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

D7.2

Effectiveness Evaluation of the Interventions

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	**D7.2**
Deliverable Title	Effectiveness Evaluation of the Interventions
Work Package	Work Package 7: Evaluation of Safety Interventions
Contractual due date	28/04/23
Actual submission date	27/04/23
Nature	Report
Dissemination level	Public
Lead Beneficiary	LOUGH
Responsible Author	Laurie Brown
Contributions from	Laurie Brown, Rachel Talbot, Sally Maynard, Evita Papazikou, Fran Pilkington-Cheney, Graham Hancox, Ashleigh Filtness (LOUGH) Muhammad Wisal Khattak, Muhammad Adnan, Veerle Ross, Ariane Cuenen, Geert Wets, Tom Brijs, Kris Brijs (UHasselt) Virginia Petraki, Eva Michelaraki, Stella Roussou, Christos Katrakazas, George Yannis (NTUA) Chiara Gruden (UM)

Please refer to the document as:

Brown, L., Talbot., R., Khattak, M. W., Petraki, V., Gruden, C. et al. (2023). *Effectiveness Evaluation of the Interventions.* Deliverable 7.2 of the EC H2020 project i-DREAMS

Revision history (including peer review & quality control)

Version	Issue date	% Complete	Changes	Contributor(s)
v0.1	06/12/22	0	Initial deliverable structure	Laurie Brown, Rachel Talbot
V0.2	10/03/23	95%	Final draft deliverable for review	See ‘contributions from’ above
V1.4	27/04/23	100%	Final version	See ‘contributions from’ above

Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the document is believed to be accurate, the authors(s) or any other participant in the *i-DREAMS* consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the *i-DREAMS* Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the i-DREAMS Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Copyright

© *i-DREAMS* Consortium, 2019-2022. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

Table of Contents

List of Figures

©i-DREAMS, 2019

List of Tables

Glossary and Abbreviations

Word / Abbreviation	Description
ADAS	Advanced Driver Assistance System
BE	Belgium
DE	Germany
DSD	Drivers Safety Device
EL	Greece
FCA	Forward Collision Avoidance
GLMM	Generalized Linear Mixed Model
GPS	Global Positioning System
IBI	Inter-Beat Interval
KPH	Kilometres per Hour
KSS	Karolinska Sleepiness Scale
PAN	Pantograph
PO	Performance Objective
PPG	Photoplethysmography
PT	Portugal
RS	Road Sharing (events)
SPAD	Signal Passed At Danger
SPD	Speeding (events)
SPG	Safety Promoting Goal
SSST	Serial Sevens Subtraction Task
STZ	Safety Tolerance Zone
TBC	Traction Brake Controller
TSR	Temporary Speed Restrictions
UK	United Kingdom
VC	Vehicle Control (events)
VRU	Vulnerable Road User

Executive Summary

The overall objective of the *i*-DREAMS project is to set up 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 was made to monitor and determine if a driver is within acceptable boundaries of safe operation (i.e., Safety Tolerance Zone). Moreover, the *i*-DREAMS platform offers 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 while driving.

This deliverable focusses on evaluating the effectiveness of the i-DREAMS interventions in improving drivers' safety outcomes. The work here will evaluate the impact of the real-time driver interventions, post-trip driver feedback, and gamification interventions, in order to assess their impact on driving behaviour and driver state. Comparisons will be made between the different countries for which data are available, between the different interventions, and between the different outcome variables.

The data collected in on-road field trials are analysed for private drivers (passenger cars) and professional drivers (trucks and busses). The analysis of the interventions is formed of two main areas: outcome evaluation and process evaluation. Outcome evaluation, also known as effect evaluation, measures the effectiveness of the intervention. More specifically, it assesses whether the targeted factors of the on-road trials changed as a result of the intervention or not. The outcome evaluation of the on-road trials will examine whether the i-DREAMS interventions influenced the following four areas: safety outcomes, safety promoting goals, performance objectives, and change objectives. These four areas are part of the logic model of change behind the i-DREAMS interventions. Process evaluation assesses which parts of the intervention were implemented as intended, and which were not.

For private drivers, the results show that there was a statistically significant decrease in events from Phase 1 to Phase 4. This suggests that the i-DREAMS system had a positive impact on the measured safety outcomes and succeeded in keeping drivers in the first level of the STZ for more of their journey. When individual phase changes are considered, the most significant results were seen from Phase 3 to Phase 4. This suggest that the addition of the gamification elements had a significant impact on safety outcomes, and further supports the conclusion that the full system provides the most effective results. However, differences were found when each country was analysed individually, which were statistically significant, though there is not a clear reason why this would be so. Furthermore, differences were also found between drivers within countries. In each country, between two thirds to three quarters of drivers showed improved outcomes (i.e., a reduction in events), but the remainder had worse outcomes (an increase in events). It's not clear from the data why some individuals responded positively to the technology and others did not, and further work is needed to understand why the system has such varied effects on different drivers.

For all countries, drivers engaged more with the app in Phase 4 of the trial compared with Phase 3, after the introduction of the gamification features. Although the 'trips' and 'scores' menu were the functions most used by drivers (functions that were available in both phases), the data suggests that the gamification functions were more engaging and held attention more consistently. The generic information in the app (hints, tips etc.) was less appealing to users. They found more interest in personalised feedback such as their trip information, goals, and position on the leader board. The data also suggests a link between app usage and performance outcome; nearly all the drivers who used the app heavily showed improved

outcomes. It would be interesting to investigate this further to determine whether there is a causal effect between these results.

Generally, the i-DREAMS system showed less positive impact with professional drivers compared to private drivers. Specifically, a lower proportion of the professional drivers showed improved outcomes, and little significant change was seen in terms of safety outcomes. Where there were significant results, these were most often increases in events, i.e., worse outcome. Again, it is not obvious why this result is observed. The only statistically significant improved outcome was for truck drivers, which was for ‘total’ high severity events specifically between Phase 1 and Phase 2. Therefore, it can tentatively be concluded that the system had a positive impact on the most severe events. Process evaluation results were only available for Truck drivers, but showed similar results to private drivers, with more app engagement in Phase 4 compared to Phase 3, after the introduction of gamification features. This further supports the value of gamification features.

The intention was to use the results to inform the ranking of interventions and provide an assessment of which intervention schemes are most effective. However, given the varied results between countries and transport modes, it is difficult to conclude a definitive ranking of the different interventions. The results indicate that the full system (real-time warnings plus app feedback plus gamification features in the app) provides the most significant positive impact on driver outcome. For private drivers, the analysis showed that most significant positive change was seen in Phase 4 of the trial, i.e., the gamification features, however it cannot be said that those alone were the most effective, as they were tested in combination with the other interventions. However, the data does suggest that app feedback on its own is less effective than when the app also includes gamification features. For truck drivers, we can tentatively conclude that the real-time interventions had the most impact, however more data is needed to support this.

Lastly, the rail mode was included in i-DREAMS to broaden the application of the i-DREAMS platform which was originally designed for use in road vehicles. Trams operate within a mixed-traffic environment, driving on both segregated track, and shared, multi-user road. Therefore, aspects of the i-DREAMS platform can be applied to trams and may be beneficial to tram driving safety and risk mitigation. Two main studies were carried out to assess the use of the i-DREAMS platform in trams. The first was a simulator study to test the real-time element of the platform and the second was a focus group study to assess the potential use of the post-trip feedback app in the tram context.

The tram simulator study suggests that the i-DREAMS system and associated warnings offer several benefits for tram driving operations. Firstly, as instances of speeding are rare, the speed alert would be more helpful as a warning before the occurrence of speeding, alerting the drivers they are approaching the limit, or more effective as a constant in-cab reminder of the current speed limit. The concept of a vulnerable road user (VRU) warning could be beneficial to tram drivers operating in mixed traffic environments, however, it was clear that the VRU warning needs to be developed to take into account specific aspects of tram driving and there is a concern about it being triggered too often. The fatigue warning could also potentially be beneficial as a warning before the existing fatigue monitoring device alerts, as a prompt to drivers to consider their alertness or take a break. Tram drivers suggested that the app would be most useful in identifying issues that were common to drivers and as a self-evaluation tool. They were more sceptical about the gamification elements, in particular the leader board, and expressed views that competition could have a negative impact on safety and is therefore not desired.

1 Introduction

1.1 About the project

The overall objective of the *i*-DREAMS project is to set up 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 was made to monitor and determine if a driver is within acceptable boundaries of safe operation (i.e., Safety Tolerance Zone). Moreover, the *i*-DREAMS platform offers 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 while driving. The safety-oriented interventions were 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 acquisition of safer driving habits/behaviours. Consequently, the *i*-DREAMS platform allows 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 intended 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 to the present circumstances. A more intrusive warning signal is 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 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 so-called gamification features intended 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 with a total of 600 participants representing car, bus, truck and tram/train drivers.

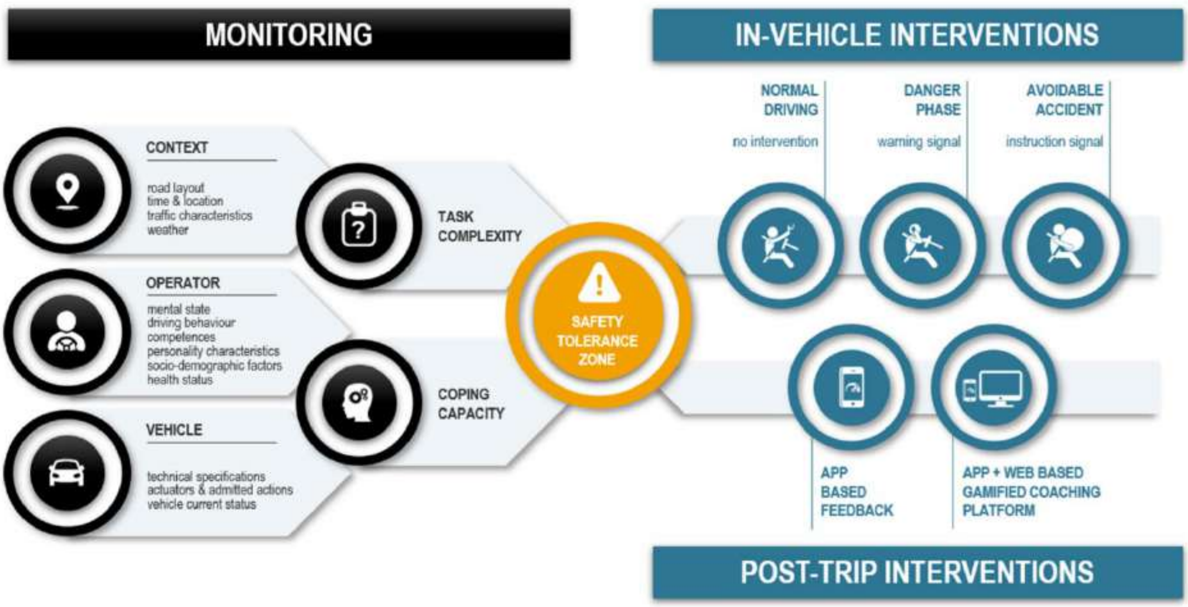


Figure 1: Conceptual framework of the i-DREAMS platform.

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 user-license Human Factors database with anonymized data from the simulator and field experiments will be developed.

1.2 About this report

This deliverable focusses on evaluating the effectiveness of the i-DREAMS interventions in improving drivers' safety outcomes, i.e., the right half of the conceptual framework shown in Figure 1. Results for analysis of driver monitoring (the left half of the conceptual framework) can be found in Deliverable 6.1 (task complexity) (Papazikou et al., 2023), Deliverable 6.2 (coping capacity) (Michelaraki et al., 2023) and Deliverable 6.3 (synthesis of task complexity and coping capacity) (Michelaraki et al., 2023b).

The work here will evaluate the impact of the real-time driver interventions, post-trip driver feedback, and gamification interventions, in order to assess their impact on driving behaviour and driver state. Comparisons will be made between the different countries for which data are available, between the different interventions, and between the different outcome variables. The report will aim to provide a complete assessment and ranking of the different interventions, to identify the most promising intervention schemes for improving driver behaviour.

The structure of this deliverable is presented as follows: Section 2 gives an overview of the framework underpinning the i-DREAMS system and the protocols for on-road data collection, as well as describing the analysis methods used for this deliverable. Section 3 presents the results for car drivers, Section 4 then gives the results for truck drivers, Section 5 presents the results for bus drivers, and Section 6 gives the methods and results for rail studies. Finally, Section 7 discussed the results and offers conclusions and recommendations.

2 Methodology

Section 2 gives an overview of the methodology used, specifically as it relates to the data collection and interpretation of results. The intention is not to go into detail regarding the specific project methodologies as these have been described in previous i-DREAMS deliverables, which will be referenced here where appropriate.

2.1 Theoretical Framework

The **Safety Tolerance Zone** (STZ) is a concept generated in the i-DREAMS project to guide the process of developing the i-DREAMS platform. The STZ as a theoretical concept originates from Fuller's Task Capability Interface model (Fuller 2000, 2005, 2011). In brief, this model states that for the driver to be fully in control of the vehicle and operate it safely, their capability (referred to here as coping capacity) has to be balanced with the task demand (referred to here as task complexity). See i-DREAMS Deliverable 3.1 (Talbot et al., 2020) for further detail.

The STZ includes three different driving phases: 'normal', 'danger' and 'avoidable crash'. As set out in Deliverable 3.2 (Katrakazas et al., 2020a), the **normal driving phase** represents the conditions in which a crash is unlikely to occur, i.e., the crash risk is low. During this phase, the driver can successfully adapt their behaviour to meet the task demand thus achieving a balance between coping capacity and task complexity. The **danger phase** is characterised by changes in normal driving that indicate that a crash may occur, therefore, the crash risk is increased. Finally, the **avoidable accident phase** occurs when a collision scenario develops but there is still time for the driver to intervene and avoid the crash. The need for action is more urgent than in the danger phase and if the driver does not adapt their behaviour to the current circumstances, a crash is very likely to occur.

