separate SEM models were estimated in order to explore the relationship between the latent variables of task complexity, coping capacity and risk (expressed as the three stages of the STZ) of all event variables, such as speeding, headway, overtaking and fatigue (level 1 ‘normal driving’ used as the reference case). Data from Belgian, German and UK car drivers were analyzed. Each model corresponds with one of the phases of the i-DREAMS experiment namely:

- Phase 1: monitoring - 120 car drivers, 5,643 trips (104,195 minutes)
- Phase 2: real-time interventions - 125 car drivers, 6,188 trips (109,341 minutes)
- Phase 3: real-time & post-trip interventions - 130 car drivers, 6,519 trips (117,381 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 130 car drivers, 8,558 trips (169,695 minutes)

To begin with, the results for phase 1 are shown in Figure 41 below. Risk is measured by means of the STZ levels for speeding, headway, overtaking and fatigue (level 1 ‘normal driving’ used as the reference case; level 2 refers to ‘dangerous driving’, while no incidents with regards to level 3 ‘avoidable accident driving’ were found).

To begin with, the latent variable task complexity is measured by means of the environmental indicator of time of the day, lighting conditions and weather. Furthermore, it is shown that the latent coping capacity is measured by means of both vehicle state indicators, such as “VehicleAge” (indicating the age of the vehicle), “Gearbox” (indicating the type of gearbox; automatic or manual) and “Fuel_type” (indicating the type of fuel; diesel, hybrid electric, petrol). At the same time, operator state indicators, such as “Gender” (indicating the gender of the driver; male or female), “Age” (indicating the age of the driver), distance travelled, harsh acceleration and harsh braking are included in the SEM applied.

The structural model between the latent variables shows some interesting findings: first, task complexity and coping capacity are inter-related with a positive correlation (regression coefficient=0.02) – which increases in magnitude as the driver’s progress from phases 1 though phases 2 and 3. This positive correlation indicates that higher task complexity is associated with higher coping capacity implying that drivers coping capacity increases as the complexity of driving task increases. Overall, the structural model between task complexity and risk shows a positive coefficient, which means that increased task complexity relates to increased risk according to the model (regression coefficient=2.17). On the other hand, 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 (regression coefficient=-0.55).

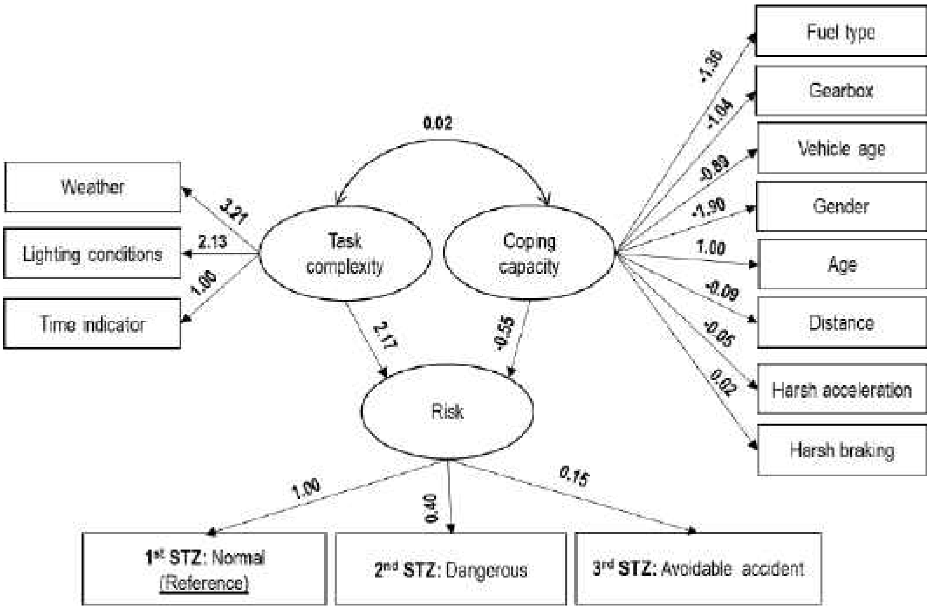

Figure 41: Results of SEM on Risk (Speeding STZ) – Belgian, German and UK car drivers – experiment Phase 1

The Comparative Fit Index (CFI) of the model is equal 0.650; TLI is 0.570 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.091. Table 88 summarizes the model fit of SEM applied for speeding.

Table 88: Model Fit Summary for speeding – Belgian, German and UK car drivers – experiment Phase 1

Model Fit measures	Value
AIC	817833.112
BIC	818194.915
CFI	0.650
TLI	0.570
RMSEA	0.091
GFI	0.918
Hoelter's critical N (α = .05)	155.529
Hoelter's critical N (α = .01)	171.977

Residual variances details are presented in Table 89 that follows.

Table 89: Residual variances for speeding – Belgian, German and UK car drivers – experiment Phase 1

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.862	0.009	100.596	< .001
ME_Car_high_beam_median	0.812	0.008	97.405	< .001
ME_Car_wipers_median	0.998	0.010	104.686	< .001
Age	0.379	0.009	41.795	< .001

Variable	Estimate	Std. Error	z-value	P(>|z|)
Fuel_type	1.000	0.013	76.545	< .001
VehicleAge	1.000	0.013	76.555	< .001
Gearbox	2.402	0.131	18.391	< .001
GPS_distances_sum	1.220	0.023	52.503	< .001
Gender	1.032	0.010	101.735	< .001
DEM_evt_ha_lvl_H_mean	0.862	0.009	100.596	< .001
STZ1	0.812	0.008	97.405	< .001
STZ2	0.998	0.010	104.686	< .001
STZ3	0.379	0.009	41.795	< .001

The following Figures show the results of the 2nd, 3rd and 4th phase of the experiment. It is observed that the measurement equations of task complexity and coping capacity are fairly consistent between the different phases. At the same time, the loadings of the observed proportions of the STZ of speeding are consistent between the different phases. The structural model between task complexity and inverse risk (normal driving) are positively correlated among the four phases, while coping capacity and risk found to have a negative relationship in all phases of the experiment. The results for phase 2 are shown in Figure 42 below.

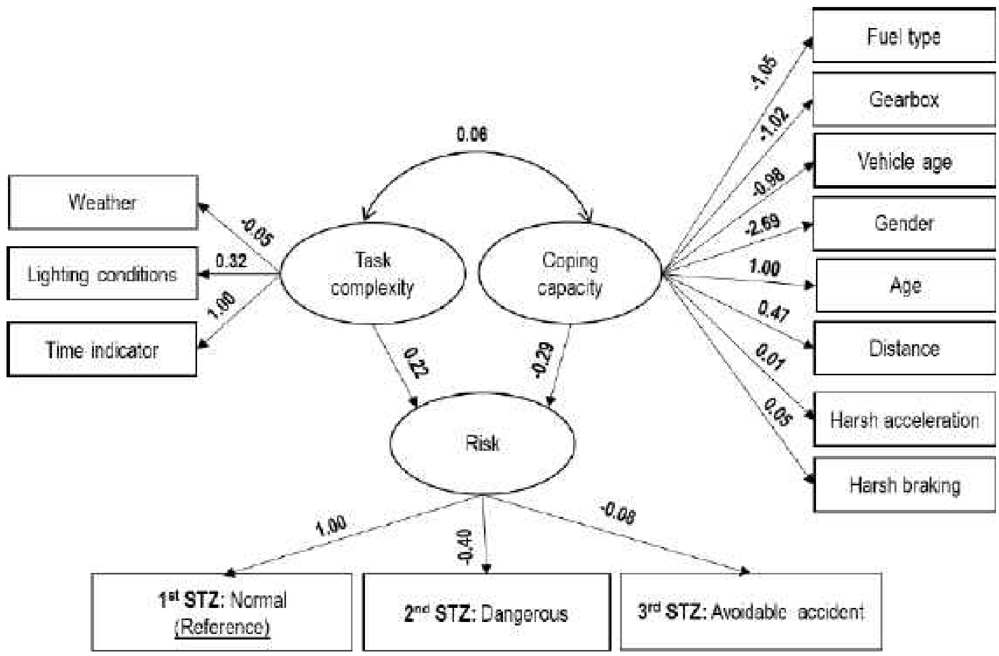

Figure 42: Results of SEM on Risk (Speeding STZ) – Belgian, German and UK car drivers – experiment Phase 2

The Comparative Fit Index (CFI) of the model is equal 0.688; TLI is 0.617 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.074. Table 90 summarizes the model fit of SEM applied for speeding.

Table 90: Model Fit Summary for speeding – Belgian, German and UK car drivers – experiment Phase 2

Model Fit measures	Value
AIC	$2.512\times10^{+6}$
BIC	$2.512\times10^{+6}$
CFI	0.688
TLI	0.617

Model Fit measures	Value
RMSEA	0.074
GFI	0.938
Hoelter's critical N (α = .05)	236.232
Hoelter's critical N (α = .01)	261.271

Residual variances details are presented in Table 91 that follows.

Table 91: Residual variances for speeding – Belgian, German and UK car drivers – experiment Phase 2

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_AWS_time_indicator_median	0.713	0.022	31.963	< .001
ME_Car_high_beam_median	0.970	0.006	171.072	< .001
ME_Car_wipers_median	0.999	0.005	187.505	< .001
Age	0.879	0.005	180.312	< .001
Fuel_type	0.867	0.005	178.940	< .001
VehicleAge	0.884	0.005	180.930	< .001
Gearbox	0.873	0.005	179.664	< .001
GPS_distances_sum	0.973	0.005	183.467	< .001
Gender	0.120	0.009	13.409	< .001
DEM_evt_ha_lvl_H_mean	1.000	0.008	123.875	< .001
DEM_evt_hb_lvl_H_mean	1.000	0.008	123.385	< .001
STZ1	-0.361	0.077	-4.690	< .001
STZ2	0.783	0.013	60.557	< .001
STZ3	0.991	0.005	187.483	< .001

The results for phase 3 are shown in Figure 43 below.

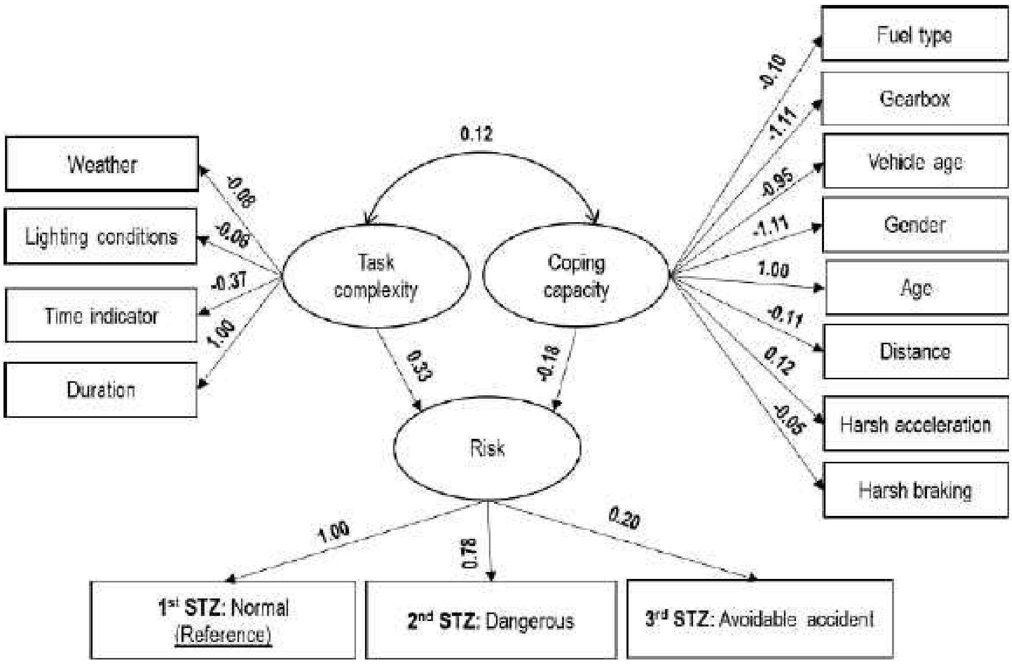

Figure 43: Results of SEM on Risk (Speeding STZ) – Belgian, German and UK car drivers – experiment Phase 3

The Comparative Fit Index (CFI) of the model is equal 0.637; TLI is 0.562 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.087. Table 92 summarizes the model fit of SEM applied for speeding.

Table 92: Model Fit Summary for speeding – Belgian, German and UK car drivers – experiment Phase 3

Model Fit measures	Value
AIC	$2.901 \times 10^{+6}$
BIC	$2.901 \times 10^{+6}$
CFI	0.637
TLI	0.562
RMSEA	0.087
GFI	0.908
Hoelter's critical N (α = .05)	166.828
Hoelter's critical N (α = .01)	183.169

Residual variances details are presented in Table 93 that follows.

Table 93: Residual variances for speeding – Belgian, German and UK car drivers – experiment Phase 3

Variable	Estimate	Std. Error	z-value	P(>|z|)
grpby_seconds	0.644	0.015	44.245	< .001
ME_AWS_time_indicator_median	0.951	0.005	179.290	< .001
ME_Car_wipers_median	0.998	0.005	193.632	< .001
ME_Car_high_beam_median	0.999	0.005	193.861	< .001
Age	0.639	0.004	153.179	< .001
Fuel_type	0.997	0.005	195.131	< .001
VehicleAge	0.674	0.004	159.380	< .001
Gearbox	0.557	0.004	135.209	< .001
GPS_distances_sum	0.996	0.005	191.177	< .001
Gender	0.554	0.004	134.476	< .001
DEM_evt_ha_lvl_H_mean	0.995	0.008	129.345	< .001
DEM_evt_hb_lvl_H_mean	0.999	0.008	129.153	< .001
STZ1	1.629	0.029	56.712	< .001
STZ2	1.386	0.018	75.676	< .001
STZ3	1.026	0.005	188.174	< .001

The results for phase 4 are shown in Figure 44 below.

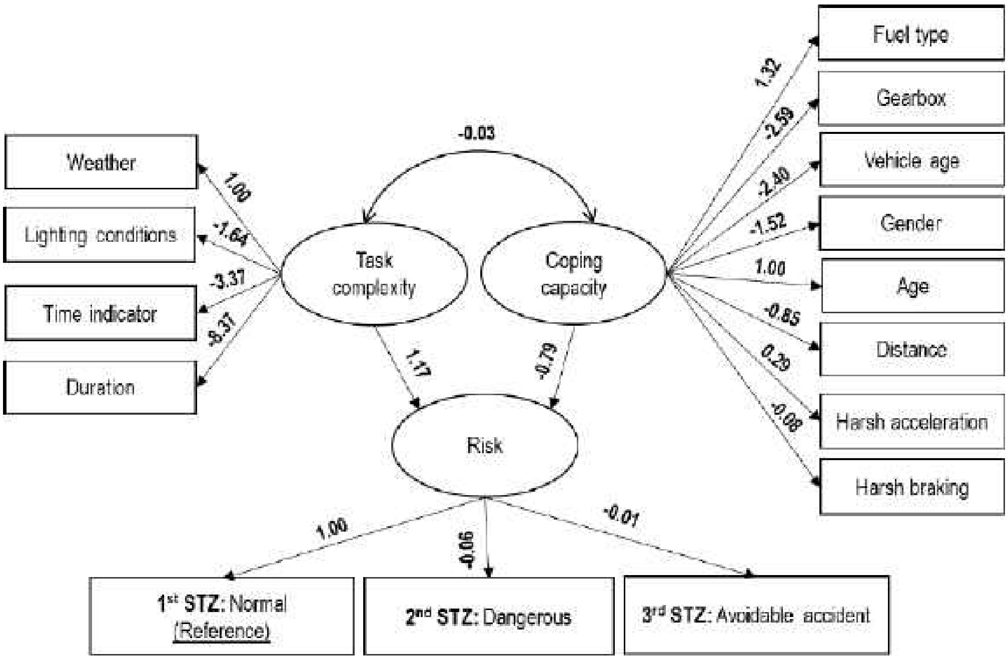

Figure 44: Results of SEM on Risk (Speeding STZ) – Belgian, German and UK car drivers – experiment Phase 4

The Comparative Fit Index (CFI) of the model is equal 0.754; TLI is 0.703 and the Root-Mean-Square-Error Approximation (RMSEA) is 0.093. Table 94 summarizes the model fit of SEM applied for speeding.

Table 94: Model Fit Summary for speeding – Belgian, German and UK car drivers – experiment Phase 4

Model Fit measures	Value
AIC	$5.729\times10^{+6}$
BIC	$5.729\times10^{+6}$
CFI	0.754
TLI	0.703
RMSEA	0.093
GFI	0.899
Hoelter's critical N (α = .05)	147.761
Hoelter's critical N (α = .01)	162.223

Residual variances details are presented in Table 95 that follows.