The fundamental goal of the i-DREAMS platform is to keep the driver in the normal driving phase for as long as possible and, where this is not possible, to prevent the transition from the danger to the avoidable accident phase. To this end, the system combines both **real-time** and **post-trip interventions** which, respectively, aim to nudge and coach the driver. These interventions aim to improve the outcomes proposed in the **logic model of change** (Figure 2).

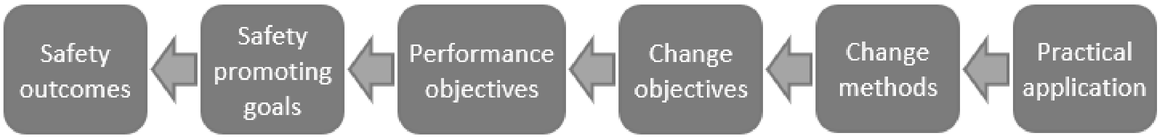


Figure 2: Structural overview of the compartments inside the operational toolbox for the i-DREAMS interventions

For the evaluation of interventions, we focus on the four highest levels of the logic model of change. The highest level targeted by the interventions is the **safety outcomes**, such as the likelihood of crash occurrence (e.g., forward crashes and rear-to-end crashes). The second-highest level is the **safety promoting goals** (SPGs). These are the behaviours that need to change for the safety outcomes to be realised. The second-lowest level is the **performance objectives** (POs), these are the more specific actions or behavioural parameters that need to change for the safety promoting goals to be achievable. The lowest level is the **change objectives.** These underlying behavioural determinants need to change for the performance objectives to become realisable. For a detailed description, see Deliverable 3.3 (Brijs et al., 2020).

The i-DREAMS platform is a warning-based driver assistance system, and it does not actively intervene physically in any way with the driving task. To estimate in which STZ phase the driver is, and which interventions should be provided, the i-DREAMS platform uses two modules. Firstly, it uses the **monitoring module**, which takes measurements related to the context, the operator, and the vehicle, to derive the demands of the driving task and the driver's ability to

©i-DREAMS, 2019

cope with these demands. This module estimates at which stage of the STZ the driver is operating at any given time. More specifically, the monitoring module registers driving behaviour related to a list of performance objectives as shown in Table 1. Secondly, the in-vehicle **intervention module** is responsible for keeping the driver within the normal phase of the STZ, either by providing a warning or alert during the trip (real-time intervention) or by providing feedback about the journey after completion of the driving task (post-trip intervention). In case of real-time interventions, a different type of in-vehicle warning is being delivered to the driver depending on the severity of the detected event.

Table 1: Mapping of safety promoting goals and performance objectives.

Safety Promoting Goal	Performance Objectives	Drivers informed via
Vehicle Control	Acceleration	Post-trip feedback only
	Deceleration	
	Steering	
Speed Management	Speeding	Real-time warnings and post-trip feedback
Sharing the Road with Others	Tailgating	Real-time warnings and post-trip feedback
	Lane departure	
	Forward collision avoidance	
	Vulnerable road user collision avoidance	
	Illegal overtaking	
Driver Fitness	Fatigue	Real-time warnings and post-trip feedback
	Distraction (hand-held phone use only)	

Therefore, the research questions that aim to be answered by this Deliverable can be summarised as:

- **SPG1**: Performance in terms of vehicle control (expressed as a numerical score) will significantly improve for vehicles equipped with and exposed to the i-DREAMS interventions.
- **SPG2**: Performance in terms of sharing the road with others (expressed as a numerical score) will significantly improve for vehicles equipped with and exposed to the i-DREAMS interventions.
- **SPG3**: Performance in terms of speed management (expressed as a numerical score) will significantly improve for vehicles equipped with and exposed to the i-DREAMS interventions.
- **SPG4**: Performance in terms of driving under conditions where one is fit enough (expressed as a numerical score) will significantly improve for vehicles equipped with and exposed to the i-DREAMS interventions.

2.2 Data Collection for On-Road Field Trials

This section provides a summary of the on-road field trials data collection, with particular focus on elements of the protocol that are relevant to understanding the analysis results. For more detailed information on the trial design, readers may refer to previous i-DREAMS deliverables: D3.4 (Pilkington-Cheney et al., 2020), D5.1 (Hancox et al., 2020), and D5.3 (Hancox et al., 2021). More detail on the in-vehicle technologies being used can further be found in D4.1 (Lourenço et al., 2021) and D4.4 (Lourenço et al., 2020).

Field trials were carried out in **five countries** (Belgium, Germany, Greece, Portugal, and the UK), and across **four transport modes** (cars, trucks, busses, and rail). The trial protocols were broadly the same for cars, trucks, and busses, and will be described here. Different methods were employed for rail; these are described in Section 6 alongside the rail results.

It is also noted here that all data collection partners followed the required ethical approval processes in their institution, and ethical approval was granted.

The on-road trials focussed on monitoring driving behaviour and the impact of real-time interventions (i.e., in-vehicle warnings) and post-trip interventions (i.e., post-trip-feedback & gamification) on driving behaviour. These interventions were assessed at different points of the trial, therefore the trial consisted of four phases:

- **Phase 1**: Baseline measurement.
- **Phase 2**: Introduction of real-time interventions via the in-vehicle display.
- **Phase 3**: Real-time intervention + introduction of post-trip feedback via the i-DREAMS smartphone app.
- **Phase 4**: Real-time intervention + post-trip feedback + introduction of gamification features in the smartphone app.

Phases 1, 2 and 3 each had a duration of four weeks. Phase 4 had an extended duration of six weeks.

Real-time warnings were given to drivers via the in-vehicle i-DREAMS display. A warning was triggered if drivers entered the ‘danger phase’ (STZ level 2) for a particular PO, and a more intrusive warning would further be triggered if they moved to the ‘avoidable crash phase’ (STZ level 3). Both visual and audible warnings were used; the complete set of possible warnings are described in the ‘Real-time interventions manual’ which can be found in Deliverable 5.3 (Hancox et al, 2021), and example is given below for ‘headway’ warnings. Real-time warnings were given for the POs related to ‘speeding’, ‘road sharing’ and ‘driver fitness’. There were no real-time warnings for ‘vehicle control’.

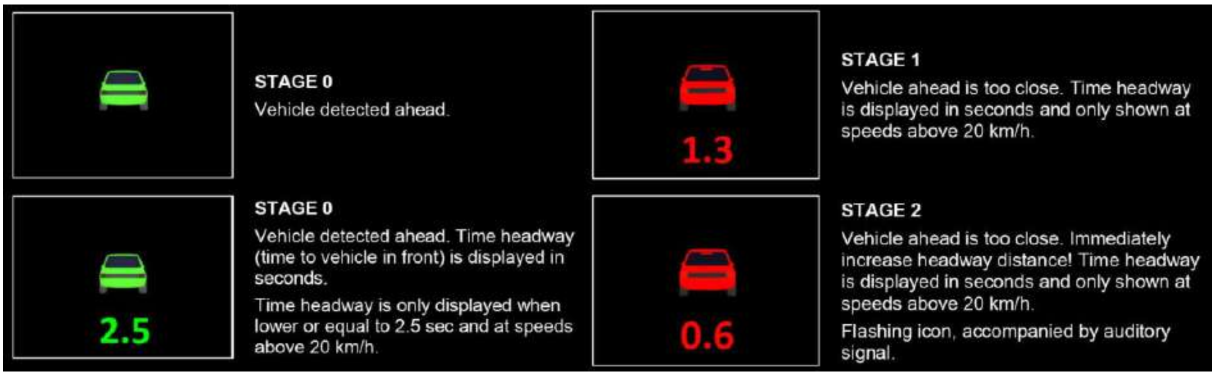


Figure 3: Example of in-vehicle warning icons - headway

Post-trip feedback was given to drivers via the i-DREAMS smartphone app. The app functions are described in detail in Deliverable 4.5 (Vanrompay et al., 2020), some example screenshots are provided below (Figure 4).

Briefly, the functions available to users during both Phase 3 and Phase 4 of the trial were:

- **‘Scores’** – scores calculated for each trip, for each PO and SPG (based on events registered). Drivers can also view aggregated scores over a chosen time period, and how their scores change over time.
- **‘Trips’** – drivers can see a list of all their trips with basic information and the overall scores for that trip. Within each trip, users can see a breakdown of the scores for each PO/SPG for that trip, and also can visualise the trip on a map, with markers placed where events were registered. More detail on these events can be viewed, as well as any video recordings (which are captured by the dashcam for some events).

©i-DREAMS, 2019

During Phase 4 of the trial, additional features became available as follows:

- **‘Info’ & ‘Tips’** – the app was programmed with helpful information and tips regarding safe driving behaviour (in relation to the SPGs), which users could browse to learn more. Push notifications would direct users to view particular information, to help increase engagement. Users also had the option to ‘like’ or ‘dislike’ information units.
- **‘Leader board’** – as part of the gamification element, a leader board was used, which would show all the drivers in a certain project (e.g., Belgian wave 1 drivers) ranked by their overall average score. Users had the option to select ‘daily’, ‘weekly’, ‘monthly’ or ‘lifetime’ leader boards, to see how they ranked against other drivers, and how their position changed each day/week etc.
- **‘Goals and badges’** – the second part of the gamification element was to offer drivers driving goals, which if completed would earn the user virtual badges. Goals were offered for multiple POs and would require a certain average score to be maintained over a certain distance of driving. If a goal was completed, a harder goal would then be offered, and this would repeat (i.e., a behaviour change technique called ‘scaffolding), allowing users to earn ‘bronze’, ‘silver’, ‘gold’, or ‘platinum’ badges.

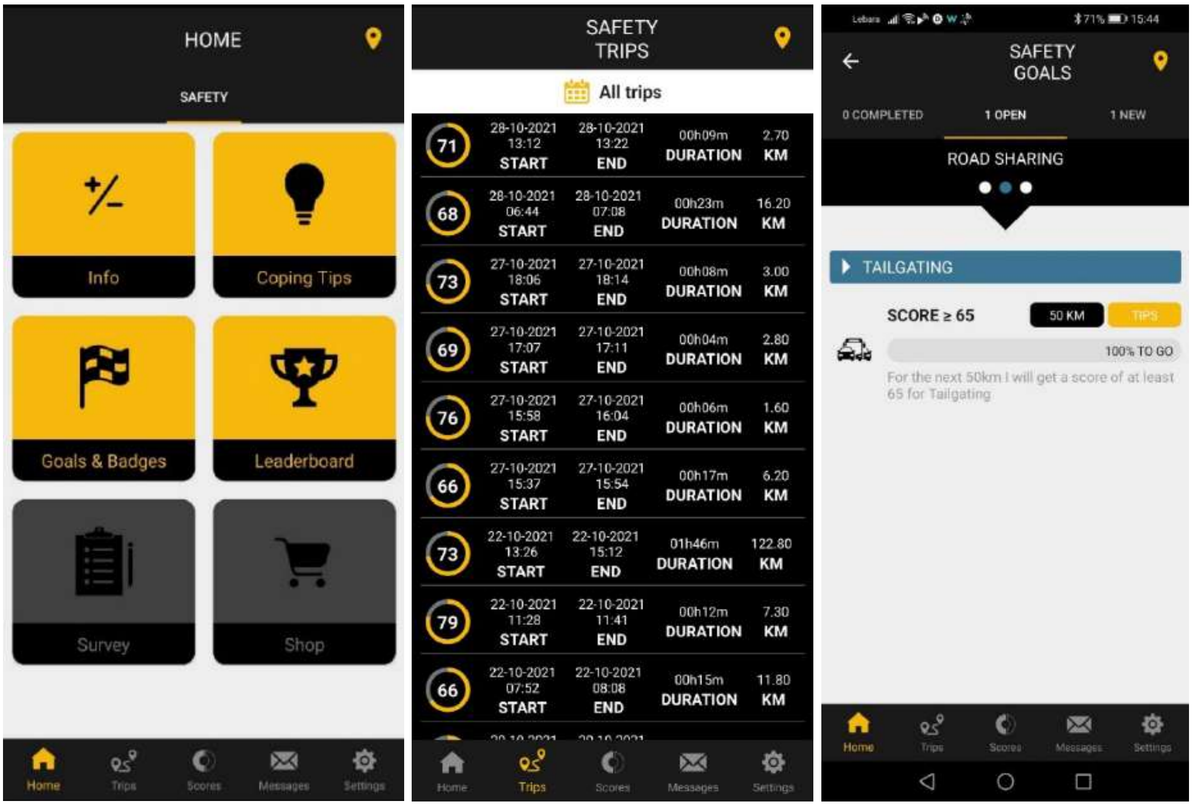


Figure 4: Example of i-DREAMS App functions - home, trips, goals

It should also be noted that, due to a limited number of equipment sets, it was not possible for all participants to take part at the same time. For some countries participants took part in two ‘waves’, with the first half of the participants having the equipment installed and experiencing the four trial phases, then that equipment being removed and installed for the second half of the participants. Where this occurred, both groups experienced the same interventions and phases – the intention was not to expose different participants to different trials, but simply to allow greatest amount of data to be collected with the available equipment. However, it is further noted that there was the option to make small adjustments to the trial design for the ‘wave 2’ participants, if the first wave identified any issues or areas for improvement. Only

Belgium and the UK had two distinct waves of participants, and generally the design remained the same, with only small software tweaks, e.g., to improve connectivity of the i-DREAMS wearable to the rest of the system.

Finally, in addition to the vehicle and app data, **questionnaire** data were collected both before and after the trial. The full questionnaires are given in Annex 2, and information collected pre-trial included:

- **Screening questionnaire**: driver details (age, gender, driving experience, employment status, etc.), vehicle details (make, 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 behaviours (e.g., speeding, using phone), accident and offence history, sleepiness and driving, medical conditions.

Information collected post-trial included:

- **User experience questionnaire**: opinions on the i-DREAMS system (ease of use, works as described), opinions on the i-DREAMS app (ease of use, usefulness).
- **Exit questionnaire**: opinions on the i-DREAMS system (did it improve their driving, usefulness, trust, clarity of warnings, etc.,), experience of driving situations, driver behaviour (driving and non-driving related behaviours), 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), to allow analysis of before and after responses. Following the reasoning behind the logic model of change adopted by the i-DREAMS interventions (Figure 2), these variables constituted the conceptual basis of the so-called **'change objectives'**. In other words, positive change in these socio-cognitive variables was assumed to generate positive change at the level of behavioural parameters monitored by the i-DREAMS sensors.

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., 2020b) and Deliverable 3.3 (Brijs et al., 2020). Briefly, **Self-efficacy** is to be understood as a person's judgment of his or her ability to cope effectively in different circumstances. **Attitude** stands for an individual's positive or negative assessment about performing a certain behaviour, in this case, using a new technology systemin a real-time or post-trip setting. **Personal norm** is when the motivation to perform a certain behaviour (or not) is dependent upon one's own personal value system. Before engaging in a particular behaviour, an individual will consider the potential consequences for his or her self-image. In case there is a perceived conflict with a set of deeply engrained moral values, anticipated regret will refrain a person from carrying out the behaviour. In the case of **subjective norm**, motivation is believed to be dependent on the extent to which a person complies (or not) with the opinion of important social referents (e.g. colleagues, friends, partner) about performing a particular behaviour.

2.2.1 Issues with data collection

As would be expected with any on-road trial, minor issues occasionally occurred that needed resolving. These typically only applied to an individual participant and were resolved without impacting the data. If a participant did experience issues to the extent that data were impacted, they would be excluded from relevant analyses. Issues that were more widespread, and affect the analysis, are detailed here.

- 'Driver fitness' data: (1) Distraction events, specifically hand-held phone use, were detected via the i-DREAMS app. If a driver was not able to install the app, or did not have their Bluetooth enabled, events were not detected. Furthermore, there was a period of ~3 months where a software issue with the app prevented events being registered, which impacted Belgium, Germany, and the UK. (2) Fatigue events were registered if the driver took a long trip without a break, or if their heart rate dropped to a certain level. For most drivers, heart rate was measured via a bracelet worn on their wrist, however drivers did not always remember to wear this, and furthermore it did not always connect to the rest of the system. For these reasons, ‘driver fitness’ results are presented separately from other SPGs.

- COVID-19: The trials were initially delayed due to the pandemic. Therefore, drivers who had been recruited had to wait some time before the trials actually began, and as a result some no longer wanted to take part. This made recruitment more difficult, particularly in Germany. Additionally, the pandemic resulted in a shortage of certain electrical components needed for the i-DREAMS system, which caused further delays in some countries, and meant that the initial planned number of participants was not reached by all countries. Varying restrictions across locations also impacted the traffic environment.
- Specifically for the Belgian drivers, they were most impacted by COVID-19. Belgium was the first country to start data collection, and their first wave of participants started the trial when some restrictions were still in operation. For some participants, these restrictions eased during the trial. Therefore, traffic density increased as more people returned to work and other activities, which would have presented a more complex environment than was experienced in their baseline phase. Furthermore, since Belgium was the first to start the trials, some of the early participants experienced delays within the trial as technology issues were identified and had to be resolved. These were subsequently reduced in countries which began data collection at a later date.

2.2.2 Variations between countries

Although all countries aimed to follow the same experimental protocol as set out in the WP5 deliverables, difficulties in practice meant some countries had to make adjustments.

- Belgium and the UK followed the standard protocols.
- For Germany, there was an issue with the installations that meant ‘road sharing’ data was not captured. Therefore, results are only presented for ‘total’, ‘vehicle control’, ‘speeding’, and ‘driver fitness’ data.
- For Greece, only the app was used to collect data due to difficulties with hardware installation, therefore only partial data were collected. As there were no real-time warnings, the data collection period was shortened and only included Phases 1, 3 & 4.
- For Portugal, only in-vehicle data were collected due to issues with app installation. Therefore, data are only analysed for Phases 1 and 2, for outcome evaluation only.

2.3 Analysis Methods

The analysis of the interventions is formed of two main areas: outcome evaluation and process evaluation.

Outcome evaluation, also known as effect evaluation, measures the effectiveness of the intervention. More specifically, it assesses whether the targeted factors of the on-road trials changed as a result of the intervention or not. The outcome evaluation of the on-road trials will examine whether the i-DREAMS interventions influenced the following four areas: safety outcomes, safety promoting goals, performance objectives, and change objectives. These four areas are part of the logic model of change behind the i-DREAMS interventions (Figure 2), which is described in more detail in D3.3 (Brijs et al., 2020). Ideally, we would like to detect a statistically significant impact on the safety outcomes (i.e., crash occurrence). However, this is not very likely to be detected during the on-road trials due to the rare nature of crashes and because the on-road trials have a total duration of only five months. It is more likely that the i-

DREAMS interventions will impact the underlying outcome variables (safety promoting goals, performance objectives, and change objectives), therefore analyses will be focussed on these.

Process evaluation assesses which parts of the intervention were implemented as intended, and which were not. More specifically, the quality of implementation and adoption of the intervention is investigated. The RE-AIM Framework variables (Glasgow et al., 1999) are the main focus of the process evaluation of the on-road trials. RE-AIM is a widely known framework for process evaluation and stands for: Reach, Effectiveness, Adaption, Implementation and Maintenance. User acceptance will also be a key component to be investigated in terms of process evaluation.

A more detailed description of the methodology informing the analyses can be found in D7.1 (Katrakazas et al., 2020b)

More practically, the outcome evaluation will be conducted by determining change in events and scores, both overall and for the different SPGs / POs. A 'positive' outcome is seen if the number of events <u>decreased</u>, and/or the score <u>increased</u>. Scores are derived from the events, with a driver starting a trip with a perfect score of 100 and losing points for events, with the number of points lost varying according to the type and severity of the event.

There are a few points to note regarding these analyses:

- For events analyses, 'events per 100km' were calculated and are used as opposed to raw numbers of events, so that a measure of exposure is taken into account and results are more comparable.
- Typically, events results will be presented for 'high severity', 'medium severity', and 'all' (medium + high) events. 'Medium' events correspond to the second level of the STZ, and 'high' events correspond to the third level.
- The total number of events is calculated as the sum of events for each PO, but excludes 'distraction' events, due to the issues described previously.
- The POs 'Distraction', 'Forward Collision Avoidance', 'VRU Collision Avoidance' and 'Lane departure' do not have different severity levels but are considered to be 'high' severity events for totals.
- Regarding scores, the 'overall' score is calculated as the average of the scores for 'vehicle control', 'speeding', 'road sharing' (for BE&UK), and 'driver fitness' – fatigue only (i.e., excluding distraction).
- For 'speeding' events, the data were additionally processed before analysing. It was found that the Mobileye system did not always identify speed limit signs correctly, therefore, some false positives and false negatives occurred. In the post-processing, GPS data from each trip was used to map-match and identify the correct speed limit, and then the recorded vehicle speed at each GPS point informed whether or not the driver was speeding at that time.
- For most POs, there will be more 'medium' events than 'high' events, as a driver must usually go through the 'danger' phase before reaching the 'avoidable accident phase'. The exception to this is 'speeding' events, as the data post-processing determined the most severe level reached for each instance of 'speeding', and therefore 'medium' events for 'speeding' are only seen if the driver did not then have a 'high' event.

It is also noted that Greece and Portugal are still collecting data at the time of writing, therefore results only include data that was available at the time of analysis.

For the outcome evaluation, the majority of analyses will be descriptive, and will focus on the change in events/scores across phases, and 'before and after' differences. However, statistical methods will also be used to determine if the changes observed are statistically significant. Specifically, repeated measures ANOVA will be used to determine if there are significant differences between the four data collection phases, and, where data are not normally distributed (as is the case with the majority of the data here), the Friedman test is used instead. As there is also value in determining any significant changes between individual phases, pairwise comparisons will also be carried out using either paired t-test or Wilcoxon signed rank test, depending on whether data are normally distributed or not. Finally, Generalized Linear Mixed Model (GLMM) analysis will be carried out on a dataset of combined countries data, to look for significant differences between phases and between countries.

For the process evaluation, again descriptive methods will be used. Firstly, data were collected each time a user visited the app, which will allow analysis of how often the app was used, and which functions were most visited. Secondly, use of technology and user acceptance will be analysed using questionnaire data.

2.3.1 Data cleaning

Before starting the analyses, data for each country were cleaned in a consistent manner. For outcome evaluation:

- Trips that were 'outside phase' were removed. During the trial, data collection partners recorded when each participant entered each phase. End dates could then be calculated (start date + 4 weeks for Phases 1, 2 and 3, and + 6 weeks for Phase 4), so that analyses only included trips within the set length of the phase. If a participant experienced delays moving from one phase to another, trips were only counted if they were in the correct date range, to avoid biasing results. This was mostly needed for Phase 4, as there was often a small delay between the participant finishing the trial and attending to have the equipment de-installed.
- If a driver experienced significant issues during the trial, e.g., repeated technology malfunction, they were not included in the analysis.
- If a driver did not have trip data in all four phases they were excluded, as it would not be possible to see behaviour change for each phase.
- For each driver, their average number of events and standard deviation were calculated (for all trips made by that driver, i.e., trips in all phases). Any trips that were outliers, defined as the mean +/- three standard deviations, were removed.
- Trips with a distance of less than 1km were removed.

For process evaluation:

- Drivers were excluded if they did not engage with the app at all.
- App visits 'outside phase' were removed, in the same way as for outcome evaluation.
- For total usage, 'visits' that occurred within 2 minutes of the previous one were removed, as this was thought to be a user engaging with different functions but within the same 'visit'.

Table 2: Data collection time period and participant numbers

	# Drivers Participated	Data Collection Start State	Data Collection End Date	# Drivers Excluded (Outcome)	# Drivers Excluded (Process)
BE Cars	52	20/04/2021	04/07/2022	4	3
DE Cars	29	25/01/2022	19/09/2022	4	6
UK Cars	54	04/10/2021	13/09/2022	5	3
EL Cars	80	01/01/2023	22/04/2023	24	9
BE Trucks	52	20/09/2021	25/10/2022	15	28
PT Busses	43	17/10/2022	26/07/2023	22	NA

3 Results – Cars

Section 3 presents the analysis results for car drivers. In the following analyses the sample size for each country may vary, as some drivers are excluded from certain analyses (for example if they had no trips in one or more data collection phases, missing questionnaire data, etc.). All analyses will state the sample used.

Car data were analysed for four countries: Belgium (BE), Germany (DE), Greece (EL) and the United Kingdom (UK). It is reiterated here that data were collected differently in Greece (as described in Section 2.2.2), where drivers used the app for post-trip feedback but had no in-vehicle real-time warnings. Therefore, results are not fully comparable with other countries data and care should be taken when interpreting them.

3.1 Data Sample

The full participant sample for each country is described below. All countries had a similar gender distribution; however, the German and Greek drivers were typically younger and had less driving experience than the Belgian and UK drivers. It is noted that these countries struggled more with recruitment, therefore could not sample for desired age distribution.

Table 3: Cars sample, per country

		BE	DE	EL	UK
Number of participants (drivers)		52	29	80	54
Participant gender	Male	32 (62%)	19 (66%)	48 (60%)	33 (61%)
	Female	18 (35%)	10 (34%)	32 (40%)	21 (39%)
	Unknown	2 (4%)	-	-	-
Mean age of participants (years)		47.3	32.2	31.5	45.4
Standard deviation of age (years)		17.9	9.6	10.1	13.6
Years' driving experience (range, avg.)		2 - 55, 27.6	1 - 35, 11.4	1 - 41, 10.7	2 - 60, 25.0

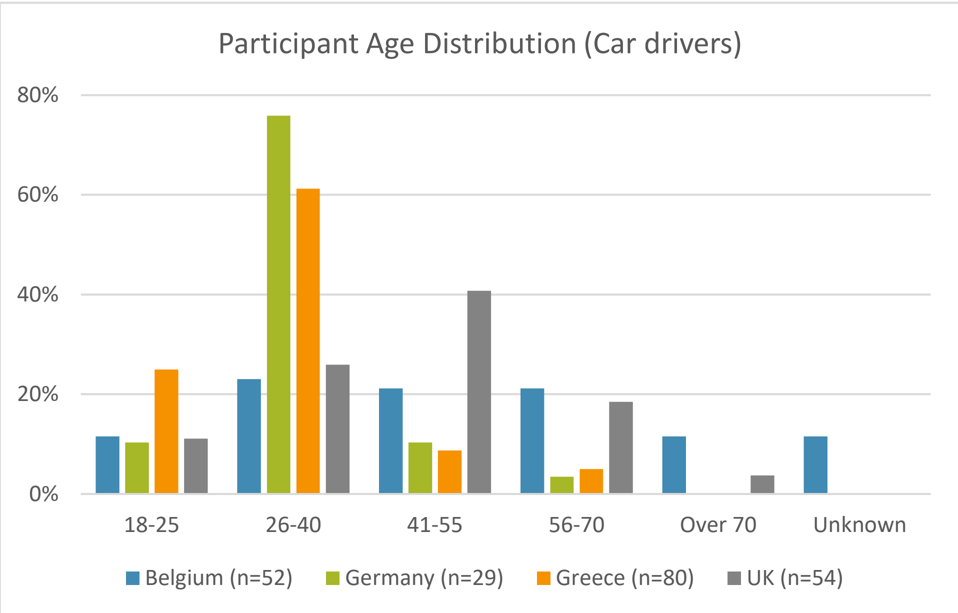


Figure 5: Participant age distribution, per country (car drivers)

Before taking part in the trial drivers were asked a number of questions, including about their current use of advanced driver assistance systems (ADAS), previous accidents and offences, and how they would describe their driving style and confidence. The responses to these are shown in Table 4 below, which highlights some differences between countries.