Table 95: Residual variances for speeding – Belgian, German and UK car drivers – experiment Phase 4

Variable	Estimate	Std. Error	z-value	P(>|z|)
ME_Car_wipers_median	0.987	0.004	279.552	< .001
ME_Car_high_beam_median	0.966	0.003	278.492	< .001
grpby_seconds	0.112	0.006	18.858	< .001
ME_AWS_time_indicator_median	0.855	0.003	265.935	< .001
Age	0.888	0.003	275.077	< .001
Fuel_type	0.806	0.003	270.034	< .001
VehicleAge	0.355	0.002	189.990	< .001
Gearbox	0.245	0.002	137.894	< .001
GPS_distances_sum	0.917	0.003	266.204	< .001

Variable	Estimate	Std. Error	z-value	P(>|z|)
Gender	0.742	0.003	265.255	< .001
DEM_evt_ha_lvl_H_mean	0.995	0.006	168.002	< .001
DEM_evt_hb_lvl_H_mean	1.000	0.006	167.719	< .001
STZ1	-7.362	0.990	-7.435	< .001
STZ2	0.974	0.005	211.023	< .001
STZ3	0.999	0.004	280.639	< .001

4.3 Summary

The analyses demonstrated that in Belgium task complexity and coping capacity were positively correlated in the majority of the models, which means that with higher task complexity comes higher coping capacity, a non-intuitive result. Task complexity was found to have greater loadings on risk, but that effect dropped when observing trips from phase 1 to phase 4 of the experiment. Furthermore, in many of the developed models the loadings revealed a spike in their values during phase 3 of the experiment and a small drop in phase 4, which points to the fact that the combination of real-time and post-trip feedback significantly influenced the relationship between task complexity, coping capacity and risk, whereas gamification in some cases might have confused drivers. It should be noted that there might be the danger for drivers to have been experienced a learning effect through the stages, as they gained experience and familiarity with the driving task. However, if the drivers did not go through the stages in the same order, it was possible that their results in later phases may have been different. Thus, the same order of phases for all drivers was used in order to draw a definitive conclusion.

In UK, loadings from the SEM models demonstrate that coping capacity and task complexity were positively correlated in phase 1 and 3, but had no significant relationship in phase 2 and phase 4. Similarly with Belgium, task complexity had a stronger impact on risk, with phase 3 showing the greatest effect on driving risk. The difference in the relationship between variables across different countries could be due to a variety of factors, such as cultural differences, economic factors, or variations in driving behaviors and infrastructure.

In Germany, the model for speeding revealed a positive correlation of task complexity and coping capacity, but with the largest correlation on phase 2 of the experiment, where real-time warnings were introduced. At the end of the experiment (phase 4), coping capacity was found to have its largest correlation with risk, while task complexity had its greatest loading during phase 3 of the experiment.

In Greece, i phase 1, task complexity and coping capacity were inter-related with a positive correlation which implies that drivers' coping capacity increases as the complexity of driving task increases. On the other hand, in phase 4, task complexity and coping capacity were negatively correlated. The effect of task complexity was generally greater than the one of coping capacity, whereas the peak of the contributions from task complexity and coping capacity was observed in phase 4.

Lastly, in Portugal, task complexity was positively associated with the latent variable risk, which was defined by different levels of headway. The higher the complexity, the higher the chance to drive normally and more carefully. On the other hand, coping capacity was negatively associated with risk (or normal driving) which implied that higher coping capacity might encourage normal driving and reduce risk. Task complexity and coping capacity were inter-related with a positive correlation – which reduced in magnitude as the driver's progress from phase 1 though phase 4. Similar patterns of professional drivers (in terms of loadings and signs among phases for Belgian truck and Portuguese bus drivers) were observed.

Looking at the observed risk factors, it was demonstrated that for the speeding and the headway models, the correlation of task complexity and coping capacity was positive, with the main exceptions being observed in phases 2 and 3 in Greece, Germany and Belgium. For harsh accelerations in Belgian trucks, the correlation of coping capacity and task complexity was in general positive along the same magnitude for all phases.

According to the overall model applied for cars, the latent variable risk was measured by means of the STZ levels for speeding, headway, overtaking and fatigue. The positive correlation of task complexity and coping capacity implied that drivers' coping capacity increased as the complexity of driving task increases. This finding may be a sign of risk compensating behavior of drivers when the complexity of driving task is high, and is in line with the theoretical model of i-DREAMS, validating the assumption that risk (or its' inverse, the normal driving) is an outcome of the interaction between the two variables in addition to their separate effect. A positive correlation of risk with the STZ indicators was identified in phase 1, while a negative correlation was found in phase 4 which showed that the latent variable risk could in fact be representing an inverse of risk, more like a normal driving.

With regards to the overall model, results showed that higher task complexity levels lead to higher coping capacity. This means that drivers, when faced with difficult conditions, tend to regulate well their capacity to apprehend potential difficulties, while driving. It was revealed that the SEM applied between task complexity and inverse risk were positively correlated in all phases of the experiment, which means that increased task complexity relates to increased risk. On the other hand, coping capacity and inverse risk found to have a negative relationship in all phases, which means that increased coping capacity relates to decreased risk. Overall, **the interventions had a positive influence on risk**, increasing the coping capacity of the operators and reducing the risk of dangerous driving behavior.

5 Other analyses

5.1 Real-time

5.1.1 Neural Networks

In order to investigate if real-time prediction of the STZ is also feasible, two feed-forward multi-layer perceptrons were also applied on a subset from the total dataset of the UK car drivers ($N_{drivers}$=30, trips=5340). In order to identify the effect of phase on the prediction, the analysis considered phase as an independent variable and the analysis was performed for the whole dataset, rather than per phase as the analyses in Chapter 4. The algorithms, had an accuracy of more than 94% with a false alarm rate of up to only 6%. The Neural Networks (NNs) classification algorithms acted as preparatory step towards the LSTM classification that is shown in the next subsection. The predictors utilized for the models are shown in Table 96.

Table 96: Predictors utilized for Neural Networks

Variables	Headway	Speeding
Phase	x	x
SQ_Age	x	x
ME_Car_speed_mean	x	x
DEM_evt_ha_lvl_L_mean	x	x
DrivingEvents_Map_lvl_L_mean	x	
iDreams_Headway_Map_level_total_mean	x	
iDreams_Speeding_Map_level_0_mean		x
iDreams_Speeding_Map_level_total_mean		x

After the application of the models, the identified confusion matrix was produced for the two independent variables (i.e. headway and speeding), as shown in Table 97.

Table 97: Confusion data matrix for headway and speeding

Variable	TP	FP	FN	TN	Sum
Headway	33378	0	1400	82	34860
Speeding	2178	1987	63	30632	34860

From the confusion matrix, the following metrics were estimated and are depicted in Table 98.

Table 98: Assessment of classification model for headway and speeding

Variable	Accuracy	Precision	Recall	f1-score	G-Means	FA Rate
Headway	95.98%	100.00%	95.97%	97.95%	97.97%	0.00%
Speeding	94.12%	52.29%	97.19%	68.00%	71.29%	6.09%

Figure 45 illustrates the performance of Neural Network classification on headway and speeding STZ level.

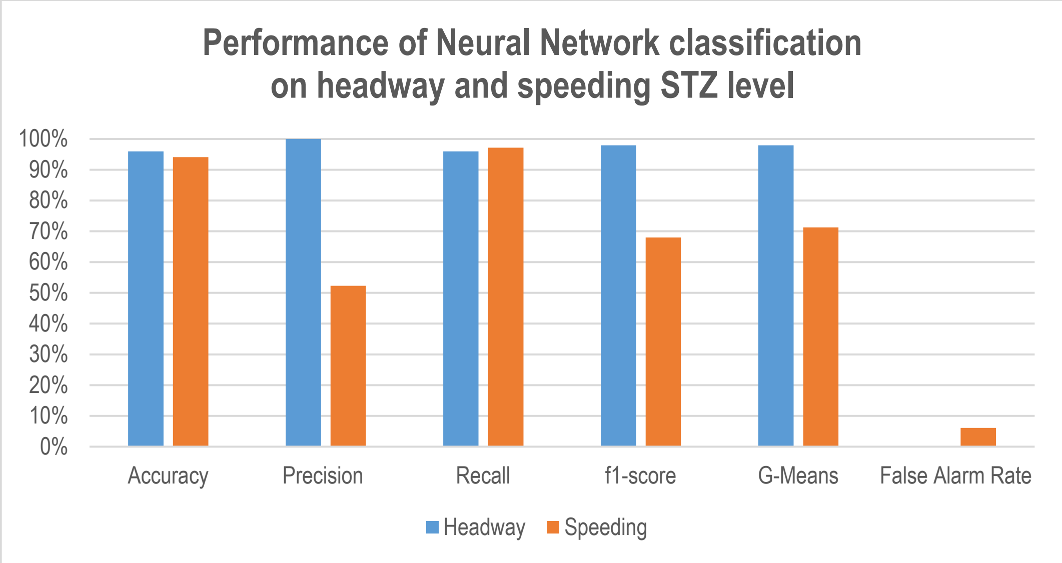

Figure 45: Performance of Neural Network classification for headway and speeding

The results shown in Figure 45 are in line with relevant literature on real-time safety evaluations (Silva et al., 2020), as well as previous project analyses utilized on simulator data (Garefalakis et al., 2022). Precision, f1-score and G-means metrics are probably lower due to the greater amount of 'normal' STZ level instances as compared with 'dangerous' conditions.

5.1.2 Long Short-Term Memory Networks

5.1.2.1 Speeding

Following the development of simple NN classifiers, Long Short-Term Memory Networks (LSTMs) were trained in order to predict 'dangerous' speeding level. As shown in Table 99, the speeding LSTM did not achieve significant results, only reaching 57.82% accuracy after the developed trials. Although LSTM is often used for sequence modeling, it is worth mentioning that the sequence may not always be explicitly visible in the predictors themselves. In some cases, the sequence may be implicit in the way that the data is organized or structured. For example, in time series data, the sequence is often defined by the order in which the data was collected over time. In this case, the LSTM is used to model and make predictions based on the temporal dependencies and patterns in the data. In other cases, the sequence may be less obviously related to time, but still exist in the way that the data is organized. For example, in natural language processing, the sequence may be defined by the order of words in a sentence or text document. Thus, the sequence is implicit in the way that the data was collected or organized, even if it's not immediately apparent from the predictors themselves. An LSTM could still be used in this case to model and make predictions based on the implicit sequence in the data. The predictors utilized for the models applied for speeding are shown in Table 99.

Table 99: Predictors utilized for Long Short-Term Memory Networks for speeding

Variables	v1	v2	v3	v4	v5
Phase	x	x	x	x	x
SQ_Age	x	x	x	x	x
GPS_spd_mean	x			x	x
ME_Car_speed_mean	x	x		x	x
Driving Events_Map_evt_ha_mean	x	x	x		x
DEM_evt_ha_lvl_L_mean	x	x	x		
iDreams_Speeding_Map_level_0_mean	x	x	x		
iDreams_Speeding_Map_level_1_mean					x
iDreams_Speeding_Map_level_total_mean	x	x	x	x	x
iDreams_Headway_Map_level_total_mean				x	
Accuracy (%)	**57.82**	**57.82**	**57.82**	**57.11**	**57.82**

5.1.2.2 Headway

Similarly with speeding, LSTMs could not find the dangerous level of headway as well. Perhaps this is because of a lack of data or speed-related indicators to identify the different levels. The predictors utilized for the models applied for headway are shown in Table 100.

Table 100: Predictors utilized for Long Short-Term Memory Networks for headway

Variables	v1	v2	v3	v4
Phase	x	X	x	x
SQ_Age	x	X	x	x
GPS_spd_mean	x		x	
ME_Car_speed_mean		X	x	x
DrivingEvents_Map_lvl_H_mean	x		x	
DrivingEvents_Map_lvl_L_mean		X		
DEM_evt_ha_lvl_L_mean		X		
iDreams_Headway_Map_level_-1_mean			x	
iDreams_Headway_Map_level_0_mean	x			
iDreams_Headway_Map_level_total_mean	x	X	x	x
iDreams_Speeding_Map_level_total_mean				x
Accuracy (%)	**57.39**	**55.5**	**57.82**	**57.82**

It should be noted that an accuracy of less than 60% may not be sufficient for a high-performance intervention system, as it could result in a relatively high number of false alarms or missed detections. However, the required level of accuracy depends on the specific use case and the risks involved. For instance, in a system designed to detect potential crashes or safety hazards, a higher level of accuracy may be necessary in order to ensure the safety of drivers and other road users. As for the use of prediction models by an intervention system, the output of the models can be used in a variety of ways. In particular, the prediction models can generate real-time alerts or warnings to drivers or other stakeholders, such as traffic control centers or emergency responders. The models can also be used to trigger automated interventions, such as adjusting the speed of a vehicle or activating safety features like automatic braking systems. In addition, the output of prediction models can be used for ongoing analysis and monitoring of road safety performance, in order to identify trends and patterns that can inform future interventions and improvements.

5.2 Post-trip

5.2.1 Grouped Random Parameters Binary Logit Model

5.2.1.1 Near-Misses

Occurrence of a near-miss is one of the indicators of risk that can be used in i-DREAMS. A near-miss can be defined in various ways (Papazikou et al., 2019). For the purpose of analysis in i-DREAMS, near-miss is defined as a binary indicator in which (in any 60 seconds interval) at least two harsh events occur including harsh acceleration, harsh braking, or harsh cornering.

Four separate models were fitted for the occurrence of near-misses, one for each phase of the experiment. The sample used was that of Belgian car drivers. More specifically:

- Phase 1: 29 drivers, 456 trips (3,735 minutes) – 19 observed near-misses
- Phase 2: 35 drivers, 462 trips (3,673 minutes) – 23 near-misses
- Phase 3: 36 drivers, 424 trips (3,584 minutes) – 24 near-misses
- Phase 4: 30 drivers, 436 trips (3,659 minutes) – 26 near-misses

The results of each phase are shown in the following Table 101, Table 102, Table 103 and Table 104. In order to show the evolution of models over the four phases, shaded cells in each phase are the differences of the parameter estimates (either one has become significant, or one has become insignificant) with the previous phase.

Table 101: GRPL model for near-misses – Belgian car drivers – Experiment phase 1

Variable	Estimate	Std.Error	z value	p value
Years of holding a driving licence	-0.280	0.023	-12.423	0.000
Night-time driving	-1.847	0.742	-2.488	0.013
Long headway (more than 2 seconds)	-6.208	1.979	-3.137	0.002
Inter-Beat Interval (IBI) [mean]	0.418	0.218	1.917	0.055
Inter-Beat Interval (IBI) [SD]	-0.001	0.336	-0.003	0.997
Driving faster than the speed limit in urban areas over the last year [mean]	-4.278	1.229	-3.480	0.001
Driving faster than the speed limit in urban areas over the last year [SD]	1.697	0.867	1.958	0.050
Violation: used a hand-held mobile phone while driving over the last year [mean]	0.253	0.159	1.590	0.112
Violation: used a hand-held mobile phone while driving over the last year [SD]	-1.915	0.266	-7.190	0.000

Variable	Estimate	Std.Error	z value	p value
Fixed effect (trip)	-0.930	4.412	-0.211	0.833
Goodness of fit				
LL0	-468.5			
LL	-129.5			
P	10			
N	3735			
AIC	279.0			
BIC	294.7			

The fitted model of experiment phase 1 performs well as indicated by the large reduction of the LL vs. the initial LL0. A number of explanatory variables are significant in terms of their impact on the probability of near-miss, including driver related variables (including physiological and self-reported behavioral indicators), trip related variables but also observation / minute level variables:

- The years of holding a driving license are associated with reduced probability of near miss, which is intuitive as more experienced drivers are known to have lower involvement in critical events.
- A higher mean IBI is associated with higher probability of near miss; although this effect is marginally significant at 95%, it may reflect the effects of driver sleepiness or fatigue, which is known to be associated with higher IBI (slower heart rate). The variance of this random parameter is not statistically significant indicating that the effect of IBI on near-misses is almost the same for all drivers.
- Driving faster than the speed limit in urban areas over the last year is associated with a lower probability of near-miss over this sample of drivers; this effect has a significant random variation, as indicated by its SD estimate, however the overall distribution lies within the positive scale. It can be assumed that more frequent speed limit exceedance in urban areas may be associated with lower traffic volumes, which leads to fewer interactions with other vehicles and other road users. This may also due to the fact that drivers actually had a real crash rather than a near-miss.
- The mean parameter estimate of using a mobile phone while driving over the last year is not statistically significant, however its standard deviation over this sample significant and negative. This indicates that for half of the trips mobile phone use is associated with higher probability of near-miss, while the opposite is the case for the other half. This may reflect the known mechanism of mobile phone use which may on the one hand lead to impaired reaction time and crash probability, but on the other hand indicate a successful compensatory behavior of drivers, who may reduce their speed and increase their headway while using the mobile phone.
- Night time driving is associated with lower probability of near miss, possibly due to lower traffic volumes at night.
- Finally, headways longer than 2 seconds over the 60-second intervals are found to reduce the probability of near-misses, which is intuitive.

- The fixed variable “trip” is associated with lower probability of near-miss, indicating that drivers who performed more trips during the experiment were less likely to have a near-miss.

Table 102: GRPL model for near-misses – Belgian car drivers – Experiment phase 2

Variable	Estimate	Std.Error	z value	p value
Gender: male	-1.468	0.322	-4.558	0.000
Years of holding a driving licence	-0.065	0.012	-5.337	0.000
Night-time driving	-1.023	0.600	-1.706	0.088
Long headway (more than 2 seconds)	-2.751	0.934	-2.946	0.003
Driving more than 10 km/h over the limit	2.184	0.359	6.083	0.000
Inter-Beat Interval (IBI) [mean]	0.814	0.210	3.871	0.000
Inter-Beat Interval (IBI) [SD]	0.001	0.387	0.001	0.999
Driving faster than the speed limit in urban areas over the last year [mean]	-0.589	0.216	-2.730	0.006
Driving faster than the speed limit in urban areas over the last year [SD]	0.000	0.250	0.000	1.000
Violation: used a hand-held mobile phone while driving over the last year [mean]	-0.579	0.185	-3.133	0.002
Violation: used a hand-held mobile phone while driving over the last year [SD]	0.000	0.360	-0.001	0.999
Fixed effect (trip)				
Goodness of fit				
LL0	-350.6			
LL	-148.9			
P	12			
N	3673			
AIC	321.8			
BIC	340.6			

The fitted model for the 2nd phase of the experiment (receiving real-time feedback through the i-DREAMS gateway) reveals the following differentiations compared to the 1st phase (no intervention):

- The male gender is associated with lower probability of near miss in this phase. It should be noted that males are over-represented in the sample, so this effect may express the general effectiveness of the real-time intervention.
- A higher mean IBI is now clearly associated with higher probability of near miss, suggesting that the real-time interventions may not fully address the sleepiness/fatigue risk, or that there are other human factors related to IBI which are not explicitly identified.

- Driving faster than the speed limit in urban areas over the last year is associated with a lower probability of near-miss over this sample of drivers, and its random variation is now non-significant.
- At the same time, however, the real-time measurement of exceeding the speed limit more than 10 Km/h over one minute is associated with higher probability of near-miss, suggesting that the interventions seems less effective at higher stages of the STZ.
- The mean parameter estimate of Using a mobile phone while driving over the last year is now statistically significant and reducing the probability of near-miss. It appears that the effect of compensatory behavior of mobile phone use while driving is stronger when drivers receive real-time interventions, although these are not targeting mobile phone use in real time.

It is noted that the effect of headways is still significant in this model, but with a smaller magnitude, suggesting the effectiveness of the headway warnings received by drivers.

Table 103: GRPL model for near-misses – Belgian car drivers – Experiment phase 3

Variable	Estimate	Std.Error	z value	p value
Gender: male	-1.674	0.356	-4.707	0.000
Years of holding a driving licence	-0.080	0.012	-6.783	0.000
Night-time driving	-0.930	0.762	-1.220	0.222
Long headway (more than 2 seconds)	-2.247	0.929	-2.418	0.016
Driving more than 10 km/h over the limit	1.952	0.340	5.733	0.000
Inter-Beat Interval (IBI) [mean]	0.784	0.188	4.178	0.000
Inter-Beat Interval (IBI) [SD]	-0.245	0.201	-1.219	0.223
Driving faster than the speed limit in urban areas over the last year [mean]	-0.653	0.250	-2.614	0.009
Driving faster than the speed limit in urban areas over the last year [SD]	-0.006	0.318	-0.020	0.984
Violation: used a hand-held mobile phone while driving over the last year [mean]	-0.588	0.191	-3.073	0.002
Violation: used a hand-held mobile phone while driving over the last year [SD]	0.000	0.258	0.001	1.000
Fixed effect (trip)	1.093	0.773	1.414	0.157

Goodness of fit	
LL0	-364.7
LL	-157.1
P	12
N	3584

Variable	Estimate	Std.Error	z value	p value
AIC	338.2			
BIC	356.8			

The model fitted for the 3rd phase of the experiment, in which post-trip feedback is added, presents one main differentiation: the impact of night time driving becomes non significant. The effect of long headways is statistically significant again, together with the effect of driving more than 10Km/h over the speed limit, without notable differentiation in their magnitude. It can be assumed that the effect of post-trip feedback together with the real-time interventions does not show strikingly different results from the previous phase.

Table 104: GRPL model for near-misses – Belgian car drivers – Experiment phase 4

Variable	Estimate	Std.Error	z value	p value
Gender: male	-3.270	0.765	-4.275	0.000
Years of holding a driving licence	-0.072	0.019	-3.821	0.000
Night-time driving	-	-	-	-
Long headway (more than 2 seconds)	-3.677	1.276	-2.883	0.004
Driving more than 10 km/h over the limit	1.319	0.533	2.477	0.013
Inter-Beat Interval (IBI) [mean]	0.540	0.225	2.395	0.017
Inter-Beat Interval (IBI) [SD]	-0.482	0.492	-0.979	0.327
Driving faster than the speed limit in urban areas over the last year [mean]	-2.926	1.324	-2.210	0.027
Driving faster than the speed limit in urban areas over the last year [SD]	1.500	0.632	2.373	0.018
Violation: used a hand-held mobile phone while driving over the last year [mean]	0.144	0.245	0.587	0.557
Violation: used a hand-held mobile phone while driving over the last year [SD]	1.013	0.394	2.569	0.010
Fixed effect (trip)	-18.377	6.176	-2.976	0.003
Goodness of fit				
LL0	-387.3			
LL	-137.4			
P	12			
N	3659			
AIC	298.8			
BIC	317.6			

The results of the respective model of the 4th phase of the project, in which real-time and post-trip feedback is combined with gamification show no substantial difference from the previous two phases. Only minor differentiations occur, e.g. the variance of the impact of self-reported use of mobile phone over the last year becomes significant, as in the 1st phase model.

It appears that lighter traffic conditions, as those of night-time driving and speeding in urban areas reduce the probability of near-miss at the driver habits level. On the other hand, at the microscopic level, drivers exceeding the speed limit by more than 10Km/h within 60 seconds of a trip are more likely to have a near-miss while doing so, while on the contrary drivers keeping longer headways within 60 seconds of trip are less likely to have a near-miss. Physiological indicators that are known to be associated with sleepiness increase the probability of near-miss. There is heterogeneity among drivers as regards the impacts of self-reported behaviors and habits, as well as the impact of physiological indicators, without however changes on the sign of these impacts (positive or negative) in each case.

Overall, it can be concluded that near-misses appear to be random events whose explanatory factors do not differentiate between different phases of the experiment. This finding can be attributed to the adopted definition of near-misses, i.e. the occurrence of two harsh events in 60 seconds of driving, which may not be representative of all types of near-misses. It is noted, however, that these explanatory factors concern the near-misses which did occur, and do not directly relate to the effectiveness of the applied interventions as per this type of outcome nor do they relate to actual crashes that may have occurred. In fact, the number of this type of near-misses remains fairly constant along the four phases of the experiment, and it is proportional to the number of drivers / minutes of trips recorded in each phase.

5.2.2 Ordered Probit Fractional Split Model

5.2.2.1 Speeding

Speeding (i.e. driving over the speed limit) is one of the indicators of risk that is used in i-DREAMS. According to this definition, this indicator of risk is discrete and is ordered. The STZ for speeding has been defined in i-DREAMS as:

- STZ 1: driving less than 5 km/h over the speed limit
- STZ 2: driving between 5 km/h and 10 km/h over the speed limit and
- STZ 3: driving more than 10 km/h over the speed limit

Four separate models were fitted for the propensity of speeding as per the above model (Ordered Probit Fractional Split – OPFS), one for each phase of the experiment. More specifically:

- Phase 1: monitoring - 39 drivers, 1173 trips (23,725 minutes)
- Phase 2: real-time interventions - 43 Belgian car drivers, 1549 trips
- Phase 3: real-time & post-trip interventions - 51 Belgian car drivers, 1973 trips (40,121 minutes)
- Phase 4: real-time, post-trip interventions & gamification - 49 Belgian car drivers, 2468 trips (52,077 minutes)

The results of each phase are shown in the following Table 105, Table 106, Table 107 and Table 108. In order to show the evolution of models over the four phases, shaded cells in each phase are the differences of the parameter estimates (either one has become significant, or one has become insignificant) with the previous phase.

Table 105: OPFS model for speeding – Belgian car drivers – Experiment phase 1

Variable	Estimate	Std.Error	z value	p value
Age	-0.193	0.027	-7.012	0.000
Gender: male	0.059	0.024	2.449	0.014
Education: college or above	-0.124	0.033	-3.744	0.000
Driving style: sportive	0.201	0.023	8.708	0.000
Belief about driving	-0.058	0.011	-5.164	0.000
Perceived competence: general driving skills than the average driver	0.067	0.012	5.740	0.000
Familiarity with the benefits of safe driving	-0.032	0.012	-2.640	0.008
Violation: driving faster than the speed limit over the last year	0.055	0.012	4.642	0.000
Average weekly km driven on rural roads	0.150	0.026	5.731	0.000
Night-time driving	0.201	0.026	7.797	0.000
Fixed effect 1: driver	0.004	0.001	3.875	0.000
Fixed effect 2: trip	-0.233	0.443	-0.527	0.598
Threshold 1: STZ1 to STZ2	1.090	0.062	17.581	0.000
Threshold 2: STZ2 to STZ3	1.292	0.062	20.770	0.000
Goodness of fit				
LL0	-11441.6			
LL	-11224.3			
P	14.0			
N	23725.0			
AIC	22476.6			
BIC	22509.9			

A number of explanatory variables are significant in terms of their impact on the propensity of exceeding the STZ levels of speeding, including driver related variables (including demographic and self-reported behavioral indicators), trip related variables and environment related variables:

- Older drivers and drivers with college education or higher are less likely to commit higher levels of the STZ for speeding over a minute, possibly due to more conservative driving
- Male drivers, as well as drivers with sportive driving style, driving faster than the speed limit over the last year and higher perceived competence compared to the average driver are more likely to exhibit higher levels of the STZ. All these variables reflect the

confident and more aggressive behaviors that are known to be associated with violations.

- Drivers who think driving is very dangerous and those who are familiar with the benefits of safe driving have lower propensity of exceeding the normal STZ of speeding.
- A higher exposure per week on rural roads is associated with higher propensity of speeding, possibly because rural roads have lower traffic and moderate speed limits, leading familiar drivers to tend to exceed speed limits in higher levels.
- Night time driving also leads to higher propensity of speeding, possibly due to lower traffic during these hours.

During the 2nd phase of the experiment (see Table 106), the results remain practically the same. The only difference is that the effect of exposure on rural roads becomes non significant. This might be an indication that the real-time intervention has counterbalanced the effect of experience and exposure.

Table 106: OPFS model for speeding – Belgian car drivers – Experiment phase 2

Variable	Estimate	Std.Error	z value	p value
Age	-0.047	0.012	-3.900	0.000
Gender: male	0.045	0.023	1.943	0.052
Education: college or above	-0.098	0.025	-3.861	0.000
Driving style: sportive	0.031	0.022	1.415	0.157
Belief about driving	-0.070	0.010	-6.964	0.000
Perceived competence: general driving skills than the average driver	0.047	0.010	4.723	0.000
Familiarity with the benefits of safe driving	-0.010	0.011	-0.982	0.326
Violation: driving faster than the speed limit over the last year	0.110	0.011	9.766	0.000
Average weekly km driven on rural roads	-0.007	0.012	-0.555	0.579
Night-time driving	0.161	0.028	5.770	0.000
Fixed effect 1: driver	0.000	0.001	0.392	0.695
Fixed effect 2: trip	-0.747	0.379	-1.971	0.049
Threshold 1: STZ1 to STZ2	0.925	0.053	17.590	0.000
Threshold 2: STZ2 to STZ3	1.081	0.053	20.533	0.000

Goodness of fit	
LL0	-13715.9
LL	-13572.3
P	14
N	30188
AIC	27172.6

Variable	Estimate	Std.Error	z value	p value
BIC	27207.3			

In the model of the 3rd phase of the experiment (real-time interventions and post-trip feedback through the mobile phone app – see Table 107) there is an interesting finding; the demographic variables of age and education level also become non significant. This suggests that personality traits (beliefs, attitudes, aggressive driving and violations) remain explanatory factors of the propensity of speeding despite the presence of the intervention scheme. In this model, night time driving is also not associated with exceeding safe STZ boundaries.

Table 107: OPFS model for speeding – Belgian car drivers – Experiment phase 3

Variable	Estimate	Std.Error	z value	p value
Age	-0.002	0.010	-0.164	0.870
Gender: male	0.052	0.021	2.464	0.014
Education: college or above	0.009	0.020	0.475	0.635
Driving style: sportive	0.165	0.018	9.203	0.000
Belief about driving	-0.059	0.008	-7.042	0.000
Perceived competence: general driving skills than the average driver	0.048	0.008	5.739	0.000
Familiarity with the benefits of safe driving	-0.034	0.008	-4.110	0.000
Violation: driving faster than the speed limit over the last year	0.029	0.010	3.029	0.002
Average weekly km driven on rural roads	-0.022	0.011	-2.069	0.039
Night-time driving	-	-	-	-
Fixed effect 1: driver	0.001	0.001	1.501	0.133
Fixed effect 2: trip	-0.380	0.378	-1.004	0.316
Threshold 1: STZ1 to STZ2	1.054	0.045	23.403	0.000
Threshold 2: STZ2 to STZ3	1.185	0.045	26.260	0.000
Goodness of fit				
LL0	-18417.7			
LL	-18277.8			
P	13			
N	40121			
AIC	36581.6			
BIC	36615.4			

Finally, in the 4th phase of the experiment (where gamification is introduced – see Table 108), there is only slight variation of the previous findings. Age is significant again, college education is marginally significant, and familiarity with the benefits of safe driving becomes non significant.

Table 108: OPFS model for speeding – Belgian car drivers – Experiment phase 4

Variable	Estimate	Std.Error	z value	p value
Age	-0.042	0.010	-4.092	0.000
Gender: male	0.068	0.018	3.832	0.000
Education: college or above	0.032	0.018	1.798	0.072
Driving style: sportive	0.041	0.016	2.560	0.010
Belief about driving	-0.046	0.007	-6.237	0.000
Perceived competence: general driving skills than the average driver	0.036	0.008	4.608	0.000
Familiarity with the benefits of safe driving	-0.002	0.008	-0.207	0.836
Violation: driving faster than the speed limit over the last year	0.045	0.008	5.398	0.000
Average weekly km driven on rural roads	0.021	0.010	2.087	0.037
Night-time driving	-	-	-	-
Fixed effect 1: driver	0.000	0.000	-0.894	0.371
Fixed effect 2: trip	-0.310	0.378	-0.821	0.411
Threshold 1: STZ1 to STZ2	1.096	0.039	28.080	0.000
Threshold 2: STZ2 to STZ3	1.231	0.039	31.449	0.000
Goodness of fit				
LL0	-22399.4			
LL	-22308.0			
P	12			
N	52077			
AIC	44640.0			
BIC	44672.6			

Although there is indication that the introduction of interventions reduces the role of the environmental variable (night-time) and the drivers' general characteristics, and strengthens the role of personality characteristics, the small samples do not allow for a final conclusion. It is possible that these fluctuations are due to the differences in sample sizes and other unobserved factors.

It is indicated however that this type of model is appropriate for monitoring the proportions the STZ levels at a relatively small time-scale, which gives a more accurate representation of speeding behavior that what would have been obtained by a discrete ordinal model.

6 Conclusions

This deliverable aimed at developing an integrated model of driver-vehicle-environment interaction and risk by:

(i) identifying the most critical precursors of risk from both the task complexity and the coping capacity side,
(ii) implement an integrated model for understanding the effect of the inter-relationship of task complexity and coping capacity with risk, and
(iii) compare the performance of such models on different countries.

The ultimate goal of the analyses in this project was to identify the impact that the balance between task complexity and coping capacity has on the risk of a crash. For that reason, a vast library of data from naturalistic driving experiments was created in five countries (i.e. Belgium, UK, Germany, Greece and Portugal) to investigate the most prominent driving behavior indicators available, including speeding, headway, overtaking, duration, distance and harsh events (i.e. harsh acceleration and harsh braking).

The analysis team answered several of the research questions dealing with the comprehension of the relationship between task complexity and coping capacity, which can be summarized in the following conclusions:

- For the majority of the risk factors investigated it was found that higher task complexity levels lead to higher coping capacity by the vehicle operators. This means that drivers when faced with difficult conditions tend to regulate well their capacity to react to potential difficulties, while driving.
- When looking into the relationship between the interaction of task complexity and coping capacity and its effect on risk, in Belgium, Greece and Germany, it was shown that the influence of task complexity on risk was greater than the effect of coping capacity. Mixed results were observed in the UK and Portugal.
- The comparison of models fitted on data from the different phases of the experiments, validated that in the majority of the countries the interventions had a positive influence on risk compensation, increasing the coping capacity of the operators and reducing the risk of dangerous driving behavior.
- Predictive real-time analyses demonstrated that it is possible to predict the level of STZ with an accuracy of up to 95%, while post-trip explanatory studies showcased the capacity of state-of-the-art econometric models to shed light on the complex relationship of risk with the interdependence of task complexity and coping capacity.