Table 4: Driver ADAS, accident and offence history, and confidence, per country (cars)

Question / Response Option		BE (n=50[1])	DE (n=29)	EL (n=80)	UK (n=54)
Which ADAS are present in your car? (Percentage replied equipped)	Automatic emergency braking	26%	31%	16%	2%
	Blind spot warning	10%	21%	5%	0%
	Drowsiness alert	14%	21%	9%	0%
	Forward collision warning	28%	34%	20%	9%
	High speed alert	28%	28%	14%	0%
	Lane keeping assistance	24%	28%	15%	0%
	Night vision & pedestrian detection	2%	7%	3%	0%
In the last three years, have you been involved in an accident with your car, which was self-inflicted?	No	90%	86%	83%	85%
	Yes, once	10%	7%	16%	13%
	Yes, twice	0%	7%	1%	2%
Within the last three years, have you been fined for a traffic offence while driving your car? (Excluding parking offences)	No	40%	72%	93%	80%
	Yes - not specified	12%	0%	0%	0%
	Yes - speeding	44%	24%	6%	19%
	Yes – DUI (intoxicated)	0%	0%	1%	0%
	Yes - running a red light	0%	0%	0%	2%
	Yes - multiple offences (speeding + phone offence, speeding + running red light)	4%	3%	0%	0%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	0%	7%	9%	2%
	Discreet, average	62%	62%	70%	74%
	Sportive, ambitioned	38%	28%	19%	15%
	Risk-taking, offensive	0%	3%	3%	9%
How confident you are concerning your own driving skills?	Insecure	0%	0%	0%	2%
	Neutral	28%	21%	26%	11%
	Confident	52%	59%	53%	61%
	Very confident	20%	21%	21%	26%

A very small proportion of UK drivers currently used ADAS in their vehicle. This was intentional, as the UK tried to exclude drivers with ADAS where possible, to avoid conflicting messages and existing behavioural influence. A slightly higher proportion of German, Greek and UK drivers had been involved in a recent accident compared with Belgian drivers, however Belgian drivers reported substantially higher incidences of traffic offences; over half of the Belgian drivers had a recent offence, these were mostly speeding offences. Drivers in Greece reported the fewest recent traffic offences.

UK drivers more often described themselves as ‘confident’ or ‘very confident’ compared with Belgian, German and Greek drivers, who had a higher proportion of ‘neutral’ confidence. In all countries, the majority of drivers identified themselves as having a ‘discreet, average’ driving style. Germany and Greece had a higher proportion of drivers who identified as ‘less experienced, hesitant’ compared with Belgium and the UK, whereas the UK had a relatively higher proportion of ‘risk-taking, offensive’ drivers.

Drivers were also asked how often they believed they engaged in certain risk-taking behaviours (question EQ4 in Annex 2). Participants were asked to estimate how often they had engaged in these behaviours over the previous year and responded using a 5-point scale of: ‘almost always’, ‘usually’, ‘about half the time’, ‘seldom’ and ‘never’.

[1] Two participants from Belgium did not complete the full entry questionnaires.

Figure 6 shows the responses for each country. Responses have been grouped into 'almost always / usually', 'about half the time' and 'seldom / never' to make the data more readable, the full tables of responses for each country are given in Annex 1 (Table 57).

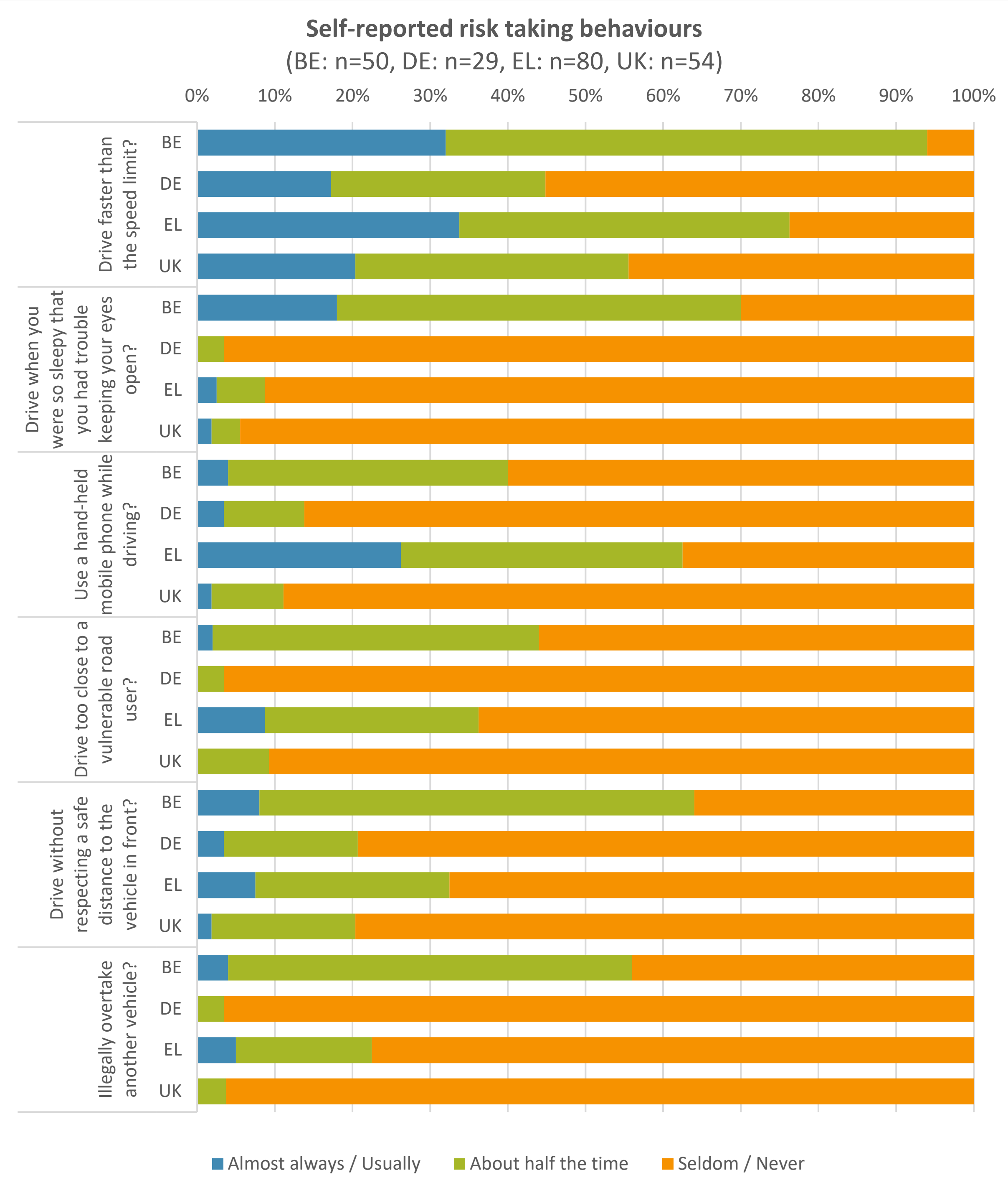


Figure 6: Questionnaire data – self-reported risk-taking behaviours, per country (cars)

Notably, drivers in Belgium reported substantially more risk-taking behaviour compared with Germany and the UK. In particular, when the responses 'about half the time', 'usually' or 'always' are considered together; 94% of Belgian drivers said they exceed the speed limit, 70% said they drove while being so sleepy they struggled to keep their eyes open, and 64% said they drive close to the vehicle in front. Drivers in Greece also reported a relatively high amount of speeding (76% replied that they speed 'about half the time, 'usually' or 'always'), and of all countries they reported the most use of their phone while driving (63%).

3.1.1 Results overview

The tables below describe the valid trip data available for analysis for each country, and also give overview results for each data collection phase. As a reminder, drivers were excluded here if sufficient data were not available for each phase, and also outlier trips were removed (as described in section 2.3.1).

The following sections explore the results in more detail, in relation to both outcome and process evaluation, as well as examining differences between the drivers in each country.

Table 5: Results overview – Belgium cars

Data Collection Phase	Belgium Cars (n=48 drivers)					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	3,366	47,534	180.89	94.47	85.89	5.88
Phase 2	3,200	45,039	185.73	97.50	85.98	5.49
Phase 3	3,805	55,555	188.01	107.01	85.83	5.52
Phase 4	4,395	65,926	177.17	105.46	86.37	5.96
TOT/AV	**14,766**	**214,055**	**183.72**		**86.02**	

Table 6: Results overview – Germany cars

Data Collection Phase	Germany Cars (n=25 drivers)					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	997	17,957	152.72	153.75	81.53	10.09
Phase 2	1,054	14,800	151.05	114.74	80.09	9.90
Phase 3	1,073	17,357	137.38	123.62	81.10	11.59
Phase 4	1,422	21,818	149.59	126.16	79.34	11.63
TOT/AV	**4,546**	**71,933**	**147.73**		**80.42**	

Table 7: Results overview – Greece cars

Data Collection Phase	Greece Cars (n=56 drivers)					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	2,970	25,938	68.60	82.85	78.42	19.14
Phase 2	NA	NA	NA	NA	NA	NA
Phase 3	4,096	36,351	70.16	77.17	77.28	19.36
Phase 4	4,665	47,183	62.86	69.30	77.10	19.17
TOT/AV	**11,731**	**109,472**	**66.86**		**77.50**	

Table 8: Results overview – UK cars

Data Collection Phase	UK Cars (n=49 drivers)					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	3,620	46,661	275.30	249.61	83.81	8.85
Phase 2	3,789	44,388	261.32	223.87	84.43	8.96
Phase 3	3,878	46,401	251.05	225.23	84.43	9.14
Phase 4	5,175	68,619	240.75	219.28	84.68	8.89
TOT/AV	**16,462**	**206,070**	**255.51**		**84.37**	

3.2 Outcome Evaluation

3.2.1 Events and scores analysis

This section presents the descriptive analyses of events and scores data. The following section (3.2.2) gives more information on which results were statistically significant.

First, we consider the **total events** / 100km and overall scores for each country. These are given above but are repeated in the table below to compare each country more easily. Figure 7 and Figure 8 further break down the total events into 'medium' and 'high' severity.

Further analysis of events was also carried out on a combined countries dataset. Due to the different methodology used in Greece, they are not included in the combined data. Therefore, where results refer to combined countries, this comprises Belgium, Germany and the UK.

When looking at the total number of events, it is important to remember that data for Germany and Greece doesn't include the full parameter set, so it is expected their total is lower. It should also be reiterated here that the Belgian drivers were most impacted by delays and varying COVID-19 restrictions, and it's difficult to know whether and how these affected the results.

Table 9: Total events / 100km and overall scores per country and per phase (cars)

Phase	Belgium (n=48)		Germany (n=25)		UK (n=49)		Combined (n=122)	Greece (n=56)	
	Events / 100km	Overall scores	Events / 100km	Overall scores	Events / 100km	Overall scores	Events / 100km	Events / 100km	Overall scores
Phase 1	180.89	85.89	152.72	81.53	275.30	83.81	230.04	68.60	78.42
Phase 2	185.73	85.98	151.05	80.09	261.32	84.43	220.70	NA	NA
Phase 3	188.01	85.83	137.38	81.10	251.05	84.43	214.23	70.16	77.28
Phase 4	177.17	86.37	149.59	79.34	240.75	84.68	204.34	62.86	77.10