An overview of the effects found for task complexity and coping capacity on risk among all available data can be found in Table 109 below. A positive sign means a positive correlation of task complexity or coping capacity with risk while a negative sing indicates a negative relationship between task complexity or coping capacity and risk.

Table 109: Effect of task complexity and coping capacity on risk per indicator/ phase/country/transport mode

Country (transport mode)	Indicator	Phase 1		Phase 2		Phase 3		Phase 4	
		TC	CC	TC	CC	TC	CC	TC	CC
Belgium (cars)	speeding	-	+	-	+	-	+	+	+
	headway	-	+	-	+	-	-	-	+
Belgium (trucks)	speeding	-	-	-	-	-	-	-	-
	harsh acceleration	+	-	+	-	+	-	+	-
	headway	-	-	-	-	-	-	+	-
UK (cars)	headway	-	-	+	-	-	-	-	-
Germany (cars)	speeding	+	-	+	-	+	-	+	+
Greece (cars)	speeding	+	-			+	-	+	-
Portugal (buses)	headway	+	-	-	-	+	-	+	-
Overall (cars)	speeding, headway, overtaking, fatigue	+	-	+	-	+	-	+	-

**TC refers to Task Complexity and CC refers to Coping Capacity*

From Table 109, it can be concluded that in the majority of the models the intuitive effect that task complexity along with coping capacity has on risk has been validated. It is demonstrated that in most of the models, increased task complexity decreases risk, while increased coping capacity decreases risk. The majority of 'inconsistent' effects are observed in phase 2 and phase 4 of the experiments, probably due to some of the drivers being affected by the introduction of real-time warnings or the gamification features of the i-DREAMS app.

It is worth noting that the relationship between task complexity and risk, as well as coping capacity and risk, may depend on the specific context and the type of task or activity involved. In general, higher task complexity may increase the potential for errors or crahses, as it can lead to greater cognitive or physical demands on the individual performing the task. However, it is also possible that increased experience or training can help to mitigate the risk associated with higher task complexity. Similarly, a higher coping capacity may help to reduce the risk of crahses or errors, as it can provide individuals with the resources or strategies needed to effectively manage challenging or stressful situations. However, the effectiveness of coping strategies may depend on the specific context and the individual's ability to apply them in real-world situations. Overall, it is important to consider the specific factors and context involved when assessing the relationship between task complexity, coping capacity, and risk.

The developed models presented in this deliverable can be further exploited by researchers and practitioners. Additional task complexity and coping capacity factors, such as road type, more personality traits and driving profiles could be utilized for example. Furthermore, data could be enhanced by including additional measurements such as electrocardiogram and electroengephalogram readings, traffic conflicts and transport emissions . Finally, additional methodologies such as imbalanced learning and models taking into account unobserved heterogeneity could be explored for the understanding of the relationship between task complexity, coping capacity and crash risk.

On the basis of the i-DREAMS results, a set of policy recommendations at different levels (EU, national and local authorities, industry, etc.) can be provided. Some potential policy recommendations based on the i-DREAMS results include:

- Developing and promoting standardized **guidelines and best practices** for driver monitoring and assistance systems, in order to improve their effectiveness and ensure consistency across different vehicles and manufacturers.
- Encouraging **investment in research** and development of new technologies and approaches for driver monitoring and assistance, in order to further improve their accuracy and effectiveness.
- Establishing clear **regulations and standards** for driver monitoring and assistance systems, in order to ensure their safety and effectiveness in real-world situations.
- Providing **education and training programs** for drivers and other road users, in order to increase awareness of the benefits and limitations of driver assistance systems, and to promote safe and responsible use of these technologies.
- Developing and implementing **targeted interventions and policies** to address specific road safety challenges identified by the i-DREAMS study, such as distraction, fatigue, and impaired driving.

Overall, the i-DREAMS study can provide valuable guidance and evidence-based recommendations for policymakers and stakeholders working to improve road safety and promote the widespread adoption of effective driver assistance and monitoring systems.

The i-DREAMS system itself can directly improve safety once launched, but also additional safety benefits can be envisaged in the medium and long term as it is built on and further adapted to different contexts and industry needs, thanks to its modular nature. The effectiveness of the i-DREAMS system may depend on a variety of factors, including the specific context in which it is implemented, the quality and accuracy of the data used to train the system, and the degree of integration and adoption by drivers and other stakeholders.

The integrated treatment of task complexity, coping capacity and risk can improve behavior and safety of all travelers and all transport modes, through the unobtrusive and seamless monitoring of behavior. Moreover, the feedback and training of travelers can also improve travel behavior, shift to safer and eco-friendly modes and eventually reduce their risk. Thus, authorities may use data systems at population level to plan mobility and safety interventions, set up road user incentives, optimize enforcement and enhance community building on safe traveling.

All in all, it is expected that the i-DREAMS platform, as a part of an Intelligent Transportation System (ITS) which is the cornerstone of the Information and Communication Technology (ICT) infrastructure, will improve the efficiency of traffic and road safety (i.e. reducing the number of road crashes, or the severity of crashes in the form of a decrease in fatalities, seriously injured and injured, or the number of crashes involving unprotected traffic participants - motorcyclists, cyclists and pedestrians). The introduction of this multidisciplinary approach (i.e. the integrated treatment of task complexity, coping capacity and risk) could be a solid base for the road transport safety planning and governance as well as the use of friendly techniques like Internet of Things (IoT) or gamification tools can be promoted for better travellers' connectivity and interaction with the systems and devices.

7 References

Abdel-Aty, M., Pande, A., Hsia, L. Y., & Abdalla, F. (2005). The potential for real-time traffic crash prediction. ITE Journal, 75(12), 69.

Afghari, A. P., Haque, M. M., & Washington, S. (2018). Applying fractional split model to examine the effects of roadway geometric and traffic characteristics on speeding behavior. Traffic injury prevention, 19(8), 860-866.

Afghari, A. P., Haque, M. M., & Washington, S. (2020). Applying a joint model of crash count and crash severity to identify road segments with high risk of fatal and serious injury crashes. Accident Analysis & Prevention, 144, 105615.

Afghari, A. P., Papadimitriou, E., Pilkington-Cheney, F., Filtness, A., Brijs, T., Brijs, K., ... & Rodrigues, L. (2022). Investigating the effects of sleepiness in truck drivers on their headway: An instrumental variable model with grouped random parameters and heterogeneity in their means. Analytic methods in accident research, 36, 100241.

Baumgartner, H., & Hombur, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. International Journal of Research in Marketing, 13, 139-161.

Brown, L., Papazikou, E., Talbot R., … , & Brijs, T. (2023). Effectiveness evaluation of the interventions. Deliverable 7.2 of the EC H2020 project i-DREAMS.

CardioID Technologies (2022). Heartmetrics Experts. Last assessed 15/05.22. Retrieved from: https://www.cardio-id.com/automotive/.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural equation modeling: a multidisciplinary journal, 14(3), 464-504.

Collins, M., Dasgupta, S., & Schapire, R. E. (2001). A generalization of principal components analysis to the exponential family. Advances in neural information processing systems, 14.

Fleming, J. M., Allison, C. K., Yan, X., Lot, R., & Stanton, N. A. (2019). Adaptive driver modelling in ADAS to improve user acceptance: A study using naturalistic data. Safety Science, 119, 76–83. https://doi.org/10.1016/j.ssci.2018.08.023

Franke, T., & Krems, J. F. (2013). Understanding charging behaviour of electric vehicle users. Transportation Research Part F: Traffic Psychology and Behaviour, 21, 75-89.

Garefalakis, T., Katrakazas, C., & Yannis, G. (2022). Data-driven estimation of a driving safety tolerance zone using imbalanced machine learning. Sensors, 22(14), 5309.

Girma, A., Yan, X., & Homaifar, A. (2019). Driver identification based on vehicle telematics data using LSTM-recurrent neural network. In 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 894-902). IEEE.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8), 681-698.

Habtemichael, F. G., & de Picado Santos, L. (2012). The need for transition from macroscopic to microscopic traffic management schemes to improve safety and mobility. Procedia-Social and Behavioral Sciences, 48, 3018-3029.

Hancox, G., Talbot, R., Brown, L., Filtness, A., Pilkington-Cheney, F., Brijs, K., Polders, E., Brijs, T., Ross, R., Katrakazas, C., Yannis G., De Vos, B., Gaspar, C., Lourenço, A., Carreiras, C., Al Haddad, C., Antoniou, C., Amini, R.E., & Kui, Y. (2021). Description of the on-road driving trials for identifying safety tolerance zones and the performance of in-vehicle interventions. Deliverable 5.3 of the EC H2020 project i-DREAMS.

Hastie, T. J., & Pregibon, D. (2017). Generalized linear models. In Statistical models in S (pp. 195-247). Routledge.

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. Biometrics, 1005-1016.

Hensher, D. A., Rose, J. M., Rose, J. M., & Greene, W. H. (2005). Applied choice analysis: a primer. Cambridge university press.

Katrakazas, C., Quddus, M., & Chen, W. H. (2019). A new integrated collision risk assessment methodology for autonomous vehicles. Accident Analysis & Prevention, 127, 61-79.

Katrakazas, C., Michelaraki, E., Yannis, G., Cuenen, A., Brijs, K., Brijs, T., Fitness, A., Talbot, R., Hancox, G., & Gruden, C. (2020). Methodology for the evaluation of interventions. Deliverable 7.1 of the EC H2020 project i-DREAMS.

Katrakazas, C., Michelaraki, E., Yannis, G., Kaiser, S., Senitschnig, N., Ross, V., ... & Taveira, R. (2020). Toolbox of recommended data collection tools and monitoring methods and a conceptual definition of the Safety Tolerance Zone. Deliverable 3.2 of the EC H2020 project i-DREAMS.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Wasserman, W. (2004). Applied linear regression models (Vol. 4, pp. 563-568). New York: McGraw-Hill/Irwin.

Lee, B., & Sohn, W. (2022). Testing the performance of level-specific fit evaluation in MCFA models with different factor structures across levels. Educational and Psychological Measurement, 82(6), 1153-1179.

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. Psychological methods, 7(1), 64.

McFadden, D. (1980). Econometric models for probabilistic choice among products. Journal of Business, S13-S29.

Michelaraki, E., Katrakazas, C., Yannis, G., Papazikou, E., Brown, L., Talbot, R., Afghari, A.P., Papadimitriou, E., Muhammad, A., Khattak, W.M., Brijs, K., & Brijs, T. (2023). Analysis of coping capacity factors: vehicle and operator state. Deliverable 6.2 of the EC H2020 project i-DREAMS.

Mobileye (2022). An Intel Company. Last assessed 15/05/2022. Retrieved from: https://www.mobileye.com/

Musicant, O., & Lotan, T. (2016). Can novice drivers be motivated to use a smartphone based app that monitors their behavior? Transportation Research Part F: Traffic Psychology and Behaviour, 42, 544–557. https://doi.org/10.1016/j.trf.2015.10.023

Panou, M. C. (2018). Intelligent personalized ADAS warnings. European Transport Research Review, 10(2), 59. https://doi.org/10.1186/s12544-018-0324-6

Papazikou, E., Brown, L., Talbot, R., Filtness, A., Michelaraki, E., Katrakazas, C., Yannis, G., Afghari, A.P., Papadimitriou, E., Muhammad, A., Khattak, W.M., Brijs, K., & Brijs, T. (2023). Analysis of task complexity factors. Deliverable 6.1 of the EC H2020 project i-DREAMS.

Papazikou, E., Quddus, M., Thomas, P., & Kidd, D. (2019). What came before the crash? An investigation through SHRP2 NDS data. Safety Science, 119, 150-161.

Silva, P. B., Andrade, M., & Ferreira, S. (2020). Machine learning applied to road safety modeling: A systematic literature review. Journal of traffic and transportation engineering (English edition), 7(6), 775-790.

Sheather, S. (2009). A modern approach to regression with R. Springer Science & Business Media.

Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transportation Research Part C: Emerging Technologies, 16(3), 320–331. https://doi.org/10.1016/j.trc.2008.01.001

Vlahogianni, E. I., & Barmpounakis, E. N. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges. Transportation Research Part C: Emerging Technologies, 79, 196-206.

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychological methods, 17(2), 228.

Washington, S.P., Karlaftis, M.G., & Mannering, F.L. (2011). Statistical and Econometric Methods for Transportation Data Analysis, second edition. CRC Press.

Washington, S., Karlaftis, M., Mannering, F., & Anastasopoulos, P. (2020). Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.

Yang, K., Alam, M. R., Al Haddad, C., Ezzati Amini, R., & Antoniou C. (2020). An algorithm implementation for safety tolerance zone calculation. Deliverable 4.1 of the EC H2020 project i-DREAMS.

Annex 1: Descriptive statistics for the available parameters

Belgium (Cars)

Table 110: Descriptive statistics for the available parameters in database used for Belgium car drivers

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NA	Description
Phase 1								
Gender - male	0	0	1	0,71	1	1		
Age	20	30	44	43,8	64	79		
Income	1	3	5	4,27	5	6		1: Less than €1.000, 2: €1.000 - €2.000, 3: €2000 - €3.000, 4: €3000 - €4.000, 5: €4000 - €5.000, 6: More than €5.000
% driving on urban roads	2	20	25	26,8	30	60		
% driving on rural roads	20	25	40	42,9	60	80		
Violation item 1	1	2	2	2,14	2,14	5		how often did you as a car driver, drive faster than the speed limit inside built-up areas? (1) Never, (2) Seldom, (3) About half the time, (4) Usually, (5) (almost) Always
Violation item 2	1	3	3	3,3	4	5		how often did you as a car driver, drive faster than the speed limit? (1) Never, (2) Seldom, (3) About half the time, (4) Usually, (5) (almost) Always
Driving style	0	0	0	0,45	1	1		0: Discrete average driver or Less experienced hesitant driver 1: Sportive ambitioned driver or risk-taking offensive driver
Confidence	0	0	1	0,65	1	1		How confident you are concerning your own driving skills? 1: Very confident or confident, 0: otherwise
Competence	3	3	3	3,503	4	5		How do you think you compare to the average driver, regarding general driving skills, I am: (1) Much worse, (2) Worse, (3) Not better nor worse, (4) Better, (5) Much better
Attitude item 1	1	2	3	3,27	5	5		Driving is … (5) Very dangerous, (4) Quite dangerous, (3) Neither dangerous nor safe, (2) Quite safe, (1) Very safe
Attitude item 2	3	4	4	4,14	4	5		a. I know the benefits of safe driving: (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, (5) Strongly agree
Highest level of education	0	0	1	0,71	1	1		1: college or above, 0: otherwise
Employment status	0	0	1	0,59	1	1		1: full time or part time employed, 0: otherwise
Headway - STZ 1	0	0	0,07	0,17	0,27	1		Proportion of events in 60 seconds
Headway - STZ 2	0	0	0	0,13	0,17	1		Proportion of events in 60 seconds

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NA	Description
Headway - STZ 3	0	0	0	0,03	0	1		Proportion of events in 60 seconds
Speeding - STZ 1	0	0	0,45	0,47	0,92	1	159	Proportion of events in 60 seconds
Speeding - STZ 2	0	0	0	0,05	0,02	1	824	Proportion of events in 60 seconds
Speeding - STZ 3	0	0	0	0,1	0,05	1	895	Proportion of events in 60 seconds
Harsh acceleration events	0	0	0,53	0,5	1	1	607	Proportion of events in 60 seconds
Harsh braking events	0	0	0	0,2	0,27	1	630	Proportion of events in 60 seconds
Harsh cornering events	0	0	0	0,31	0,7	1	895	Proportion of events in 60 seconds
KSS	35	35	35	35	35	39	5598	
IBI	376	755	807	811	871	1263	1230	
Wipers on	0	0	0	0,05	0	1		Proportion of events in 60 seconds
High beams on	0	0	0	0,018	0	1		Proportion of events in 60 seconds
FCW	0	0	0	0,02	0	3		Number of events in 60 seconds
PCW	0	0	0	0	0	2		Number of events in 60 seconds
Night-time driving	0	0	0	0,2	0	1		Proportion of events in 60 seconds
Day-time driving	0	1	1	0,78	1	1		Proportion of events in 60 seconds
Phase 2								
Gender - male	0	0	1	0,7	1	1		
Age	20	34	44	42,6	54	79		
Income	1	3	5	4,421	5	6		
% driving on urban roads	2	20	25	25,4	40	60		
% driving on rural roads	20	25	40	41,6	60	80		
Violation item 1	1	2	2,559	2,559	3	5		
Violation item 2	1	3	3	3,413	4	5		
Driving style	0	0	0	0,458	1	1		
Confidence	0	0	1	0,736	1	1		
Competence	3	3	3,561	3,561	4	5		
Attitude item 1	1	2	3	3,18	3,18	5		
Attitude item 2	3	4	4	4,203	4,203	5		
Highest level of education	0	0	1	0,684	1	1		
Employment status	0	0	1	0,615	1	1		
Headway - STZ 1	0	0	0,1	0,219	0,366	1		
Headway - STZ 2	0	0	0	0,117	0,15	1		
Headway - STZ 3	0	0	0	0,023	0	1		

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NA	Description
Speeding - STZ 1	0	0	0,4	0,454	0,933	1	68	
Speeding - STZ 2	0	0	0	0,032	0,016	1	920	
Speeding - STZ 3	0	0	0	0,097	0,033	1	1015	
Harsh acceleration events	0	0	0,433	0,47	1	1	775	
Harsh braking events	0	0	0	0,23	0,35	1	760	
Harsh cornering events	0	0	0	0,311	0,733	1	1379	
KSS	0	0	0	0,071	0	1		
IBI	0	0	0	0,012	0	3		
Wipers on	0	0	0	0,001	0	2		
High beams on	0	0	0	0,006	0	1		
FCW	35	35	35	35,01	35	39	1160	
PCW	371	751	791	798	858	1478	3497	
Night-time driving	0	0	0	0,105	0	1		
Day-time driving	0	1	1	0,877	1	1		
Phase 3								
Gender - male	0	0	1	0,728	1	1		
Age	20	30	43	43	60	79		
Income	1	3	5	4,391	5	6		
% driving on urban roads	2	20	25	26,6	40	60		
% driving on rural roads	20	30	40	40,2	55	80		
Violation item 1	1	2	3	2,795	3	5		
Violation item 2	1	3	3	3,514	5	5		
Driving style	0	0	0	0,47	1	1		
Confidence	0	1	1	0,754	1	1		
Competence	3	3	4	3,654	4	5		
Attitude item 1	1	3	3	3,384	5	5		
Attitude item 2	3	4	4	4,216	4,216	5		
Highest level of education	0	0	1	0,6	1	1		
Employment status	0	0	1	0,664	1	1		
Headway - STZ 1	0	0	0,1	0,217	0,366	1		
Headway - STZ 2	0	0	0	0,105	0,133	1		
Headway - STZ 3	0	0	0	0,018	0	1		
Speeding - STZ 1	0	0	0,45	0,478	1	1	1015	

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NA	Description
Speeding - STZ 2	0	0	0	0,224	0,333	1	1123	
Speeding - STZ 3	0	0	0	0,311	0,766	1	2145	
Harsh acceleration events	0	0	0	0,071	0	1		
Harsh braking events	0	0	0	0,011	0	3		
Harsh cornering events	0	0	0	0	0	2		
KSS	0	0	0,316	0,425	0,9	1	176	
IBI	0	0	0	0,027	0	1	1018	
Wipers on	0	0	0	0,111	0,05	1	1345	
High beams on	0	0	0	0,011	0	1		
FCW	35	35	35	35,01	35	39	18150	
PCW	319	753	818	815	857	1651	8750	
Night-time driving	0	0	0	0,104	0	1		
Day-time driving	0	1	1	0,874	1	1		
Phase 4								
Gender - male	0	0	1	0,7	1	1		
Age	20	30	43	42,5	54	79		
Income	1	3	5	4,381	5	6		
% driving on urban roads	2	20	25	26	35	60		
% driving on rural roads	20	30	40	40	55	80		
Violation item 1	1	2	3	2,818	3	5		
Violation item 2	1	3	3	3,496	5	5		
Driving style	0	0	0	0,478	1	1		
Confidence	0	0	1	0,709	1	1		
Competence	3	3	4	3,571	4	5		
Attitude item 1	1	3	3	3,331	5	5		
Attitude item 2	3	4	4	4,195	4,195	5		
Highest level of education	0	0	1	0,65	1	1		
Employment status	0	0	1	0,648	1	1		
Headway - STZ 1	0	0	0,1	0,223	0,366	1		
Headway - STZ 2	0	0	0	0,095	0,1	1		
Headway - STZ 3	0	0	0	0,013	0	1		
Speeding - STZ 1	0	0	0,45	0,476	1	1	827	
Speeding - STZ 2	0	0	0	0,239	0,383	1	901	

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NA	Description
Speeding - STZ 3	0	0	0	0,295	0,683	1	1934	
Harsh acceleration events	0	0	0	0,075	0	1		
Harsh braking events	0	0	0	0,011	0	4		
Harsh cornering events	0	0	0	0,001	0	2		
KSS	0	0	0,366	0,444	0,933	1		
IBI	0	0	0	0,027	0	1	1284	
Wipers on	0	0	0	0,096	0,016	1	2003	
High beams on	0	0	0	0,026	0	1		
FCW	35	35	35	35	35	35	25649	
PCW	471	762	829	822	867	1375	15919	
Night-time driving	0	0	0	0,133	0	1		
Day-time driving	0	1	1	0,84	1	1		

Belgium (Trucks)

Table 111: Descriptive statistics for the available parameters in database used for Belgium truck drivers

Variables	Min	Mean	Median	Std. Deviation	Max
Phase 1					
Speeding_STZ1	0,865	1,000	0,319	0,000	1,000
Speeding_STZ2	0,025	0,000	0,110	0,000	1,000
Speeding_STZ3	0,110	0,000	0,282	0,000	1,000
Trip duration	69,023	49,000	65,701	1,000	503,000
Age	45	50	11	25	56
Driving Style	2,220	2,000	0,414	2,000	3,000
Driver's Confidence	2,030	2,000	0,683	1,000	3,000
Speed	15,545	18,491	10,980	0,000	41,410
Phase 2					
Speeding_STZ1	0,836	1,000	0,346	0,000	1,000
Speeding_STZ2	0,013	0,000	0,072	0,000	1,000
Speeding_STZ3	0,151	0,000	0,335	0,000	1,000
Trip duration	81,190	54,000	84,452	1,000	749,000
Age	46	50	10	25	66
Driving Style	2,250	2,000	0,430	2,000	3,000

Driver's Confidence	2,140	2,000	0,700	1,000	3,000
Speed	15,937	18,491	11,848	0,000	43,750
Wipers	0,010	0,000	0,097	0,000	1,000
Phase 3					
Speeding_STZ1	0,885	1,000	0,295	0,000	1,000
Speeding_STZ2	0,011	0,000	0,067	0,000	1,000
Speeding_STZ3	0,104	0,000	0,282	0,000	1,000
Trip duration	92,083	59,000	102,783	1,000	791,000
Age	44	46	10	25	56
Driving Style	2,250	2,000	0,433	2,000	3,000
Driver's Confidence	2,040	2,000	0,707	1,000	3,000
Speed	15,086	17,977	11,283	0,000	39,820
Wipers	0,008	0,000	0,087	0,000	1,000
Phase 4					
Speeding_STZ1	0,820	1,000	0,360	0,000	1,000
Speeding_STZ2	0,014	0,000	0,073	0,000	1,000
Speeding_STZ3	0,167	0,000	0,350	0,000	1,000
Trip duration	99,532	68,000	100,835	1,000	779,000
Age	46	47	11	25	66
Driving Style	2,300	2,000	0,459	2,000	3,000
Driver's Confidence	2,190	2,000	0,743	1,000	3,000
Speed	16,379	19,412	12,102	0,000	41,220
Wipers	0,005	0,000	0,071	0,000	1,000

UK (Cars)

Table 112: Descriptive statistics for the available parameters in database used for UK car drivers

Variables	Min	Mean	Std. Deviation	Max	1st quartile	2nd quartile	3rd quartile
Phase 1 (total observations 113705)							
iDreams_Headway_Map_level_1_mean	0	0,151	0,266	1	0	0	0,2
iDreams_Headway_Map_level_1_0_mean	0	0,764	0,351	1	0,6	1	1
iDreams_Headway_Map_level_23_mean	0	0,085	0,215	1	0	0	0
ME_Car_wipers_median	0	0,063	0,243	1	0	0	0
ME_Car_high_beam_median	0	0,004	0,063	1	0	0	0

Variables	Min	Mean	Std. Deviation	Max	1st quartile	2nd quartile	3rd quartile
DrivingEvents_Map_evt_ha_mean	0	0,444	0,454	1	0	0,267	1
ME_LDW_Map_type_R_mean	0	0,163	0,365	1	0	0	0
SQ_Vehicle_age	3	9,48	3,138	16	7	9	11
EQ17_General_sleep_rating	0	0,035	0,184	1	0	0	0
EQ1a_Adaptive_cruise_control	0	0,011	0,103	1	0	0	0
EQ1b_Forward_collision_warning	0	0,067	0,249	1	0	0	0
EQ4b_Speed_limit	1	1,747	0,937	4	1	1	2
EQ4e_Mobile_phone	0	0,702	0,664	3	0	1	1
EQ4g_Illegal_overtake	0	0,348	0,53	2	0	0	1
EQ5_Driving_style	1	2,39	0,712	4	2	2	3
EQ6_Driving_confidence	2	4,038	0,653	5	4	4	4
Hour	0	12,844	4,263	23	9	13	16
Day_of_week	0	2,947	1,859	6	1	3	4
Month	3	7,847	3,03	11	4	10	10
Phase 2 (total observations 116917)							
iDreams_Headway_Map_level_1_0_mean	0	0,78	0,339	1	0,633	1	1
iDreams_Headway_Map_level_1_mean	0	0,51	0,425	1	0	0,5	1
iDreams_Headway_Map_level_23_mean	0	0,072	0,193	1	0	0	0
ME_Car_wipers_median	0	0,08	0,27	1	0	0	0
ME_Car_high_beam_median	0	0,004	0,066	1	0	0	0
DrivingEvents_Map_evt_ha_mean	0	0,435	0,452	1	0	0,233	1
ME_LDW_Map_type_R_mean	0	0,142	0,343	1	0	0	0
SQ_Vehicle_age	3	9,227	2,952	16	7	9	11
EQ17_General_sleep_rating	0	0,056	0,229	1	0	0	0
EQ1a_Adaptive_cruise_control	0	0,015	0,123	1	0	0	0
EQ1b_Forward_collision_warning	0	0,085	0,28	1	0	0	0
EQ4b_Speed_limit	1	1,741	0,895	4	1	1	2
EQ4e_Mobile_phone	0	0,727	0,717	3	0	1	1
EQ4g_Illegal_overtake	0	0,365	0,543	2	0	0	1
EQ5_Driving_style	1	2,391	0,733	4	2	2	3
EQ6_Driving_confidence	2	4,063	0,638	5	4	4	4
Hour	0	13,144	4,401	23	9	14	16

Variables	Min	Mean	Std. Deviation	Max	1st quartile	2nd quartile	3rd quartile
Day_of_week	0	2,852	1,928	6	1	3	5
Month	4	8,751	3,038	12	5	11	11
Phase 3 (total observations 119112)							
iDreams_Headway_Map_level_1_mean	0	0,138	0,254	1	0	0	0,167
iDreams_Headway_Map_level_1_0_mean	0	0,788	0,333	1	0,667	1	1
iDreams_Headway_Map_level_23_mean	0	0,074	0,198	1	0	0	0
ME_Car_wipers_median	0	0,098	0,297	1	0	0	0
ME_Car_high_beam_median	0	0,005	0,067	1	0	0	0
DrivingEvents_Map_evt_ha_mean	0	0,432	0,452	1	0	0,233	1
ME_LDW_Map_type_R_mean	0	0,154	0,357	1	0	0	0
SQ_Vehicle_age	3	9,267	3,149	16	7	9	11
EQ17_General_sleep_rating	0	0,054	0,225	1	0	0	0
EQ1a_Adaptive_cruise_control	0	0,02	0,141	1	0	0	0
EQ1b_Forward_collision_warning	0	0,083	0,275	1	0	0	0
EQ4b_Speed_limit	1	1,831	0,909	4	1	2	2
EQ4e_Mobile_phone	0	0,715	0,663	3	0	1	1
EQ4g_Illegal_overtake	0	0,367	0,555	2	0	0	1
EQ5_Driving_style	1	2,431	0,749	4	2	2	3
EQ6_Driving_confidence	2	4,091	0,66	5	4	4	4
Hour	0	12,909	4,356	23	10	13	16
Day_of_week	0	2,963	1,935	6	1	3	5
Month	1	8,452	3,644	12	6	7	12
Phase 4 (total observations 187948)							
iDreams_Headway_Map_level_1_0_mean	0	0,795	0,325	1	0,667	1	1
iDreams_Headway_Map_level_1_mean	0	0,551	0,42	1	0,067	0,667	1
iDreams_Headway_Map_level_23_mean	0	0,062	0,176	1	0	0	0
ME_Car_wipers_median	0	0,056	0,23	1	0	0	0
ME_Car_high_beam_median	0	0,005	0,067	1	0	0	0
DrivingEvents_Map_evt_ha_mean	0	0,431	0,449	1	0	0,233	1
ME_LDW_Map_type_R_mean	0	0,116	0,315	1	0	0	0
SQ_Vehicle_age	3	10,089	3,552	16	7	9	13
EQ17_General_sleep_rating	0	0,033	0,178	1	0	0	0

Variables	Min	Mean	Std. Deviation	Max	1st quartile	2nd quartile	3rd quartile
EQ1a_Adaptive_cruise_control	0	0,022	0,147	1	0	0	0
EQ1b_Forward_collision_warning	0	0,063	0,243	1	0	0	0
EQ4b_Speed_limit	1	1,806	0,866	4	1	2	2
EQ4e_Mobile_phone	0	0,758	0,642	3	0	1	1
EQ4g_Illegal_overtake	0	0,331	0,541	2	0	0	1
EQ5_Driving_style	1	2,379	0,714	4	2	2	3
EQ6_Driving_confidence	2	4,171	0,705	5	4	4	5
Hour	0	13,076	4,322	23	10	14	16
Day_of_week	0	2,984	1,891	6	1	3	5
Month	1	3,629	2,892	9	1	2	7

Germany (Cars)

Table 113: Descriptive statistics for the available parameters in database used for Germany car drivers

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
Phase 1 (total observations 48629)						
grpby_seconds	0	270	720	1333	1560	14610
iDreams_Headway_Map_level_.1_mean	1	1	1	1	1	1
iDreams_Headway_Map_level_.1_sum	2	30	30	29,98	30	30
iDreams_Headway_Map_level_0_mean		0	0	0	0	0
iDreams_Headway_Map_level_0_sum	0	0	0	0	0	0
iDreams_Headway_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_total_mean	0	0	0	0,0473	0	1
iDreams_Headway_Map_level_total_sum	0	0	0	1.419	0	30
iDreams_Overtaking_Map_level_0_mean	1	1	1	1	1	1
iDreams_Overtaking_Map_level_0_sum	2	30	30	29,98	30	30

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Overtaking_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_total_mean	0	0	0	0,0461	0	1
iDreams_Overtaking_Map_level_total_sum	0	0	0	1.383	0	30
iDreams_Speeding_Map_level_0_mean	0	0	0	0,317	1.000	1
iDreams_Speeding_Map_level_0_sum	0	0	0	9.523	30.000	30
iDreams_Speeding_Map_level_1_mean	0	0	1.000	0,691	1.000	1
iDreams_Speeding_Map_level_1_sum	0	0	30	20,74	30	30
iDreams_Speeding_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_total_mean	0	10.000	10.000	0,8718	10.000	1
iDreams_Speeding_Map_level_total_sum	0	30	30	26,15	30	30
iDreams_Fatigue_Map_level_0_mean	0	1	1	0,79	1	1
iDreams_Fatigue_Map_level_0_sum	0	30	30	23,68	30	30
iDreams_Fatigue_Map_level_1_mean	0	0	0	0,25	0	1
iDreams_Fatigue_Map_level_1_sum	0	0	0	7,36	0	30
iDreams_Fatigue_Map_level_2_mean	0	0	0	0,14	0	1
iDreams_Fatigue_Map_level_2_sum	0	0	0	4,2	0	30
iDreams_Fatigue_Map_level_3_mean	0	0	0	0	0	0
iDreams_Fatigue_Map_level_3_sum	0	0	0	0	0	0
iDreams_Fatigue_Map_level_total_mean	0	0	0	0,1385	0	1
iDreams_Fatigue_Map_level_total_sum	0	0	0	4.155	0	30
DrivingEvents_Map_lvl_L_mean	0	0,433	1.000	0,736	1.000	1
DrivingEvents_Map_lvl_M_mean	0	0	0	0,242	0,4	1
DrivingEvents_Map_lvl_H_mean	0	0	0	0,062	0	1
ME_Car_speed_mean	0	0	0	0	0	0