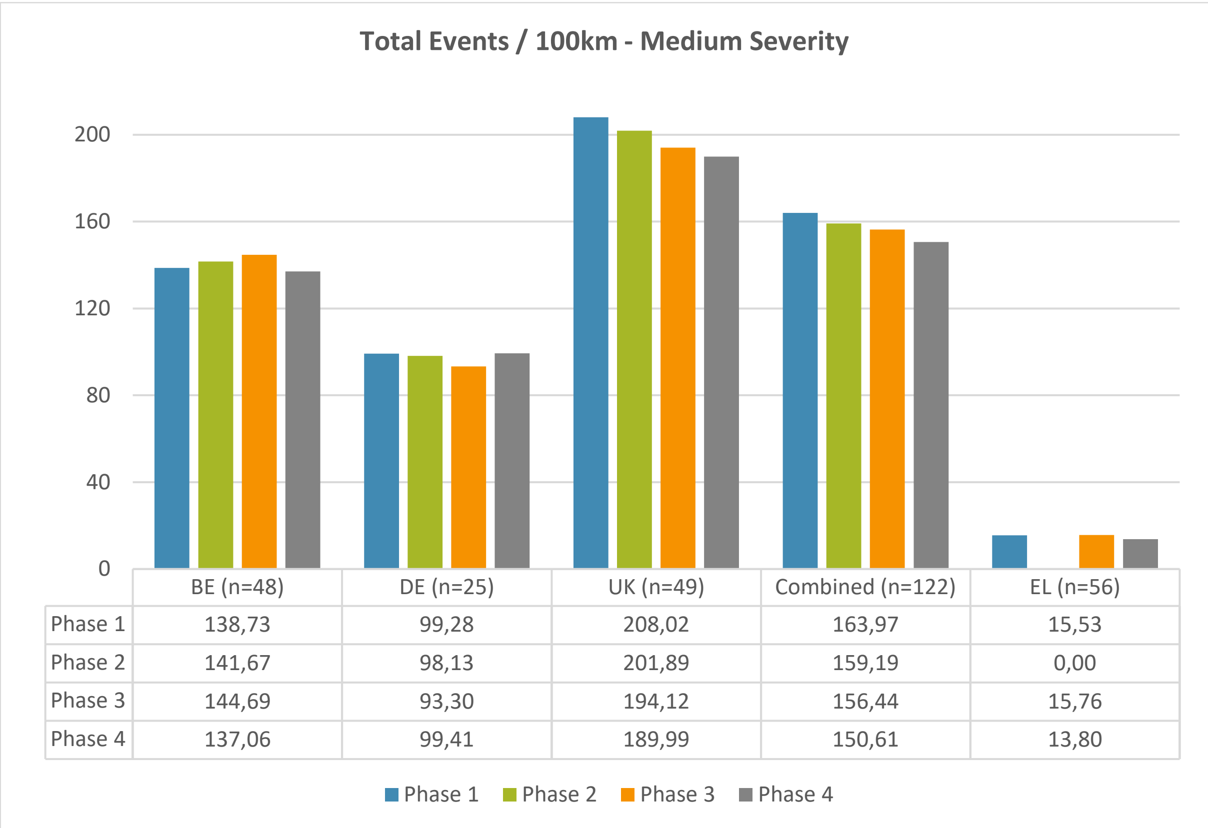
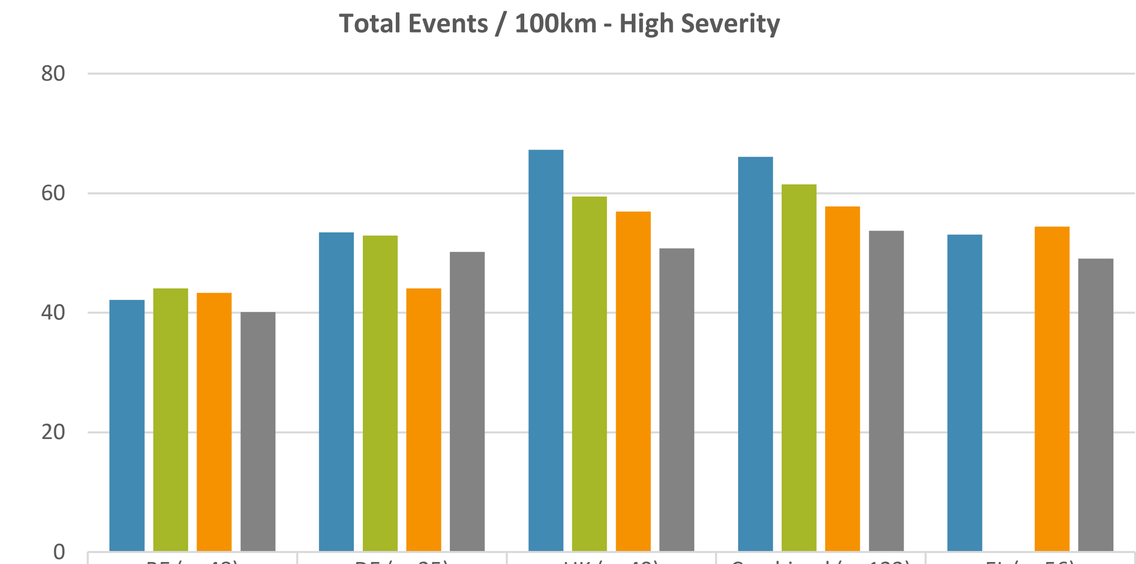


Figure 7: Total medium events / 100km per country and per phase (cars)



	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	42,16	53,44	67,28	66,08	53,07
Phase 2	44,06	52,91	59,43	61,51	0,00
Phase 3	43,32	44,08	56,93	57,79	54,40
Phase 4	40,11	50,17	50,76	53,73	49,07

Figure 8: Total high events / 100km per country and per phase (cars)

It can be seen that, for most countries, ‘medium’ severity events were more frequent than ‘high’ severity, though in Greece it is the opposite. Greek drivers had particularly low numbers of ‘medium’ events, which is discussed in more detail below. For Belgium and the UK, the ‘high’ events accounted for around a quarter of the total, in Germany this was a third, and in Greece it was two thirds. This partly explains the lower scores seen in Germany and Greece, as ‘high’ severity events have a greater impact on scores.

It is noted that the UK had the highest total number of events, however they also had the greatest decrease, and that decrease was consistent across the Phases. The same trend is observed when countries are combined. In all countries there was an overall decrease in events from Phase 1 to Phase 4, however in Belgium and Germany there was an initial increase in events, and in Germany there was an increase in Phase 4.

When comparing the UK and Belgium, the tables in section 3.1.1 show that UK drivers had more trips, but less distance travelled, i.e., a higher proportion of shorter trips. This could suggest that UK drivers had more trips in urban areas, and more interaction with other road users, parked cars, pedestrians etc., which could partly explain the higher number of events.

To further explore these results, it is useful to look at the event changes for each safety promoting goal (SPG) individually and see how they contribute to the total.

The **'vehicle control'** (VC) events and scores are given below in Table 10, Figure 9 and Figure 10, for 'all', 'medium', and 'high' events respectively.

When considering 'vehicle control' events, it should be noted that drivers did not receive real-time warnings in relation to these events, only feedback within the i-DREAMS app. Furthermore, 'vehicle control' behaviour may have been impacted by other behaviour changes, such as drivers responding to warnings about their speed or headway.

Finally, it is reiterated here that 'vehicle control' data for Greece does not include events related to 'steering', only events for 'acceleration' and 'deceleration'.

Table 10: Vehicle control events / 100km and scores per country and per phase (cars)

Phase	Belgium (n=48)		Germany (n=25)		UK (n=49)		Combined (n=122)	Greece (n=56)	
	VC Events / 100km	VC scores	VC Events / 100km	VC scores	VC Events / 100km	VC scores	VC Events / 100km	VC Events / 100km	VC scores
Phase 1	101.53	61.57	96.81	65.04	136.69	56.87	108.98	13.55	87.02
Phase 2	107.96	60.73	94.08	62.08	131.67	56.70	107.68	NA	NA
Phase 3	109.94	59.80	89.46	65.53	130.68	56.79	106.63	14.01	86.75
Phase 4	102.74	61.43	97.26	64.39	130.60	56.19	102.48	11.91	87.83

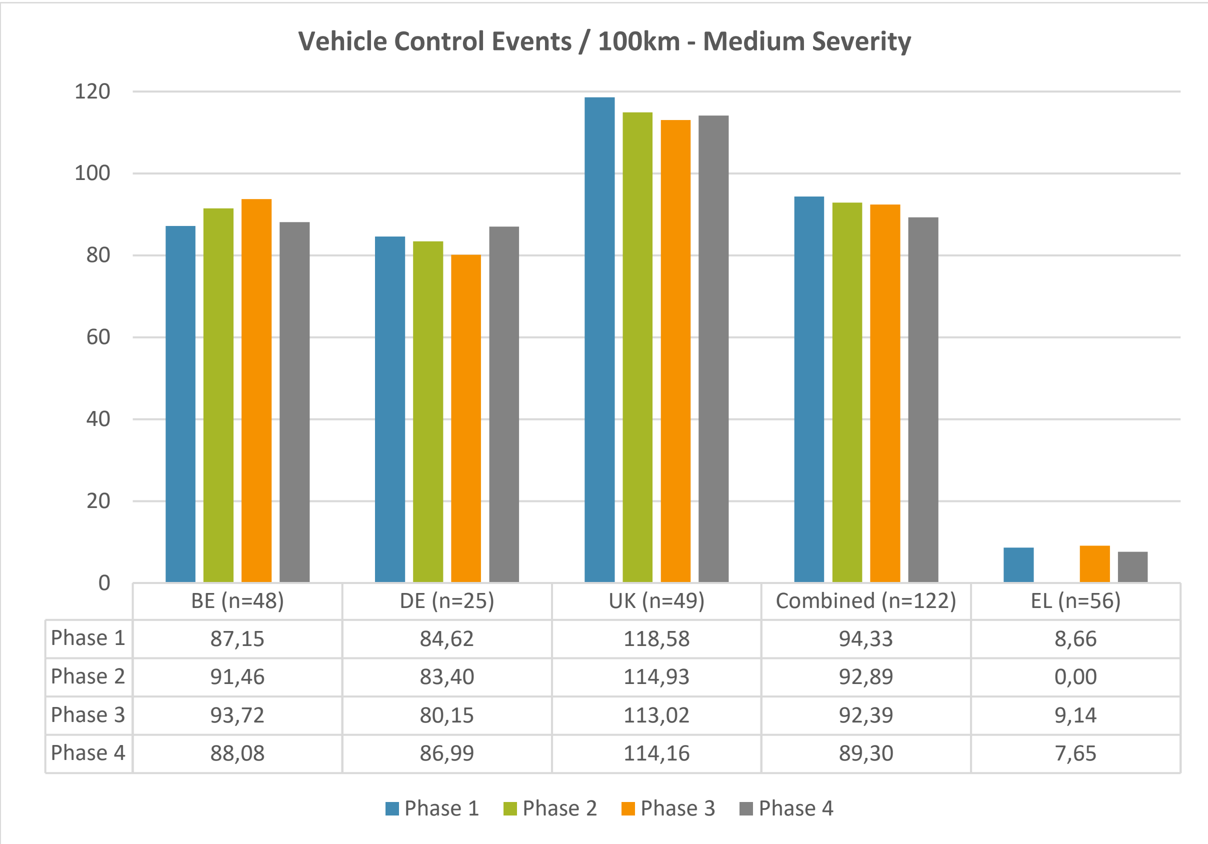
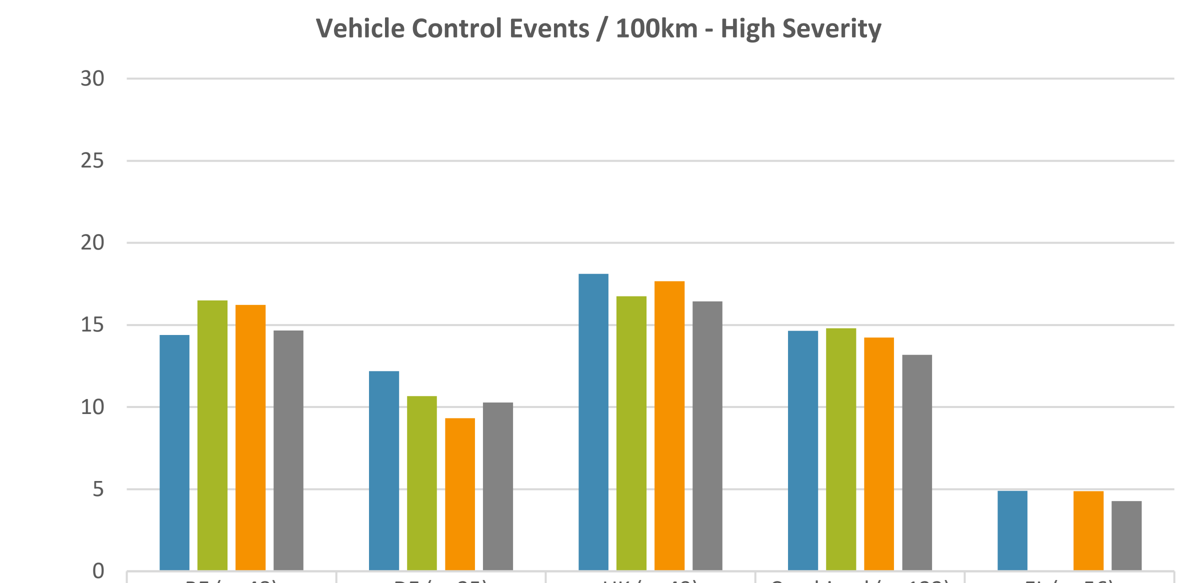


Figure 9: Medium vehicle control events / 100km per country and per phase (cars)



	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	14,38	12,19	18,11	14,65	4,90
Phase 2	16,49	10,67	16,75	14,79	0,00
Phase 3	16,22	9,31	17,65	14,24	4,87
Phase 4	14,67	10,27	16,44	13,18	4,26

Figure 10: High vehicle control events / 100km per country and per phase (cars)

Firstly, the results again show particularly low event numbers for drivers in Greece, which is not fully explained by the lack of steering events. The most likely explanation is that events are not registered as effectively when the app is used compared to when the i-DREAMS in-vehicle system (accelerometer) is used. Therefore, this data may not be comparable to other countries, however the relative change within the country is still valid.

When ‘vehicle control’ events are analysed in isolation, there are some small differences compared to when total events are considered. For Belgium and Germany there is an overall increase in events from Phase 1 to Phase 4, though in Germany there is a decrease if only high events are considered, and there are some decreases between individual phases for each country. However, some of the increases are small, and the statistical significance of these changes is explored in section 3.2.2. For Greece and the UK there is an overall decrease in events, but it is not consistent across each phase. It is interesting to note in the combined countries data there is an overall decrease in events, which is consistent across phases for ‘medium’ severity events.

The ‘vehicle control’ SPG can be further subdivided into the performance objectives (POs) of ‘acceleration’, ‘deceleration’ and ‘steering’. The detailed results for these are given in Annex 1 (Table 58 and Figure 31 - Figure 35) and summarised here.