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
ME_Car_wipers_median	0	0	0	0	0	0
ME_Car_high_beam_median	0	0	0	0	0	0
ME_AWS_hw_measurement_mean	0	0	0	0	0	0
ME_AWS_tsr_level_mean	0	0	0	0	0	0
ME_AWS_fcw_mean	0	0	0	0	0	0
ME_AWS_pcw_mean	0	0	0	0	0	0
ME_AWS_pedestrian_dz_mean	0	0	0	0	0	0
ME_AWS_time_indicator_median	1.000	1.000	1.000	1.709	3.000	3.000
ME_TSR_tsr_1_speed_median	0	5	9	64,33	39	254
GPS_spd_mean	0	26,05	52,68	53,2	71,74	198,58
GPS_distances_sum	0	221	450,6	455,7	611,5	14239,8
DrivingEvents_Map_evt_ha_mean	0	0	0,567	0,509	1.000	1
DEM_evt_ha_lvl_L_mean	0	0	0	0,403	1.000	1
DEM_evt_ha_lvl_L_sum	0	0	0	12,09	30	30
DEM_evt_ha_lvl_M_mean	0	0	0	0,088	0	1
DEM_evt_ha_lvl_M_sum	0	0	0	2.631	0	30
DEM_evt_ha_lvl_H_mean	0	0	0	0,018	0	1
DEM_evt_ha_lvl_H_sum	0	0	0	0,55	0	30
DrivingEvents_Map_evt_hc_mean	0	0	0	0,309	0,967	1
DEM_evt_hc_lvl_L_mean	0	0	0	0,154	0	1
DEM_evt_hc_lvl_L_sum	0	0	0	4.623	0	30
DEM_evt_hc_lvl_M_mean	0	0	0	0,143	0	1
DEM_evt_hc_lvl_M_sum	0	0	0	4.301	0	30
DEM_evt_hc_lvl_H_mean	0	0	0	0,012	0	1
DEM_evt_hc_lvl_H_sum	0	0	0	0,355	0	30
DrivingEvents_Map_evt_hb_mean	0	0	0	0,209	0,233	1
DEM_evt_hb_lvl_L_mean	0	0	0	0,197	0,167	1
DEM_evt_hb_lvl_L_sum	0	0	0	5.913	5.000	30
DEM_evt_hb_lvl_M_mean	0	0	0	0,009	0	1
DEM_evt_hb_lvl_M_sum	0	0	0	0,264	0	30

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
DEM_evt_hb_lvl_H_mean	0	0	0	0,003	0	1
DEM_evt_hb_lvl_H_sum	0	0	0	0,082	0	30
Drowsiness_level_median	35	35	35	35,1	35	39
IBI_value_mean	421,9	728,5	794,8	797,6	861,5	1788,1
ME_LDW_Map_type_L_mean	-9999	-9999	-9999	-9999	-9999	-9999
ME_LDW_Map_type_R_mean	-9999	-9999	-9999	-9999	-9999	-9999
Phase 2 (total observations 48629)						
grpby_seconds	0	240	570	1141	1230	13500
iDreams_Headway_Map_level_.1_mean	1	1	1	1	1	1
iDreams_Headway_Map_level_.1_sum	26	30	30	30	30	30
iDreams_Headway_Map_level_0_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_0_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_total_mean	0	0	0	0,0591	0	1
iDreams_Headway_Map_level_total_sum	0	0	0	1.773	0	30
iDreams_Overtaking_Map_level_0_mean	1	1	1	1	1	1
iDreams_Overtaking_Map_level_0_sum	26	30	30	30	30	30
iDreams_Overtaking_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_total_mean	0	0	0	0,0588	0	1
iDreams_Overtaking_Map_level_total_sum	0	0	0	1.763	0	30
iDreams_Speeding_Map_level_0_mean	0	0	0	0,243	0,175	1
iDreams_Speeding_Map_level_0_sum	0	0	0	7,29	5,25	30

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Speeding_Map_level_1_mean	0	1.000	1.000	0,768	1.000	1
iDreams_Speeding_Map_level_1_sum	0	30	30	23,05	30	30
iDreams_Speeding_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_total_mean	0	10.000	10.000	0,8771	10.000	1
iDreams_Speeding_Map_level_total_sum	0	30	30	26,31	30	30
iDreams_Fatigue_Map_level_0_mean	0	1	1	0,78	1	1
iDreams_Fatigue_Map_level_0_sum	0	30	30	23,36	30	30
iDreams_Fatigue_Map_level_1_mean	0	0	0	0,32	1	1
iDreams_Fatigue_Map_level_1_sum	0	0	0	9,48	30	30
iDreams_Fatigue_Map_level_2_mean	0	0	0	0,07	0	1
iDreams_Fatigue_Map_level_2_sum	0	0	0	1,97	0	30
iDreams_Fatigue_Map_level_3_mean	0	0	0	0,02	0	1
iDreams_Fatigue_Map_level_3_sum	0	0	0	0,61	0	30
iDreams_Fatigue_Map_level_total_mean	0	0	0	0,1304	0	1
iDreams_Fatigue_Map_level_total_sum	0	0	0	3.912	0	30
DrivingEvents_Map_lvl_L_mean	0	0,367	1.000	0,725	1.000	1
DrivingEvents_Map_lvl_M_mean	0	0	0	0,254	0,5	1
DrivingEvents_Map_lvl_H_mean	0	0	0	0,058	0	1
ME_Car_speed_mean	0	0	0	0	0	0
ME_Car_wipers_median	0	0	0	0	0	0
ME_Car_high_beam_median	0	0	0	0	0	0
ME_AWS_hw_measurement_mean	0	0	0	0	0	0
ME_AWS_tsr_level_mean	0	0	0	0	0	0
ME_AWS_fcw_mean	0	0	0	0	0	0
ME_AWS_pcw_mean	0	0	0	0	0	0
ME_AWS_pedestrian_dz_mean	0	0	0	0	0	0
ME_AWS_time_indicator_median	1.000	1.000	1.000	1.497	2.000	3.000

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
ME_TSR_tsr_1_speed_median	0	5	7	64,27	39	254
GPS_spd_mean	0	23,35	52,07	50,43	68,85	224,05
GPS_distances_sum	0	199,2	439,2	432,2	585,2	30601,2
DrivingEvents_Map_evt_ha_mean	0	0	0,567	0,511	1.000	1
DEM_evt_ha_lvl_L_mean	0	0	0,033	0,404	1.000	1
DEM_evt_ha_lvl_L_sum	0	0	1	12,12	30	30
DEM_evt_ha_lvl_M_mean	0	0	0	0,089	0	1
DEM_evt_ha_lvl_M_sum	0	0	0	2.681	0	30
DEM_evt_ha_lvl_H_mean	0	0	0	0,017	0	1
DEM_evt_ha_lvl_H_sum	0	0	0	0,519	0	30
DrivingEvents_Map_evt_hc_mean	0	0	0	0,318	0,933	1
DEM_evt_hc_lvl_L_mean	0	0	0	0,15	0,033	1
DEM_evt_hc_lvl_L_sum	0	0	0	4.511	1.000	30
DEM_evt_hc_lvl_M_mean	0	0	0	0,155	0	1
DEM_evt_hc_lvl_M_sum	0	0	0	4,66	0	30
DEM_evt_hc_lvl_H_mean	0	0	0	0,012	0	1
DEM_evt_hc_lvl_H_sum	0	0	0	0,357	0	30
DrivingEvents_Map_evt_hb_mean	0	0	0	0,197	0,167	1
DEM_evt_hb_lvl_L_mean	0	0	0	0,188	0,133	1
DEM_evt_hb_lvl_L_sum	0	0	0	5.634	4.000	30
DEM_evt_hb_lvl_M_mean	0	0	0	0,008	0	1
DEM_evt_hb_lvl_M_sum	0	0	0	0,235	0	30
DEM_evt_hb_lvl_H_mean	0	0	0	0,001	0	1
DEM_evt_hb_lvl_H_sum	0	0	0	0,026	0	30
Drowsiness_level_median	35	35	35	35,17	35	39
IBI_value_mean	342,8	738,1	809,2	806,5	877,3	1636,7
ME_LDW_Map_type_L_mean	-9999	-9999	-9999	-9999	-9999	-9999
ME_LDW_Map_type_R_mean	-9999	-9999	-9999	-9999	-9999	-9999
Phase 3 (total observations 36606)						
grpby_seconds	0	240	630	1329	1500	12270

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Headway_Map_level_.1_mean	1	1	1	1	1	1
iDreams_Headway_Map_level_.1_sum	1	30	30	29,96	30	30
iDreams_Headway_Map_level_0_mean	0	0	0	0	0	0
iDreams_Headway_Map_level_0_sum	0	0	0	0	0	0
iDreams_Headway_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_total_mean	0	0	0	0,0211	0	1
iDreams_Headway_Map_level_total_sum	0	0	0	0,6327	0	30
iDreams_Overtaking_Map_level_0_mean	1	1	1	1	1	1
iDreams_Overtaking_Map_level_0_sum	1	30	30	29,96	30	30
iDreams_Overtaking_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_total_mean	0	0	0	0,021	0	1
iDreams_Overtaking_Map_level_total_sum	0	0	0	0,6294	0	30
iDreams_Speeding_Map_level_0_mean	0	0	0	0,234	0	1
iDreams_Speeding_Map_level_0_sum	0	0	0	7.022	0	30
iDreams_Speeding_Map_level_1_mean	0	1.000	1.000	0,777	1.000	1
iDreams_Speeding_Map_level_1_sum	0	30	30	23,32	30	30
iDreams_Speeding_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_total_mean	0	1	1	0,8829	1	1
iDreams_Speeding_Map_level_total_sum	0	30	30	26,49	30	30

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Fatigue_Map_level_0_mean	0	0	1.000	0,687	1.000	1
iDreams_Fatigue_Map_level_0_sum	0	0	30	20,6	30	30
iDreams_Fatigue_Map_level_1_mean	0	0	0	0,31	1	1
iDreams_Fatigue_Map_level_1_sum	0	0	0	9,29	30	30
iDreams_Fatigue_Map_level_2_mean	0	0	0	0,33	1	1
iDreams_Fatigue_Map_level_2_sum	0	0	0	10,01	30	30
iDreams_Fatigue_Map_level_3_mean	0	0	0	0	0	0
iDreams_Fatigue_Map_level_3_sum	0	0	0	0	0	0
iDreams_Fatigue_Map_level_total_mean	0	0	0	0,141	0	1
iDreams_Fatigue_Map_level_total_sum	0	0	0	4.229	0	30
DrivingEvents_Map_lvl_L_mean	0	0,533	1.000	0,752	1.000	1
DrivingEvents_Map_lvl_M_mean	0	0	0	0,235	0,367	1
DrivingEvents_Map_lvl_H_mean	0	0	0	0,064	0	1
ME_Car_speed_mean	0	0	0	0	0	0
ME_Car_wipers_median	0	0	0	0	0	0
ME_Car_high_beam_median	0	0	0	0	0	0
ME_AWS_hw_measurement_mean	0	0	0	0	0	0
ME_AWS_tsr_level_mean	0	0	0	0	0	0
ME_AWS_fcw_mean	0	0	0	0	0	0
ME_AWS_pcw_mean	0	0	0	0	0	0
ME_AWS_pedestrian_dz_mean	0	0	0	0	0	0
ME_AWS_time_indicator_median	1.000	1.000	1.000	1.456	2.000	3.000
ME_TSR_tsr_1_speed_median	0	4	7	62,37	39	254
GPS_spd_mean	0	27,65	52,68	54,86	77,51	200,69
GPS_distances_sum	0	234,2	450,6	468,7	657,3	14773,7
DrivingEvents_Map_evt_ha_mean	0	0	0,533	0,501	1.000	1
DEM_evt_ha_lvl_L_mean	0	0	0	0,41	1	1
DEM_evt_ha_lvl_L_sum	0	0	0	12,3	30	30
DEM_evt_ha_lvl_M_mean	0	0	0	0,075	0	1
DEM_evt_ha_lvl_M_sum	0	0	0	2.256	0	30

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
DEM_evt_ha_lvl_H_mean	0	0	0	0,015	0	1
DEM_evt_ha_lvl_H_sum	0	0	0	0,459	0	30
DrivingEvents_Map_evt_hc_mean	0	0	0	0,311	0,933	1
DEM_evt_hc_lvl_L_mean	0	0	0	0,156	0,033	1
DEM_evt_hc_lvl_L_sum	0	0	0	4.685	1.000	30
DEM_evt_hc_lvl_M_mean	0	0	0	0,141	0	1
DEM_evt_hc_lvl_M_sum	0	0	0	4.226	0	30
DEM_evt_hc_lvl_H_mean	0	0	0	0,014	0	1
DEM_evt_hc_lvl_H_sum	0	0	0	0,426	0	30
DrivingEvents_Map_evt_hb_mean	0	0	0	0,221	0,267	1
DEM_evt_hb_lvl_L_mean	0	0	0	0,209	0,2	1
DEM_evt_hb_lvl_L_sum	0	0	0	6.256	6.000	30
DEM_evt_hb_lvl_M_mean	0	0	0	0,011	0	1
DEM_evt_hb_lvl_M_sum	0	0	0	0,326	0	30
DEM_evt_hb_lvl_H_mean	0	0	0	0,001	0	1
DEM_evt_hb_lvl_H_sum	0	0	0	0,038	0	30
Drowsiness_level_median	35	35	35	35,21	35	39
IBI_value_mean	338,9	722,3	783,8	787,9	848,3	1265,6
ME_LDW_Map_type_L_mean	-9999	-9999	-9999	-9999	-9999	-9999
ME_LDW_Map_type_R_mean	-9999	-9999	-9999	-9999	-9999	-9999
Phase 4 (total observations 48784)						
grpby_seconds	0	270	660	1162	1410	11220
iDreams_Headway_Map_level_.1_mean	0,97	1	1	1	1	1
iDreams_Headway_Map_level_.1_sum	29	30	30	29,99	30	30
iDreams_Headway_Map_level_0_mean	0	0	0	0	0	0,03
iDreams_Headway_Map_level_0_sum	0	0	0	0,14	0	1
iDreams_Headway_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Headway_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Headway_Map_level_total_mean	0	0	0	0,0428	0	1
iDreams_Headway_Map_level_total_sum	0	0	0	1.285	0	30
iDreams_Overtaking_Map_level_0_mean	1	1	1	1	1	1
iDreams_Overtaking_Map_level_0_sum	30	30	30	30	30	30
iDreams_Overtaking_Map_level_1_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_1_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Overtaking_Map_level_total_mean	0	0	0	0,0412	0	1
iDreams_Overtaking_Map_level_total_sum	0	0	0	1.235	0	30
iDreams_Speeding_Map_level_0_mean	0	0	0	0,236	0	1
iDreams_Speeding_Map_level_0_sum	0	0	0	7.067	0	30
iDreams_Speeding_Map_level_1_mean	0	1.000	1.000	0,772	1.000	1
iDreams_Speeding_Map_level_1_sum	0	30	30	23,16	30	30
iDreams_Speeding_Map_level_2_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_2_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_mean	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_3_sum	-9999	-9999	-9999	-9999	-9999	-9999
iDreams_Speeding_Map_level_total_mean	0	1	1	0,8881	1	1
iDreams_Speeding_Map_level_total_sum	0	30	30	26,64	30	30
iDreams_Fatigue_Map_level_0_mean	0	1	1	0,8	1	1
iDreams_Fatigue_Map_level_0_sum	0	30	30	24,08	30	30
iDreams_Fatigue_Map_level_1_mean	0	0	0	0,17	0	1
iDreams_Fatigue_Map_level_1_sum	0	0	0	5,04	0	30
iDreams_Fatigue_Map_level_2_mean	0	0	1	0,57	1	1
iDreams_Fatigue_Map_level_2_sum	0	0	30	17,11	30	30
iDreams_Fatigue_Map_level_3_mean	0	0	0	0	0	0
iDreams_Fatigue_Map_level_3_sum	0	0	0	0	0	0

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Fatigue_Map_level_total_mean	0	0	0	0,1138	0	1
iDreams_Fatigue_Map_level_total_sum	0	0	0	3.414	0	30
DrivingEvents_Map_lvl_L_mean	0	0,5	1	0,744	1	1
DrivingEvents_Map_lvl_M_mean	0	0	0	0,25	0,467	1
DrivingEvents_Map_lvl_H_mean	0	0	0	0,055	0	1
ME_Car_speed_mean	0	0	0	0	0	0
ME_Car_wipers_median	0	0	0	0	0	0
ME_Car_high_beam_median	0	0	0	0	0	0
ME_AWS_hw_measurement_mean	0	0	0	0	0	0
ME_AWS_tsr_level_mean	0	0	0	0	0	0
ME_AWS_fcw_mean	0	0	0	0	0	0
ME_AWS_pcw_mean	0	0	0	0	0	0
ME_AWS_pedestrian_dz_mean	0	0	0	0	0	0
ME_AWS_time_indicator_median	1.000	1.000	1.000	1.529	2.000	3.000
ME_TSR_tsr_1_speed_median	0	4	7	60,54	39	254
GPS_spd_mean	0	25,45	51,16	52,39	73,94	172,63
GPS_distances_sum	0	215,2	430,3	447,1	627,8	8162,7
DrivingEvents_Map_evt_ha_mean	0	0	0,333	0,465	1.000	1
DEM_evt_ha_lvl_L_mean	0	0	0	0,378	1.000	1
DEM_evt_ha_lvl_L_sum	0	0	0	11,34	30	30
DEM_evt_ha_lvl_M_mean	0	0	0	0,077	0	1
DEM_evt_ha_lvl_M_sum	0	0	0	2.307	0	30
DEM_evt_ha_lvl_H_mean	0	0	0	0,01	0	1
DEM_evt_ha_lvl_H_sum	0	0	0	0,314	0	30
DrivingEvents_Map_evt_hc_mean	0	0	0	0,345	1.000	1
DEM_evt_hc_lvl_L_mean	0	0	0	0,178	0,033	30
DEM_evt_hc_lvl_L_sum	0	0	0	5.336	1.000	30
DEM_evt_hc_lvl_M_mean	0	0	0	0,155	0	1
DEM_evt_hc_lvl_M_sum	0	0	0	4.661	0	30
DEM_evt_hc_lvl_H_mean	0	0	0	0,012	0	1