- In Greece, there were few events at all, and the majority were ‘deceleration’. It is clear from the data that ‘acceleration’ events were registered differently in Greece compared to other countries. There was an overall decrease in ‘acceleration’ and ‘deceleration’ events, for both severities.
- For all countries except Greece, ‘deceleration’ accounted for a very small proportion of VC events (~5% of ‘medium’ and 3% of ‘high’ VC events). The majority were ‘acceleration’ (~43% of ‘medium’ and 65% of ‘high’ VC events) and ‘steering’ (~52% of ‘medium’ and 32% of ‘high’ VC events). It is likely that this is due to the way events

were calculated: although the algorithms were developed based on relevant literature, in practice, harsh braking did not trigger events in the same way as harsh acceleration.

- ‘Acceleration’ events: more ‘medium’ events occurred than ‘high’ events. Overall (Phase 1 to Phase 4), there was an increase in events in Belgium, but a decrease in Germany, the UK, and the combined countries data.
- ‘Deceleration’ events: more ‘medium’ events than ‘high’ events, and little overall change in ‘high’ events (though numbers are very small). Overall, there was an increase in Belgium and Germany, and a decrease in the UK. When counties are combined there is an overall increase.
- ‘Steering’ events: more ‘medium’ events than ‘high’ events. Overall, there was a decrease in Belgium and when countries are combined, and an increase for Germany and the UK.

‘Speeding’ (SPD) events are considered next, the results of which are presented below in Table 11, Figure 11 and Figure 12.

Table 11: Speeding events / 100km and scores per country and per phase (cars)

Phase	Belgium (n=48)		Germany (n=25)		UK (n=49)		Combined (n=122)	Greece (n=56)	
	SPD Events / 100km	SPD scores	SPD Events / 100km	SPD scores	SPD Events / 100km	SPD scores	SPD Events / 100km	SPD Events / 100km	SPD scores
Phase 1	13.96	91.18	55.80	79.92	18.88	91.57	22.54	32.32	75.70
Phase 2	15.47	91.45	56.87	78.67	15.83	92.62	22.27	NA	NA
Phase 3	16.19	91.71	47.92	78.24	14.32	91.74	20.38	31.86	74.22
Phase 4	15.07	91.96	52.29	73.77	13.89	92.07	19.78	29.93	73.72

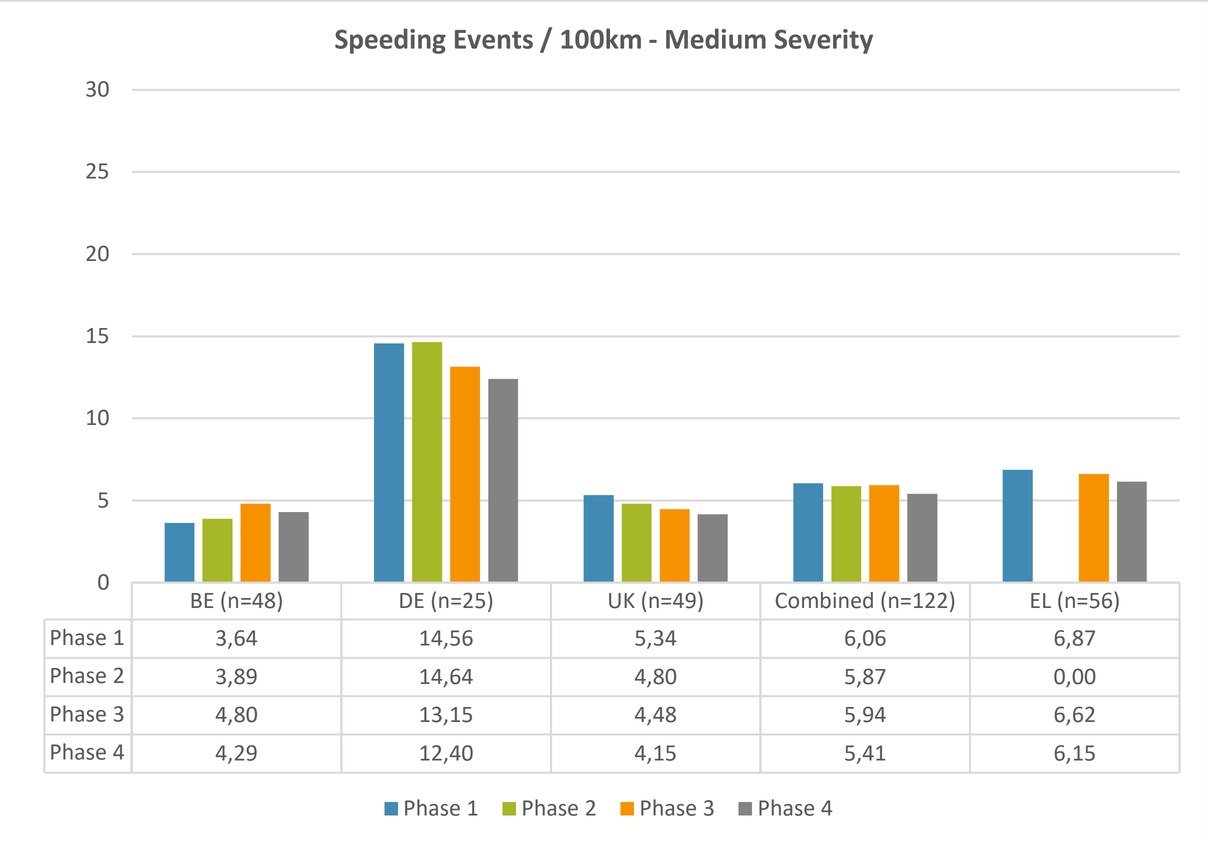
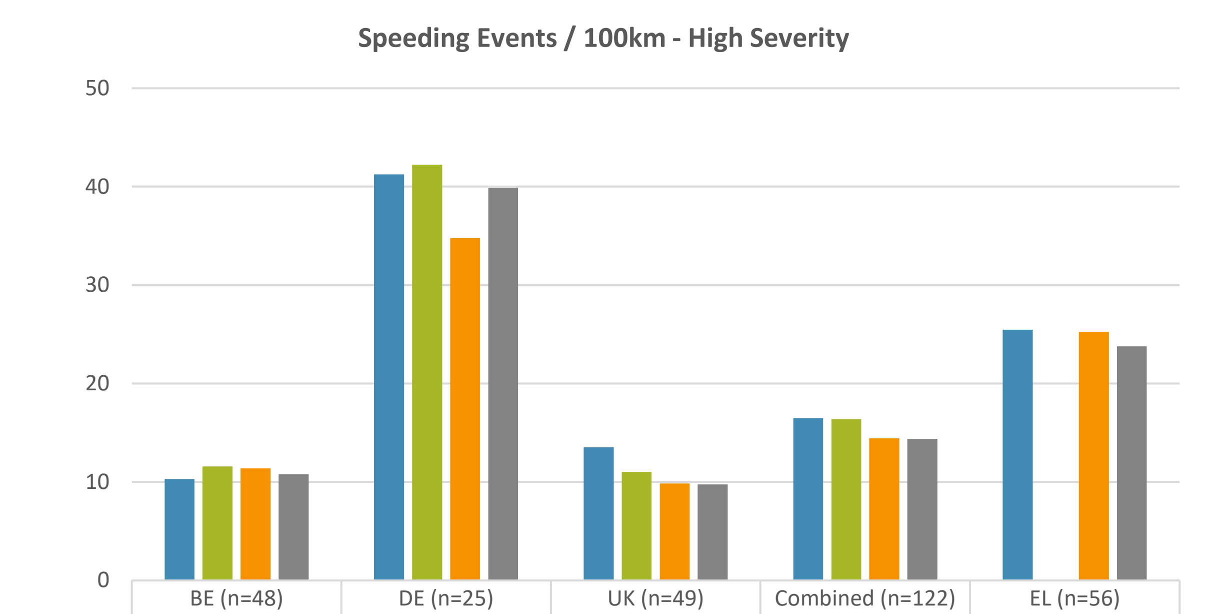


Figure 11: Medium speeding events / 100km per country and per phase (cars)



	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	10,32	41,25	13,54	16,48	25,45
Phase 2	11,58	42,24	11,04	16,39	0,00
Phase 3	11,39	34,77	9,85	14,44	25,24
Phase 4	10,78	39,90	9,74	14,37	23,78

Figure 12: High speeding events / 100km per country and per phase (cars)

For 'speeding', in contrast to other event types, there are more 'high' severity events than 'medium' severity. This is due to the way events are calculated for speeding, as explained in Section 2.3, with only the most severe event being recorded.

Regarding Greece, it is noted that vehicle speed was calculated using the app only, whereas other countries obtained speed directly from the vehicle CAN, so results may not be fully comparable to other countries.

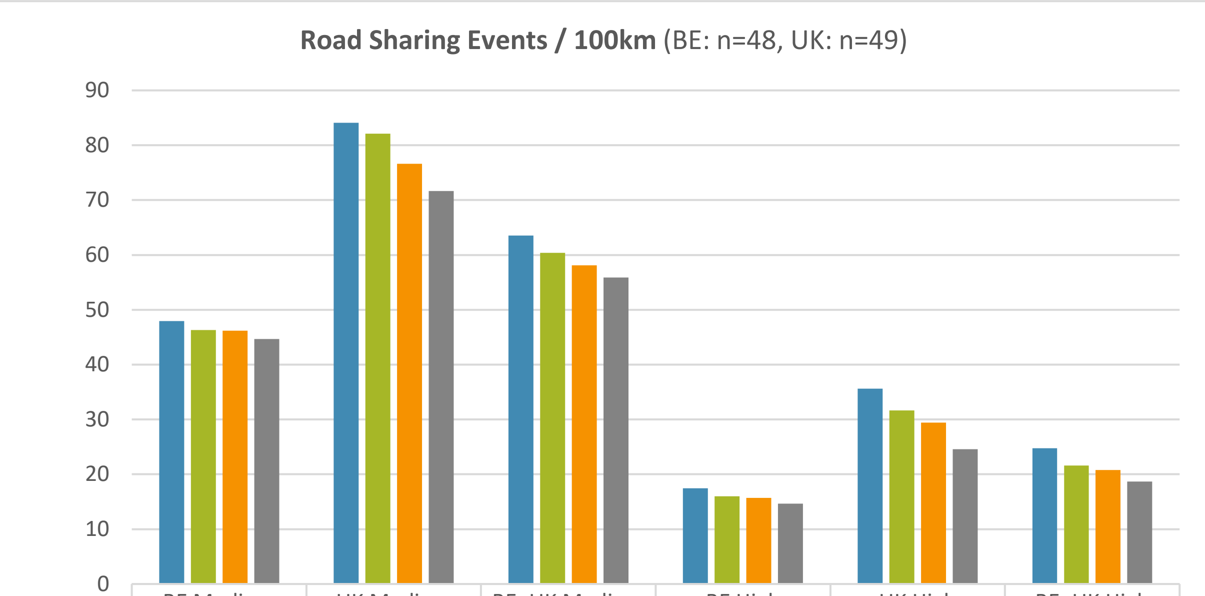
There are substantially more events in Germany compared to Belgium and the UK, particularly 'high' severity events. It's possible this is in some part due to the way events are calculated in the post-processing. In Germany, there are some sections of the highways that do not have legal speed restrictions, and further investigation is needed to understand how speeding events are calculated for these areas. Anecdotally, it is known that drivers sometimes drive very fast in these areas (which they can do legally), therefore if the 'recommended' speed is used as the limit in calculations, this could account for some of the 'high' speed events.

The 'speeding' events again show differences between countries. In most countries, there was an overall (Phase 1 to Phase 4) decrease in speeding events, which was consistent across phases for the UK and Greece. In Belgium however, there was an overall increase, though events did decrease from Phase 3 to Phase 4. When countries data were combined, there was a consistent decrease across the phases.

Finally, **‘road sharing’** (RS) events are analysed for Belgium and the UK (due to issues with data collection ‘road sharing’ data is not available for Germany, and the adapted methodology for Greece did not include ‘road sharing’ data.).

Table 12: Road sharing events / 100km and scores per country and per phase (cars)

Phase	**Belgium (n=48)**		**UK (n=49)**		**BE+UK (n=97)**
	RS Events / 100km	**RS scores**	**RS Events / 100km**	**RS scores**	**RS Events / 100km**
Phase 1	65.40	90.81	119.71	86.86	88.29
Phase 2	62.30	91.74	113.78	88.42	81.96
Phase 3	61.88	91.87	106.04	89.26	78.90
Phase 4	59.36	92.16	96.23	90.54	74.56



	BE Medium	UK Medium	BE+UK Medium	BE High	UK High	BE+UK High
Phase 1	47,94	84,09	63,55	17,46	35,63	24,73
Phase 2	46,31	82,14	60,36	15,99	31,65	21,60
Phase 3	46,17	76,62	58,10	15,71	29,43	20,79
Phase 4	44,69	71,66	55,89	14,67	24,58	18,68

Figure 13: Medium and high road sharing events / 100km per country and per phase (cars)

It can be seen that UK drivers had nearly twice as many RS events as Belgian drivers, though they did show a greater decrease. Both countries showed a consistent decrease in events across each phase, which is seen in both ‘medium’ and ‘high’ events.

The ‘road sharing’ SPG can be further subdivided into the POs of ‘tailgating’, ‘illegal overtaking’, ‘lane departure’, ‘forward collision avoidance’ (FCA), and ‘vulnerable road user collision avoidance’ (VRU CA). The detailed results for these are given in Annex 1 (Table 59, Table 60 and Figure 36) and summarised here.