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
DEM_evt_hc_lvl_H_sum	0	0	0	0,351	0	30
DrivingEvents_Map_evt_hb_mean	0	0	0	0,23	0,3	1
DEM_evt_hb_lvl_L_mean	0	0	0	0,218	0,233	1
DEM_evt_hb_lvl_L_sum	0	0	0	6.525	7.000	30
DEM_evt_hb_lvl_M_mean	0	0	0	0,01	0	1
DEM_evt_hb_lvl_M_sum	0	0	0	0,304	0	30
DEM_evt_hb_lvl_H_mean	0	0	0	0,002	0	1
DEM_evt_hb_lvl_H_sum	0	0	0	0,069	0	30
Drowsiness_level_median	35	35	35	35,07	35	39
IBI_value_mean	374,8	737,3	798,5	805,4	868,3	1776,4
ME_LDW_Map_type_L_mean	-9999	-9999	-9999	-9999	-9999	-9999
ME_LDW_Map_type_R_mean	-9999	-9999	-9999	-9999	-9999	-9999

Greece (Cars)

Table 114: Descriptive statistics for the available parameters in database used for Greek car drivers

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
Phase 1								
trip_distance	0.5	6.4	10.9	32.8	22.2	334.7	NA	
time_indicator	1.0	1.0	2.0	1.8	3.0	3.0	NA	1: day , 2: dusk, 3: night
VC_acc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_acc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_acc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
VC_dc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_dc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_dc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
Speed_high_sum	0.0	0.0	0.0	0.1	0.0	2.0	NA	
Speed_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
Speed_no_low_sum	27.0	30.0	30.0	29.9	30.0	30.0	NA	
distraction_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
VC_acc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
VC_dc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
Speed_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
distraction_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Gender	0.0	0.0	0.0	0.4	1.0	1.0	NA	0: Male, 1: Female
SQ_Nationality	1.0	1.0	1.0	1.0	1.0	1.0	NA	1: Greek
SQ_Year_of_birth	1964.0	1993.0	1995.0	1993.0	1998.0	2000.0	NA	
SQ_Age	22.0	24.0	27.0	29.2	29.0	58.0	NA	
SQ_ Age_got_driving_license	18.0	18.0	18.0	18.9	19.0	27.0	NA	
SQ_Years_driving	2.0	5.0	8.0	10.3	11.0	40.0	NA	
SQ_Vehicle_brand	1.0	6.0	14.0	13.5	21.0	22.0	NA	
SQ_Vehicle_age	0.0	4.0	7.0	9.4	16.0	22.0	116.0	
STC_Second_Nat	0.0	0.0	0.0	0.0	0.0	0.0	NA	0: No, 1: Yes
STC_Highest_lev_education	1.0	2.0	2.0	2.5	3.0	5.0	NA	1:Higher education , 2:Highest education , 3: Master of Science, 4:PhD, 5:Secondary education
STC_Current_occupation	1.0	2.0	3.0	3.2	4.0	5.0	NA	1:Civil servant, 2:Freelancer / self-employed, 3: Military service, 4:Student, 5:Private employee
STC_Employment_stat	1.0	1.0	1.0	1.9	4.0	4.0	NA	1:Employed full time, 2:Employed part time, 3: Military service, 4: Student
STC_Net_income	1.0	1.0	3.0	3.4	5.0	7.0	NA	1: Less than €1.000, 2: €1.000 - €2.000, 3: €2000 - €3.000, 4: €3000 - €4.000, 5: €4000 - €5.000, 6: More than €5.000
STC_Med_condition_decleration	0.0	1.0	1.0	0.9	1.0	1.0	NA	Can you declare that you are not suffering from a medical condition that would be considered a legal exclusion to drive? 0: No, 1: Yes

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
STC_Fuel_type	1.0	2.0	3.0	2.5	3.0	3.0	NA	1:diesel, 2: hybrid, 3: petrol
STC_Gearbox	1.0	1.0	2.0	1.7	2.0	2.0	NA	1:Manual, 2: Automatic
Phase 3								
trip_distance	0.5	6.8	11.8	38.4	26.8	319.7	NA	
time_indicator	1.0	1.0	2.0	1.7	2.0	3.0	NA	1: day , 2: dusk, 3: night
VC_acc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_acc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_acc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
VC_dc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_dc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_dc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
Speed_high_sum	0.0	0.0	0.0	0.1	0.0	2.0	NA	
Speed_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
Speed_no_low_sum	27.0	30.0	30.0	29.9	30.0	30.0	NA	
distraction_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_acc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
VC_dc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
Speed_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
distraction_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Gender	0.0	0.0	0.0	0.4	1.0	1.0	NA	0: Male, 1: Female
SQ_Nationality	1.0	1.0	1.0	1.0	1.0	1.0	NA	1: Greek
SQ_Year_of_birth	1964.0	1990.0	1995.0	1992.0	1998.0	2000.0	NA	
SQ_Age	22.0	24.0	27.0	29.9	33.0	58.0	NA	

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
SQ_ Age_got_driving_license	18.0	18.0	18.0	18.9	19.0	27.0	NA	
SQ_Years_driving	3.0	5.0	9.0	11.0	14.0	40.0	NA	
SQ_Vehicle_brand	1.0	7.0	15.0	13.9	21.0	22.0	NA	
SQ_Vehicle_age	0.0	5.0	7.0	9.7	16.0	22.0	NA	
STC_Second_Nat	0.0	0.0	0.0	0.0	0.0	0.0	NA	0: No, 1: Yes
STC_Highest_lev_education	1.0	2.0	2.0	2.5	3.0	5.0	NA	1:Higher education , 2:Highest education , 3: Master of Science, 4:PhD, 5:Secondary education
STC_Current_occupation	1.0	2.0	2.0	3.1	4.0	5.0	NA	1:Civil servant, 2:Freelancer / self-employed, 3: Military service, 4:Student, 5:Private employee
STC_Employment_stat	1.0	1.0	1.0	1.8	3.0	4.0	NA	1:Employed full time, 2:Employed part time, 3: Military service, 4: Student
STC_Net_income	1.0	1.0	5.0	3.6	5.0	7.0	NA	1: Less than €1.000, 2: €1.000 - €2.000, 3: €2000 - €3.000, 4: €3000 - €4.000, 5: €4000 - €5.000, 6: More than €5.000
STC_Med_condition_decleration	0.0	1.0	1.0	0.9	1.0	1.0	NA	Can you declare that you are not suffering from a medical condition that would be considered a legal exclusion to drive? 0: No, 1: Yes
STC_Fuel_type	1.0	2.0	3.0	2.5	3.0	3.0	NA	1:diesel, 2: hybrid, 3: petrol
STC_Gearbox	1.0	1.0	1.0	1.3	2.0	2.0	NA	1:Manual, 2: Automatic
Phase 4								
trip_distance	0.5	7.0	11.6	31.8	27.6	299.9	NA	
time_indicator	1.0	1.0	2.0	1.7	2.0	3.0	NA	1: day , 2: dusk, 3: night
VC_acc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_acc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_acc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
VC_dc_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
VC_dc_high_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_dc_no_low_sum	27.0	30.0	30.0	30.0	30.0	30.0	NA	
Speed_high_sum	0.0	0.0	0.0	0.1	0.0	2.0	NA	
Speed_medium_sum	0.0	0.0	0.0	0.0	0.0	2.0	NA	
Speed_no_low_sum	27.0	30.0	30.0	29.9	30.0	30.0	NA	

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
distraction_sum	0.0	0.0	0.0	0.0	0.0	3.0	NA	
VC_acc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_acc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
VC_dc_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
VC_dc_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
Speed_high_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_medium_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Speed_no_low_mean	0.9	1.0	1.0	1.0	1.0	1.0	NA	
distraction_mean	0.0	0.0	0.0	0.0	0.0	0.1	NA	
Gender	0.0	0.0	0.0	0.4	1.0	1.0	NA	0: Male, 1: Female
SQ_Nationality	1.0	1.0	1.0	1.0	1.0	1.0	NA	1: Greek
SQ_Year_of_birth	1964.0	1990.0	1995.0	1992.0	1999.0	2000.0	NA	
SQ_Age	22.0	23.0	27.0	29.5	33.0	58.0	NA	
SQ_ Age_got_driving_license	18.0	18.0	18.0	18.8	19.0	27.0	NA	
SQ_Years_driving	3.0	5.0	8.0	10.7	11.0	40.0	NA	
SQ_Vehicle_brand	1.0	11.0	15.0	14.0	21.0	22.0	NA	
SQ_Vehicle_age	0.0	4.0	7.0	9.6	15.0	22.0	NA	
STC_Second_Nat	0.0	0.0	0.0	0.0	0.0	0.0	NA	0: No, 1: Yes
STC_Highest_lev_education	1.0	2.0	2.0	2.4	3.0	5.0	NA	1:Higher education , 2:Highest education , 3: Master of Science, 4:PhD, 5:Secondary education
STC_Current_occupation	1.0	2.0	4.0	3.2	4.0	5.0	NA	1:Civil servant, 2:Freelancer / self-employed, 3: Military service, 4:Student, 5:Private employee
STC_Employment_stat	1.0	1.0	1.0	2.1	4.0	4.0	NA	1:Employed full time, 2:Employed part time, 3: Military service, 4: Student
STC_Net_income	1.0	1.0	5.0	3.8	5.0	7.0	NA	1: Less than €1.000, 2: €1.000 - €2.000, 3: €2000 - €3.000, 4: €3000 - €4.000, 5: €4000 - €5.000, 6: More than €5.000

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max	NAs	Description
STC_Med_condition_decleration	0.0	1.0	1.0	0.9	1.0	1.0	NA	Can you declare that you are not suffering from a medical condition that would be considered a legal exclusion to drive? 0: No, 1: Yes
STC_Fuel_type	1.0	2.0	3.0	2.5	3.0	3.0	NA	1:diesel, 2: hybrid, 3: petrol
STC_Gearbox	1.0	1.0	1.0	1.3	2.0	2.0	NA	1:Manual, 2: Automatic

Portugal (Buses)

Table 115: Descriptive statistics for the available parameters in database used for Portuguese bus drivers

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
Phase 1						
duration	785.00	7217.00	7217.00	6359.00	7217.00	16598.00
iDreams_Headway_Map_level_.1_mean	-1.00	-1.00	-1.00	-0.98	-1.00	0.20
iDreams_Headway_Map_level_0_mean	-1.00	0.00	0.00	0.00	0.00	0.20
iDreams_Headway_Map_level_1_mean	-1.00	0.00	0.00	0.00	0.00	0.20
iDreams_Headway_Map_level_2_mean	-1.00	0.00	0.00	0.00	0.00	0.20
iDreams_Headway_Map_level_3_mean	-0.64	0.00	0.00	0.00	0.00	0.15
Headway_level_initial	-1.00	-1.00	-1.00	-0.98	-1.00	3.00
Headway_level	-1.00	-1.00	-1.00	-0.98	-1.00	0.20
Headway_avg_level	-1.00	-1.00	-1.00	-0.98	-1.00	3.00
iDreams_Speeding_Map_level_0_mean	0.00	0.00	0.00	0.01	0.00	0.94
iDreams_Speeding_Map_level_1_mean	0.00	0.00	0.00	0.00	0.00	0.94
iDreams_Speeding_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	0.94
iDreams_Speeding_Map_level_3_mean	0.00	0.00	0.00	0.00	0.00	0.94
Speeding_level_Initial	0.00	0.00	0.00	0.01	0.00	3.00
Speeding_level	0.00	0.00	0.00	0.01	0.00	0.94
Speeding_avg_level	0.00	0.00	0.00	0.01	0.00	3.00
iDreams_Overtaking_Map_level_0_mean	0.00	0.00	0.00	0.00	0.00	0.20
iDreams_Overtaking_Map_level_1_mean	0.00	0.00	0.00	0.00	0.00	0.16
iDreams_Overtaking_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	0.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
Ovetaking_level_initial	0.00	0.00	0.00	0.00	0.00	2.00
Overtaking_level	0.00	0.00	0.00	0.00	0.00	0.20
Overtaking_avg_level	0.00	0.00	0.00	0.00	0.00	1.00
iDreams_Fatigue_Map_level_0_mean	0.00	0.00	0.00	0.01	0.00	2.83
iDreams_Fatigue_Map_level_1_mean	0.00	0.00	1.00	0.71	1.00	2.83
iDreams_Fatigue_Map_level_2_mean	0.00	0.00	0.00	0.04	0.00	2.83
iDreams_Fatigue_Map_level_3_mean	0.00	0.00	0.00	0.07	0.00	2.84
Fatigue_level_initial	0.00	1.00	1.00	0.83	1.00	3.00
Fatique_level	0.00	1.00	1.00	0.83	1.00	2.84
Fatigue_avg_level	0.00	1.00	1.00	0.83	1.00	3.00
DrivingEvents_Map_evt_ha_mean	0.00	0.00	0.00	0.03	0.00	0.44
DrivingEvents_Map_evt_hb_mean	-0.89	0.00	0.00	-0.05	0.00	0.00
Driving_events_maxg	-0.89	-0.24	-0.13	-0.02	0.22	0.53
GPS_alt	-54.60	82.80	140.60	145.00	212.10	333.50
GPS_hdg	0.00	105.20	175.90	184.00	275.80	360.00
GPS_spd	0.00	0.00	28.34	31.87	50.19	107.42
ME_AWS_fcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_hw_level	0.00	0.00	0.00	0.12	0.00	2.00
ME_AWS_hw_measurement	0.00	0.00	0.00	0.05	0.00	2.50
ME_AWS_hw_repeatable	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_hw_valid	0.00	0.00	0.00	0.12	0.00	1.00
ME_AWS_hmw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_left	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_off	0.00	1.00	1.00	0.92	1.00	1.00
ME_AWS_ldw_right	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pedestrian_dz	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_tamper	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_time_indicator	1.00	1.00	2.00	1.99	3.00	3.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
ME_AWS_tsr_level	0.00	0.00	0.00	0.15	0.00	7.00
ME_AWS_tsr_on	0.00	0.00	0.00	0.05	0.00	1.00
ME_AWS_zero_speed	0.00	0.00	1.00	0.71	1.00	1.00
tsr_1_speed	1.00	4.00	11.00	92.11	254.00	254.00
tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
avg_tsr_1_speed	72.52	83.43	88.74	92.19	99.00	120.39
rolling_tsr_1_speed	2.00	7.00	64.00	92.11	129.50	254.00
avg_tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
rolling__tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
tsr_2_speed	3.00	201.00	254.00	229.90	254.00	254.00
tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
avg_tsr_2_speed	224.80	227.10	229.90	229.90	231.90	238.60
rolling_tsr_2_speed	5.00	227.00	254.00	229.90	254.00	254.00
avg_tsr_2_sup	0.00	0.28	0.41	0.36	0.47	0.63
rolling_tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
tsr_3_speed	3.00	254.00	254.00	250.90	254.00	254.00
tsr_3_sup	0.00	0.00	0.00	0.10	0.00	20.00
avg_tsr_3_speed	248.10	249.50	251.10	250.90	252.70	254.00
rolling_tsr_3_speed	5.00	254.00	254.00	250.90	254.00	254.00
avg_tsr_3_sup	0.00	0.03	0.09	0.10	0.19	0.23
rolling__tsr_3_sup	0.00	0.00	0.00	0.10	0.00	13.00
tsr_4_speed	7.00	254.00	254.00	253.90	254.00	254.00
tsr_4_sup	0.00	0.00	0.00	0.01	0.00	20.00
avg_tsr_4_speed	252.80	254.00	254.00	253.90	254.00	254.00
rolling_tsr_4_speed	130.50	254.00	254.00	253.90	254.00	254.00
avg_tsr_4_sup	0.00	0.00	0.00	0.01	0.00	0.10
rolling__tsr_4_sup	0.00	0.00	0.00	0.01	0.00	10.00
Phase 2						
duration	974.00	974.00	2007.00	4074.00	2007.00	17041.00
iDreams_Headway_Map_level_.1_mean	-1.00	-1.00	-1.00	-0.97	-1.00	0.20

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
iDreams_Headway_Map_level_0_mean	-1.00	0.00	0.00	-0.01	0.00	0.20
iDreams_Headway_Map_level_1_mean	-1.00	0.00	0.00	0.00	0.00	0.20
iDreams_Headway_Map_level_2_mean	-0.98	0.00	0.00	0.00	0.00	0.20
iDreams_Headway_Map_level_3_mean	-0.94	0.00	0.00	0.00	0.00	0.17
Headway_level_initial	-1.00	-1.00	-1.00	-0.98	-1.00	3.00
Headway_level	-1.00	-1.00	-1.00	-0.98	-1.00	0.20
Headway_avg_level	-1.00	-1.00	-1.00	-0.98	-1.00	3.00
iDreams_Speeding_Map_level_0_mean	0.00	0.00	0.00	0.02	0.00	1.44
iDreams_Speeding_Map_level_1_mean	0.00	0.00	0.00	0.01	0.00	1.44
iDreams_Speeding_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	1.27
iDreams_Speeding_Map_level_3_mean	0.00	0.00	0.00	0.00	0.00	1.44
Speeding_level_Initial	0.00	0.00	0.00	0.04	0.00	3.00
Speeding_level	0.00	0.00	0.00	0.04	0.01	1.44
Speeding_avg_level	0.00	0.00	0.00	0.04	0.00	3.00
iDreams_Overtaking_Map_level_0_mean	0.00	0.00	0.00	0.00	0.00	0.20
iDreams_Overtaking_Map_level_1_mean	0.00	0.00	0.00	0.00	0.00	0.20
iDreams_Overtaking_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	0.15
Ovetaking_level_initial	0.00	0.00	0.00	0.00	0.00	2.00
Overtaking_level	0.00	0.00	0.00	0.00	0.00	0.20
Overtaking_avg_level	0.00	0.00	0.00	0.00	0.00	1.00
iDreams_Fatigue_Map_level_0_mean	0.00	0.00	0.00	0.02	0.00	2.84
iDreams_Fatigue_Map_level_1_mean	0.00	0.00	0.00	0.01	0.00	2.84
iDreams_Fatigue_Map_level_2_mean	0.00	0.00	0.00	0.18	0.00	2.84
iDreams_Fatigue_Map_level_3_mean	0.00	0.00	0.00	0.94	3.00	3.00
Fatigue_level_initial	0.00	0.00	0.00	1.15	3.00	3.00
Fatique_level	0.00	0.00	0.06	1.15	3.00	3.00
Fatigue_avg_level	0.00	0.00	0.00	1.15	3.00	3.00
DrivingEvents_Map_evt_ha_mean	0.00	0.00	0.00	0.03	0.00	0.44
DrivingEvents_Map_evt_hb_mean	-0.89	0.00	0.00	-0.05	0.00	0.00
Driving_events_maxg	-0.89	-0.23	-0.12	-0.01	0.21	0.54

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
GPS_alt	-3.50	101.90	170.40	163.40	227.30	351.40
GPS_hdg	0.00	102.80	185.30	186.30	275.50	360.00
GPS_spd	0.00	0.00	24.08	31.71	49.82	145.48
ME_AWS_fcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_hw_level	0.00	0.00	0.00	0.13	0.00	2.00
ME_AWS_hw_measurement	0.00	0.00	0.00	0.05	0.00	2.50
ME_AWS_hw_repeatable	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_hw_valid	0.00	0.00	0.00	0.12	0.00	1.00
ME_AWS_hmw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_left	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_off	0.00	1.00	1.00	0.94	1.00	1.00
ME_AWS_ldw_right	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pedestrian_dz	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_tamper	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_time_indicator	1.00	1.00	1.00	1.93	3.00	3.00
ME_AWS_tsr_level	0.00	0.00	0.00	0.13	0.00	7.00
ME_AWS_tsr_on	0.00	0.00	0.00	0.02	0.00	1.00
ME_AWS_zero_speed	0.00	0.00	1.00	0.75	1.00	1.00
tsr_1_speed	1.00	4.00	11.00	92.12	254.00	254.00
tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
avg_tsr_1_speed	72.52	83.43	88.74	92.32	99.00	120.39
rolling_tsr_1_speed	2.00	7.00	64.00	92.12	129.50	254.00
avg_tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
rolling__tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
tsr_2_speed	3.00	201.00	254.00	229.80	254.00	254.00
tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
avg_tsr_2_speed	224.80	227.10	229.00	229.80	231.90	238.60
rolling_tsr_2_speed	5.00	227.00	254.00	229.80	254.00	254.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
avg_tsr_2_sup	0.00	0.28	0.41	0.36	0.47	0.63
rolling_tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
tsr_3_speed	3.00	254.00	254.00	250.90	254.00	254.00
tsr_3_sup	0.00	0.00	0.00	0.10	0.00	20.00
avg_tsr_3_speed	248.10	249.50	251.10	250.90	252.70	254.00
rolling_tsr_3_speed	5.00	254.00	254.00	250.90	254.00	254.00
avg_tsr_3_sup	0.00	0.03	0.09	0.10	0.19	0.23
rolling__tsr_3_sup	0.00	0.00	0.00	0.10	0.00	13.00
tsr_4_speed	7.00	254.00	254.00	253.90	254.00	254.00
tsr_4_sup	0.00	0.00	0.00	0.01	0.00	20.00
avg_tsr_4_speed	252.80	254.00	254.00	253.90	254.00	254.00
rolling_tsr_4_speed	130.50	254.00	254.00	253.90	254.00	254.00
avg_tsr_4_sup	0.00	0.00	0.00	0.01	0.00	0.10
rolling__tsr_4_sup	0.00	0.00	0.00	0.01	0.00	10.00
Phase 3						
duration	115.00	11186.00	12617.00	15952.00	27908.00	27908.00
iDreams_Headway_Map_level_.1_mean	-1.00	-0.99	-0.99	-0.90	-0.95	0.08
iDreams_Headway_Map_level_0_mean	-0.99	0.00	0.00	-0.02	0.00	0.08
iDreams_Headway_Map_level_1_mean	-0.99	0.00	0.00	-0.01	0.00	0.08
iDreams_Headway_Map_level_2_mean	-0.99	0.00	0.00	0.00	0.00	0.08
iDreams_Headway_Map_level_3_mean	-0.99	0.00	0.00	0.00	0.00	0.08
Headway_level_initial	-1.00	-1.00	-1.00	-0.93	-1.00	3.00
Headway_level	-1.00	-0.99	-0.99	-0.93	-0.96	0.08
Headway_avg_level	-1.00	-1.00	-1.00	-0.93	-1.00	3.00
iDreams_Speeding_Map_level_0_mean	0.00	0.01	0.02	0.07	0.05	1.00
iDreams_Speeding_Map_level_1_mean	0.00	0.00	0.00	0.03	0.00	1.00
iDreams_Speeding_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	1.00
iDreams_Speeding_Map_level_3_mean	0.00	0.00	0.00	0.00	0.00	1.00
Speeding_level_Initial	0.00	0.00	0.00	0.10	0.00	3.00
Speeding_level	0.00	0.01	0.02	0.10	0.15	1.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
Speeding_avg_level	0.00	0.00	0.00	0.10	0.00	3.00
iDreams_Overtaking_Map_level_0_mean	0.00	0.00	0.00	0.01	0.01	0.26
iDreams_Overtaking_Map_level_1_mean	0.00	0.00	0.00	0.00	0.00	0.26
iDreams_Overtaking_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	0.00
Ovetaking_level_initial	0.00	0.00	0.00	0.01	0.00	3.00
Overtaking_level	0.00	0.00	0.00	0.01	0.01	0.26
Overtaking_avg_level	0.00	0.00	0.00	0.01	0.00	1.50
iDreams_Fatigue_Map_level_0_mean	0.00	0.00	0.00	0.11	0.00	2.96
iDreams_Fatigue_Map_level_1_mean	0.00	0.00	0.00	0.13	0.00	2.99
iDreams_Fatigue_Map_level_2_mean	0.00	0.00	0.00	0.64	1.84	2.99
iDreams_Fatigue_Map_level_3_mean	0.00	0.00	0.00	1.09	2.96	3.00
Fatigue_level_initial	0.00	1.00	2.00	1.99	3.00	3.00
Fatique_level	0.00	1.46	1.95	1.97	2.96	3.00
Fatigue_avg_level	0.00	1.00	2.00	1.99	3.00	3.00
DrivingEvents_Map_evt_ha_mean	0.00	0.00	0.00	0.03	0.00	0.44
DrivingEvents_Map_evt_hb_mean	-0.65	0.00	0.00	-0.05	0.00	0.00
Driving_events_maxg	-0.65	-0.24	-0.12	-0.01	0.22	0.53
GPS_alt	-39.60	94.60	155.00	154.30	215.70	350.60
GPS_hdg	0.00	110.00	182.30	188.30	280.90	360.00
GPS_spd	0.00	0.00	26.67	31.30	50.37	107.05
ME_AWS_fcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_hw_level	0.00	0.00	0.00	0.14	0.00	2.00
ME_AWS_hw_measurement	0.00	0.00	0.00	0.06	0.00	2.50
ME_AWS_hw_repeatable	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_hw_valid	0.00	0.00	0.00	0.14	0.00	1.00
ME_AWS_hmw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_left	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_off	0.00	1.00	1.00	0.92	1.00	1.00
ME_AWS_ldw_right	0.00	0.00	0.00	0.00	0.00	1.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
ME_AWS_pcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pedestrian_dz	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_tamper	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_time_indicator	1.00	1.00	3.00	2.08	3.00	3.00
ME_AWS_tsr_level	0.00	0.00	0.00	0.16	0.00	7.00
ME_AWS_tsr_on	0.00	0.00	0.00	0.02	0.00	1.00
ME_AWS_zero_speed	0.00	0.00	1.00	0.71	1.00	1.00
tsr_1_speed	1.00	4.00	11.00	91.69	254.00	254.00
tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
avg_tsr_1_speed	72.52	83.43	88.03	91.95	99.00	120.39
rolling_tsr_1_speed	2.00	6.50	64.00	91.71	129.50	254.00
avg_tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
rolling__tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
tsr_2_speed	3.00	201.00	254.00	229.90	254.00	254.00
tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
avg_tsr_2_speed	224.80	227.10	229.00	229.80	231.90	238.60
rolling_tsr_2_speed	5.00	227.00	254.00	229.90	254.00	254.00
avg_tsr_2_sup	0.00	0.28	0.41	0.36	0.47	0.63
rolling_tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
tsr_3_speed	3.00	254.00	254.00	250.90	254.00	254.00
tsr_3_sup	0.00	0.00	0.00	0.10	0.00	20.00
avg_tsr_3_speed	248.10	249.50	251.10	250.90	252.70	254.00
rolling_tsr_3_speed	5.00	254.00	254.00	250.90	254.00	254.00
avg_tsr_3_sup	0.00	0.03	0.09	0.10	0.19	0.23
rolling__tsr_3_sup	0.00	0.00	0.00	0.10	0.00	13.00
tsr_4_speed	7.00	254.00	254.00	253.90	254.00	254.00
tsr_4_sup	0.00	0.00	0.00	0.01	0.00	20.00
avg_tsr_4_speed	252.80	254.00	254.00	253.90	254.00	254.00
rolling_tsr_4_speed	130.50	254.00	254.00	253.90	254.00	254.00
avg_tsr_4_sup	0.00	0.00	0.00	0.01	0.00	0.10

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
rolling__tsr_4_sup	0.00	0.00	0.00	0.01	0.00	10.00
Phase 4						
duration	638.00	12094.00	13905.00	12552.00	14853.00	22958.00
iDreams_Headway_Map_level_.1_mean	-1.00	-0.97	-0.93	-0.74	-0.72	0.17
iDreams_Headway_Map_level_0_mean	-0.97	0.00	0.00	-0.04	0.00	0.17
iDreams_Headway_Map_level_1_mean	-0.97	0.00	0.00	-0.01	0.00	0.17
iDreams_Headway_Map_level_2_mean	-0.97	0.00	0.00	0.00	0.00	0.17
iDreams_Headway_Map_level_3_mean	-0.97	0.00	0.00	0.00	0.00	0.17
Headway_level_initial	-1.00	-1.00	-1.00	-0.80	-1.00	3.00
Headway_level	-1.00	-0.97	-0.93	-0.80	-0.82	0.17
Headway_avg_level	-1.00	-1.00	-1.00	-0.80	-1.00	3.00
iDreams_Speeding_Map_level_0_mean	0.00	0.00	0.08	0.13	0.14	1.05
iDreams_Speeding_Map_level_1_mean	0.00	0.00	0.00	0.09	0.00	1.05
iDreams_Speeding_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	1.05
iDreams_Speeding_Map_level_3_mean	0.00	0.00	0.00	0.01	0.00	1.05
Speeding_level_Initial	0.00	0.00	0.00	0.23	0.00	3.00
Speeding_level	0.00	0.06	0.12	0.23	0.29	1.05
Speeding_avg_level	0.00	0.00	0.00	0.23	0.00	3.00
iDreams_Overtaking_Map_level_0_mean	0.00	0.00	0.01	0.02	0.01	0.26
iDreams_Overtaking_Map_level_1_mean	0.00	0.00	0.00	0.00	0.00	0.26
iDreams_Overtaking_Map_level_2_mean	0.00	0.00	0.00	0.00	0.00	0.13
Ovetaking_level_initial	0.00	0.00	0.00	0.02	0.00	3.00
Overtaking_level	0.00	0.00	0.01	0.02	0.02	0.26
Overtaking_avg_level	0.00	0.00	0.00	0.02	0.00	1.50
iDreams_Fatigue_Map_level_0_mean	0.00	0.00	0.00	0.16	0.00	3.00
iDreams_Fatigue_Map_level_1_mean	0.00	0.00	0.00	0.20	0.00	2.95
iDreams_Fatigue_Map_level_2_mean	0.00	0.00	0.00	0.55	1.60	3.00
iDreams_Fatigue_Map_level_3_mean	0.00	0.00	0.00	0.85	2.59	3.00
Fatigue_level_initial	0.00	1.00	2.00	1.74	3.00	3.00
Fatique_level	0.00	1.00	1.76	1.76	2.64	3.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
Fatigue_avg_level	0.00	1.00	2.00	1.74	3.00	3.00
DrivingEvents_Map_evt_ha_mean	0.00	0.00	0.00	0.03	0.00	0.42
DrivingEvents_Map_evt_hb_mean	-0.53	0.00	0.00	-0.05	0.00	0.00
Driving_events_maxg	-0.53	-0.23	0.12	0.00	0.22	0.54
GPS_alt	1.50	86.30	134.40	138.20	193.60	338.70
GPS_hdg	0.00	97.67	173.10	179.38	270.00	360.00
GPS_spd	0.00	2.96	30.56	32.09	48.89	107.79
ME_AWS_fcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_hw_level	0.00	0.00	0.00	0.11	0.00	2.00
ME_AWS_hw_measurement	0.00	0.00	0.00	0.04	0.00	2.50
ME_AWS_hw_repeatable	0.00	0.00	0.00	0.00	0.00	0.00
ME_AWS_hw_valid	0.00	0.00	0.00	0.11	0.00	1.00
ME_AWS_hmw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_left	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_ldw_off	0.00	1.00	1.00	0.94	1.00	1.00
ME_AWS_ldw_right	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pcw	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_pedestrian_dz	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_tamper	0.00	0.00	0.00	0.00	0.00	1.00
ME_AWS_time_indicator	1.00	1.00	1.00	1.79	3.00	3.00
ME_AWS_tsr_level	0.00	0.00	0.00	0.12	0.00	7.00
ME_AWS_tsr_on	0.00	0.00	0.00	0.05	0.00	1.00
ME_AWS_zero_speed	0.00	1.00	1.00	0.77	1.00	1.00
tsr_1_speed	1.00	4.00	11.00	92.60	254.00	254.00
tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
avg_tsr_1_speed	72.52	83.43	88.74	92.24	99.00	120.39
rolling_tsr_1_speed	2.00	7.00	64.00	92.60	129.50	254.00
avg_tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00
rolling__tsr_1_sup	0.00	0.00	0.00	0.00	0.00	0.00

Variables	Min	1st quartile	Median	Mean	3rd quartile	Max
tsr_2_speed	3.00	201.00	254.00	229.90	254.00	254.00
tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
avg_tsr_2_speed	224.80	227.10	229.90	229.90	231.90	238.60
rolling_tsr_2_speed	5.00	227.00	254.00	229.90	254.00	254.00
avg_tsr_2_sup	0.00	0.28	0.41	0.36	0.47	0.63
rolling_tsr_2_sup	0.00	0.00	0.00	0.36	0.00	20.00
tsr_3_speed	3.00	254.00	254.00	250.80	254.00	254.00
tsr_3_sup	0.00	0.00	0.00	0.10	0.00	20.00
avg_tsr_3_speed	248.10	249.50	251.10	250.80	252.70	254.00
rolling_tsr_3_speed	5.00	254.00	254.00	250.80	254.00	254.00
avg_tsr_3_sup	0.00	0.03	0.09	0.10	0.19	0.23
rolling__tsr_3_sup	0.00	0.00	0.00	0.10	0.00	13.00
tsr_4_speed	7.00	254.00	254.00	253.90	254.00	254.00
tsr_4_sup	0.00	0.00	0.00	0.01	0.00	20.00
avg_tsr_4_speed	252.80	254.00	254.00	253.90	254.00	254.00
rolling_tsr_4_speed	130.50	254.00	254.00	253.90	254.00	254.00
avg_tsr_4_sup	0.00	0.00	0.00	0.01	0.00	0.10
rolling__tsr_4_sup	0.00	0.00	0.00	0.01	0.00	10.00