- ‘Tailgating’ accounted for the majority (82%-86%) of ‘road sharing’ events. There were also a small number of FCA events (6%-11%) and ‘lane departure’ events (6%-8%).
- For both countries there were very few VRU CA events (<1% of ‘road sharing’ events), and almost no ‘illegal overtaking’ events (<0.1%).

- 'Tailgating' events: both Belgium and the UK showed a consistent decrease across the Phases, though the change was greater for the UK.
- 'Lane departure' events: both countries showed an overall decrease, which was again greatest for the UK.
- FCA events: there was little change in the Belgian data; overall there was a very slight increase. The UK showed a consistent decrease across each phase.

3.2.2 Events summary and statistical analysis

To better visualise the range of effects for each country, the tables below give a summary of the change in events for each event type and level of severity, and between each phase. Each table first gives the significance value for the overall test (for cars all data were not normally distributed, therefore the Friedman test was used instead of repeated measures ANOVA). Then, pairwise comparisons between each phase show the direction (if the number of events/100km decreased ↓ or increased ↑), as well as the significance value from the relevant statistical test (Wilcoxon signed rank test).

Highlighted results were significant at the α=0.1 level, which has been chosen over an α of 0.05 due to the small sample sizes.

Table 13: Event change summary and statistical significance, per event type & severity and per phase – BE cars

Change in Number of Events: Belgium Cars (n=48 drivers)										
Event Type		**Friedman test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**					
					P1 - P2		**P2 - P3**		**P3 - P4**	
Total	**Medium**	p = 0.014	↓	0.207	↑	0.910	↑	0.894	↓	0.025
	High	p = 0.053	↓	0.246	↑	0.498	↓	0.436	↓	0.010
	All	p = 0.050	↓	0.151	↑	0.587	↑	0.829	↓	0.013
Vehicle Control	**Medium**	p = 0.101	↑	0.215	↑	0.601	↑	0.758	↓	0.045
	High	p = 0.355	↑	0.544	↑	0.509	↓	0.221	↓	0.267
	All	p = 0.070	↑	0.207	↑	0.430	↑	0.601	↓	0.066
Speeding	**Medium**	p = 0.108	↑	0.159	↑	0.430	↑	0.119	↓	0.512
	High	p = 0.281	↑	0.193	↑	0.140	↓	0.559	↓	0.943
	All	p = 0.122	↑	0.077	↑	0.110	↑	0.189	↓	0.878
Road Sharing	**Medium**	p = 0.228	↓	0.200	↓	0.582	↓	0.532	↓	0.077
	High	p < 0.001	↓	0.003	↓	0.083	↓	0.189	↓	0.014
	All	p = 0.017	↓	0.070	↓	0.128	↓	0.478	↓	0.047

To summarise the results for Belgian drivers, it can be seen in Table 13 above that there was an overall decrease in 'total' and 'road sharing' events, and an overall increase in 'vehicle control' and 'speeding' events. However, many of these changes were not significant.

The statistically significant changes were mostly from Phase 3 to Phase 4, which were all decreases, and 'road sharing' event decreases were also statistically significant in multiple phases / severities. The overall increase in 'speeding' events was significant.

Table 14: Event change summary and statistical significance, per event type & severity and per phase – DE cars

Change in Number of Events: Germany Cars (n=25 drivers)										
Event Type		**Friedman test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**					
					P1 - P2		**P2 - P3**		**P3 - P4**	
Total	**Medium**	p = 0.311	↑	0.790	↓	0.630	↓	0.290	↑	0.710
	High	p = 0.003	↓	0.165	↓	0.812	↓	0.002	↑	0.442
	All	p = 0.037	↓	0.275	↓	0.791	↓	0.075	↑	0.508
Vehicle Control	**Medium**	p = 0.874	↑	0.890	↓	0.710	↓	0.490	↑	0.600
	High	p = 0.647	↓	0.370	↓	0.480	↓	0.310	↑	0.350
	All	p = 0.691	↑	0.870	↓	0.790	↓	0.430	↑	0.490
Speeding	**Medium**	p = 0.323	↓	0.085	↑	0.312	↓	0.751	↓	0.287
	High	p = 0.068	↓	0.230	↑	0.672	↓	0.006	↑	0.411
	All	p = 0.218	↓	0.080	↑	0.958	↓	0.020	↑	0.916

For Germany, there was an overall decrease in most categories, apart from an overall increase in 'medium total' events, and 'medium' and 'all' 'vehicle control' events. Again, many of these changes were not significant however.

Statistically significant results were seen for decreases in overall 'speeding' events ('medium' and 'all'), in 'speeding' in Phase 2 to Phase 3 ('high' and 'all'), and also in Phase 2 to Phase 3 for 'total' events ('high' and 'all'). None of the event increases were statistically significant.

Table 15: Event change summary and statistical significance, per event type & severity and per phase – UK cars

Change in Number of Events: UK Cars (n=49 drivers)										
Event Type		**Friedman test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**					
					P1 - P2		**P2 - P3**		**P3 - P4**	
Total	**Medium**	p = <0.001	↓	0.001	↓	0.187	↓	0.251	↓	0.259
	High	p = <0.001	↓	<0.001	↓	0.003	↓	0.100	↓	0.024
	All	p = <0.001	↓	<0.001	↓	0.031	↓	0.231	↓	0.094
Vehicle Control	**Medium**	p = 0.305	↓	0.028	↓	0.644	↓	0.525	↑	0.538
	High	p = 0.428	↓	0.042	↓	0.878	↑	0.340	↓	0.436
	All	p = 0.060	↓	0.016	↓	0.845	↓	0.486	↓	0.406
Speeding	**Medium**	p = 0.079	↓	0.006	↓	0.132	↓	0.601	↓	0.401
	High	p = <0.001	↓	0.001	↓	<0.001	↓	0.807	↓	0.941
	All	p = <0.001	↓	<0.001	↓	<0.001	↓	0.800	↓	0.672
Road Sharing	**Medium**	p = 0.010	↓	<0.001	↓	0.104	↓	0.198	↓	0.100
	High	p = <0.001	↓	<0.001	↓	0.001	↓	0.198	↓	0.003
	All	p = <0.001	↓	<0.001	↓	0.007	↓	0.303	↓	0.013

The events for UK drivers decreased for nearly every event category and phase, with only a few increases in vehicle control events between some phases.

The overall decrease (Phase 1 to Phase 4) was statistically significant for every event type. Further significant decreases were seen in other phases, particularly for 'total' and 'road sharing' events. The change from Phase 2 to Phase 3 had the least significant results. None of the event increases were statistically significant.

Table 16: Event change summary and statistical significance, per event type & severity and per phase – combined cars

Change in Number of Events: Combined Cars (n=122 drivers)										
Event Type		**Friedman test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**					
					P1 - P2		**P2 - P3**		**P3 - P4**	
Total	**Medium**	p < 0.001	↓	<0.001	↓	0.582	↓	0.261	↓	0.032
	High	p < 0.001	↓	<0.001	↓	0.121	↓	0.003	↓	0.017
	All	p < 0.001	↓	<0.001	↓	0.341	↓	0.095	↓	0.020
Vehicle Control	**Medium**	p = 0.045	↓	0.020	↓	0.665	↓	0.393	↓	0.164
	High	p = 0.181	↓	0.034	↑	0.969	↓	0.066	↓	0.487
	All	p = 0.006	↓	0.010	↓	0.581	↓	0.269	↓	0.189
Speeding	**Medium**	p = 0.698	↓	0.070	↓	0.307	↑	0.672	↓	0.122
	High	p = 0.035	↓	0.062	↓	0.108	↓	0.219	↓	0.601
	All	p = 0.250	↓	0.026	↓	0.089	↓	0.654	↓	0.811
Road Sharing	**Medium**	p = 0.004	↓	<0.001	↓	0.108	↓	0.215	↓	0.025
	High	p < 0.001	↓	<0.001	↓	<0.001	↓	0.074	↓	0.001
	All	p < 0.001	↓	<0.001	↓	0.002	↓	0.156	↓	0.002

When the countries data are combined, there is a decrease in events for nearly every event category and phase, with increases only seen in Phase 1 to Phase 2 for 'high vehicle control' events, and in Phase 2 to Phase 3 for 'medium speeding' events.

The overall (Phase 1 to Phase 4) decrease was statistically significant for every event type. Further significant decreases were seen in other phases, particularly for 'total' and 'road sharing' events.

Table 17: Event change summary and statistical significance, per event type & severity and per phase – EL cars

Change in Number of Events: Greece Cars (n=56 drivers)								
Event Type		**Friedman test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**			
					P1 - P3		**P3 - P4**	
Total	**Medium**	p = 0.008	↓	0.016	↑	1.000	↓	0.036
	High	p = 0.005	↓	0.021	↑	1.000	↓	0.012
	All	p = 0.003	↓	0.016	↑	1.000	↓	0.007
Vehicle Control	**Medium**	p = 0.251	↓	0.084	↑	0.917	↓	0.044
	High	p = 0.235	↓	0.367	↓	0.741	↓	0.014
	All	p = 0.712	↓	0.130	↑	0.945	↓	0.053
Speeding	**Medium**	p = 0.223	↓	0.113	↓	0.760	↓	0.043
	High	p = 0.100	↓	0.053	↓	0.899	↓	0.018
	All	p = 0.013	↓	0.025	↓	0.990	↓	0.005

Finally for Greece, we see that there is an overall (Phase 1 to Phase 4) decrease for each event type and severity. Most of the changes between Phases were also decreases, though there were come increases from Phase 1 to Phase 3.

The decreases between Phase 3 and Phase 4 were statistically significant for all event types. Furthermore, the overall decrease was significant for the majority of event types, though less so for 'vehicle control' events. None of the event increases were statistically significant.

Statistical Analysis of Combined Countries Data

Additional statistical analysis was carried out on the combined countries data (BE, DE and UK), to determine significant differences between Phases, and also between countries.

Generalized linear mixed model (GLMM) analysis was used, as this allows for the analysis of data when (a) random effects are present (e.g., the case of repeated responses from study subjects/participants or multi-level data structure), and (b) it has a nonnormal distribution. We applied negative binomial (NB) GLMM to the data, since our independent variable (events per 100km) is a count variable, and it exhibits overdispersion.

The GLMM results are presented below, in turn for 'total', 'vehicle control', 'speeding' and 'road sharing' events, for 'medium', 'high' and 'medium + high' severities. In all analyses, the following parameters were used:

- Dependent variable = events per 100km
- Independent variables = Phase (Phase 1 as reference), Country (BE as reference)

Highlighted results were significant at the α=0.05 level (which is now used here as there is a larger sample).

Table 18: GLMM results – total events

GLMM Results: Total Events / 100km						
Fixed Effects	**Medium Events**		**High Events**		**Medium + High Events**	
	Estimate	**P-value**	**Estimate**	**P-value**	**Estimate**	**P-value**
(Intercept)	4.8305	<0.001	3.6522	<0.001	5.1140	<0.001
Phase 2 vs Phase 1	-0.0196	0.433	-0.0552	0.064	-0.0261	0.277
Phase 3 vs Phase 1	-0.0450	0.072	-0.1403	<0.001	-0.0709	0.003
Phase 4 vs Phase 1	-0.0880	<0.001	-0.1950	<0.001	-0.1157	<0.001
DE vs BE	-0.4570	0.001	0.0588	0.725	-0.3158	0.020
UK vs BE	0.2604	0.017	0.1404	0.308	0.2304	0.039
Random Effects (User_ID)	**Medium Events**		**High Events**		**Medium + High Events**	
Variance	0.2806		0.4449		0.2944	
Standard Deviation	0.5297		0.6670		0.5426	

For all severities of 'total' events, the expected log count of events per 100km decreases from Phase 1 (baseline) to each of Phase 2, Phase 3 and Phase 4. For all severities this was significant from Phase 1 to Phase 4. Furthermore, for 'high' severity and combined severities events, there was a significant decrease from Phase 1 to Phase 3.

The country variable indicates there were more events in the UK compared to Belgium for all severities, and for Germany, there were more 'high' severity events but fewer 'medium' severity events. These differences were only significant for 'medium' and 'medium + high' events, i.e., differences between countries for 'high' events only were not statistically significant.

Looking at the variance between drivers, more variance was found in 'high' events than in 'medium' events.

Table 19: GLMM results – vehicle control events

GLMM Results: Vehicle Control Events / 100km						
Fixed Effects	**Medium Events**		**High Events**		**Medium + High Events**	
	Estimate	**P-value**	**Estimate**	**P-value**	**Estimate**	**P-value**
(Intercept)	4.2769	<0.001	2.1431	<0.001	4.4187	<0.001
Phase 2 vs Phase 1	0.0060	0.848	-0.0490	0.383	0.0032	0.922
Phase 3 vs Phase 1	-0.0114	0.714	-0.1228	0.030	-0.0221	0.498
Phase 4 vs Phase 1	-0.0484	0.121	-0.1582	0.006	-0.0625	0.056
DE vs BE	-0.1061	0.560	-0.2807	0.379	-0.1395	0.469
UK vs BE	0.1407	0.348	-0.0365	0.890	0.1110	0.485
Random Effects (User_ID)	**Medium Events**		**High Events**		**Medium + High Events**	
Variance	0.5283		1.587		0.5931	
Standard Deviation	0.7268		1.260		0.7701	

For ‘vehicle control’ events, for ‘medium’ and ‘medium + high’ severities the events per 100km increases from Phase 1 to Phase 2, however the increase was not statistically significant. For all severities, events per 100km decrease from Phase 1 to Phase 3 and Phase 1 to Phase 4. For ‘medium’ and combined severities these decreases were not statistically significant, but they were significant for ‘high’ events..

None of the country differences were statistically significant for ‘vehicle control’ events.

Again, we see more variance between drivers for ‘high’ events compared to ‘medium’ events.

Table 20: GLMM results – speeding events

GLMM Results: Speeding Events / 100km						
Fixed Effects	**Medium Events**		**High Events**		**Medium + High Events**	
	Estimate	**P-value**	**Estimate**	**P-value**	**Estimate**	**P-value**
(Intercept)	1.1741	<0.001	2.1155	<0.001	2.4574	<0.001
Phase 2 vs Phase 1	-0.0512	0.432	-0.0291	0.603	-0.0377	0.472
Phase 3 vs Phase 1	-0.0093	0.886	-0.1236	0.029	-0.0897	0.090
Phase 4 vs Phase 1	-0.1131	0.086	-0.1162	0.041	-0.1109	0.037
DE vs BE	1.2644	<0.001	1.3415	<0.001	1.3301	<0.001
UK vs BE	-0.0178	0.913	-0.2294	0.216	-0.1583	0.354
Random Effects (User_ID)	**Medium Events**		**High Events**		**Medium + High Events**	
Variance	0.5326		0.7599		0.6494	
Standard Deviation	0.7298		0.8717		0.8059	

For ‘speeding’ events, for all severities there was a decrease in events per 100km between Phase 1 and each of the other Phases. The decreases in ‘medium’ events were not statistically significant. For ‘high’ and combined severity events, decreases were significant from Phase 1 to Phase 3 and to Phase 4.

Drivers in Germany had more events compared with drivers in Belgium, which was statistically significant for all severities. Drivers in the UK had less events than drivers in Belgium, but the difference was not significant.

Variance between drivers was again greater for ‘high’ events than ‘medium’ events. Furthermore, for ‘medium’ severity events the driver variance for ‘speeding’ events was similar to that for ‘vehicle control’ events. However, for ‘high’ events, there was less variance in ‘speeding’ compared with ‘vehicle control’.

Table 21: GLMM results – road sharing events

GLMM Results: Road Sharing Events / 100km						
Fixed Effects	**Medium Events**		**High Events**		**Medium + High Events**	
	Estimate	**P-value**	**Estimate**	**P-value**	**Estimate**	**P-value**
(Intercept)	3.7260	<0.001	2.6527	<0.001	4.0354	<0.001
Phase 2 vs Phase 1	-0.0622	0.017	-0.1495	<0.001	-0.0894	0.001
Phase 3 vs Phase 1	-0.1032	<0.001	-0.1951	<0.001	-0.1354	<0.001
Phase 4 vs Phase 1	-0.1462	<0.001	-0.3037	<0.001	-0.1957	<0.001
UK vs BE	0.4365	0.002	0.4942	0.002	0.4509	0.002
Random Effects (User_ID)	**Medium Events**		**High Events**		**Medium + High Events**	
Variance	0.4779		0.6172		0.5140	
Standard Deviation	0.6913		0.7856		0.7169	

For 'road sharing' events, for all severities there was a decrease in events per 100km between Phase 1 and each of the other Phases, which were all statistically significant. There were also significant differences between events for UK and Belgian drivers, for all severities.

There was again more variance observed between drivers for 'high' events compared to 'medium' events. However, variance was less for 'road sharing' events compared with 'speeding' and 'vehicle control' events.

3.2.3 Driver fitness analysis

Due to issues with data collection, the results for the SPG 'driver fitness' are presented here separately. Within 'driver fitness' there are the POs 'fatigue' and 'distraction'.

Results for **'fatigue'** are given first, for Belgium, Germany and the UK (fatigue data were not collected in Greece). As can be seen in the tables below, very few events were recorded; these results are included for completeness, but little conclusions can be drawn from them. There are several reasons why there may be so few events: firstly, drivers did not always remember to wear the heart rate monitor bracelet, secondly, the bracelet did not always connect to the i-DREAMS system, and finally, fatigue severe enough to be a safety risk is itself a rare event.

Table 22: Fatigue events / 100km and scores per country and per phase (cars)

Phase	**Belgium (n=48)**		**Germany (n=25)**		**UK (n=49)**	
	Fatigue Events / 100km	**Fatigue scores**	**Fatigue Events / 100km**	**Fatigue scores**	**Fatigue Events / 100km**	**Fatigue scores**
Phase 1	0.0018	99.98	0.1092	99.54	0.0094	99.93
Phase 2	0.0005	99.98	0.0969	99.52	0.0253	99.96
Phase 3	0.0020	99.93	0.0073	99.85	0.0062	99.94
Phase 4	0.0027	99.94	0.0324	99.64	0.0263	99.92

Table 23: Medium and high fatigue events / 100km per country and per phase (cars)

Phase	**Medium Fatigue Events / 100km**			**High Fatigue Events / 100km**		
	BE (n=48)	**DE (n=25)**	**UK (n=29)**	**BE (n=48)**	**DE (n=25)**	**UK (n=29)**
Phase 1	0.0012	0.1057	0.0093	0.0007	0.0035	0.0001
Phase 2	0.0003	0.0942	0.0252	0.0003	0.0028	0.0002
Phase 3	0.0017	0.0039	0.0059	0.0003	0.0034	0.0004
Phase 4	0.0019	0.0309	0.0230	0.0008	0.0015	0.0034

The event change summary and statistical results for 'fatigue' are given in the table below. Germany showed an overall decrease, whilst Belgium and the UK showed an overall increase. However, all changes were very small, and no results were statistically significant.

Table 24: Event change summary and statistical significance, per country, per event severity and per phase – fatigue events (cars)

Change in Number of Events: Fatigue											
Country / Severity		Friedman / ANOVA test significance	Overall Change (Phase 1 – Phase 4)		Change Between Phases						
					P1 - P2		P2 - P3		P3 - P4		
Belgium (n=48)	Medium	p = 0.353	↑	0.327	↓	0.500	↑	0.161	↑	0.424	
	High	p = 0.887	↑	1.000	↓	0.500	-	1.000	↑	0.655	
	All	p = 0.335	↑	0.208	↓	0.463	↑	0.333	↑	0.477	
Germany (n=25)	Medium	p = 0.527	↓	0.450	↓	0.580	↓	1.000	↑	0.530	
	High	p = 0.870	↓	0.590	↓	0.790	↑	1.000	↓	1.000	
	All	p = 0.527	↓	0.550	↓	0.580	↓	1.000	↑	0.530	
UK (n=49)	Medium	p = 0.194	↑	0.191	↑	0.575	↓	0.767	↑	0.307	
	High	p = 0.845	↑	0.285	↑	0.655	↑	1.000	↑	0.285	
	All	p = 0.337	↑	0.215	↑	0.638	↓	0.583	↑	0.301	

Results for **'distraction'** are only presented for Greek and some UK drivers only. For reasons that are not fully understood, valid distraction data were only available for four Belgian drivers and two German drivers, which was an insufficient sample for analysis.

Table 25: Distraction events / 100km and scores per phase (UK cars)

Phase	Greece (n=56)		UK (n=13)	
	Distraction Events / 100km	Distraction scores	Distraction Events / 100km	Distraction scores
Phase 1	22.72	72.55	25.14	73.24
Phase 2	NA	NA	14.41	80.21
Phase 3	24.29	70.87	15.88	79.01
Phase 4	21.02	69.76	17.34	79.19

Table 26: Event change summary and statistical significance, per phase – UK distraction events (cars)

Change in Number of Events: Distraction									
Country	Friedman test significance	Overall Change (Phase 1 – Phase 4)		Change Between Phases					
				P1 - P2		P2 - P3		P3 - P4	
UK (n=13)	p = 0.314	↓	0.701	↓	0.382	↑	0.116	↓	0.600
				P1 - P3			P3 - P4		
Greece (n=56)	p = 0.221	↓	0.080	↑	0.640		↓	0.020	

Although the UK data is still a small sample, results indicate an overall decrease in 'distraction' events, however this was not statistically significant. In Greece, there was an overall decrease in 'distraction' events which was statistically significant, with a further significant decrease from Phase 3 to Phase 4.

3.2.4 Questionnaire analysis

A set of 12 questions were asked identically at both trial entry and trial exit (respectively EQ11 and EX3 in Annex 2), to allow analysis of before and after responses. These questions assess the change objectives level of the logic model of change, and relate to the areas of perceived knowledge', 'self-efficacy', 'attitude', 'personal norm', and 'subjective norm', which are described in more detail in section 2.2.

Participants were asked to respond using a 5-point Likert scale, where 1 = strongly disagree and 5 = strongly agree. Figure 14 below shows the before and after average response when all participants are included. It is noted that the sample sizes here differ from the events/scores results, as analysis here was on participants who completed both entry and exit questionnaires.

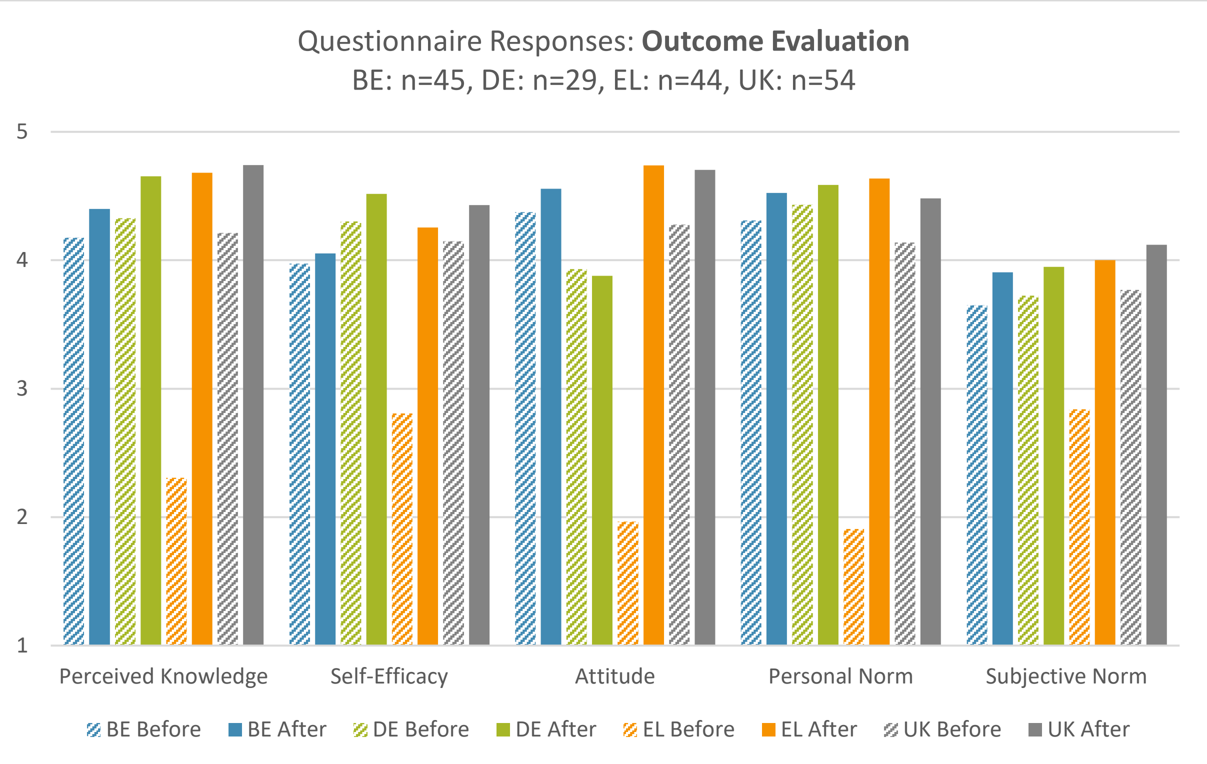


Figure 14: Questionnaire data – outcome evaluation, per country (cars)

These results are interesting, and show that drivers' subjective opinions of their socio-cognitive dispositions towards safety increased through being exposed to the i-DREAMS technology. This was for all evaluation measures for all countries, except 'attitude' for German drivers.

It can also be seen that, for BE, DE and UK, the data were relatively similar, and responses were generally positive both before and after, with most drivers selecting 'agree' or 'totally agree' in almost all evaluation areas. Responses for 'subjective norm' were a little lower, but still above 'neutral'. However, the before scores for Greek drivers were very low compared to other countries, though after scores showed a substantial increase.

Table 27 further shows the individual questions within each category, and indicates which results were significant in the before-after statistical analyses. Data were not normally distributed, therefore either the Wilcoxon sign-rank test or the paired-sample sign test was used (depending on whether or not the distribution of difference was symmetrical). Results highlighted in blue show where the 'after' value is statistically significantly different to the 'before' value at the $\alpha=0.05$ level (although the sample sizes are small, all results significant at $\alpha=0.1$ were also significant at $\alpha=0.05$). Complete statistical results (including which test was used for each measure), as well as the before and after averages for individual questions, are given in Annex 1 (Table 61, Table 62).

Table 27: Questionnaire data – outcome evaluation statistical significance, per country (cars)

Evaluation Measure	BE (n=45)		DE (n=29)		EL (n=44)		UK (n=54)	
	Before	After	Before	After	Before	After	Before	After
Perceived Knowledge								
I know the benefits of safe driving I know what is needed to drive safely	4.18	4.40	4.33	4.66	2.31	4.68	4.21	4.74
Self-Efficacy								
I have the skills to drive safely I feel competent enough to drive safely I control whether I drive safely or not For me, safe driving is easy to do	3.97	4.05	4.30	4.52	2.81	4.26	4.15	4.43
Attitude								
Safe driving is important to avoid crashes Safe driving makes me feel comfortable	4.38	4.56	3.93	3.88	1.97	4.74	4.28	4.70
Personal Norm								
For me personally, safe driving is important Safe driving should be a personal obligation	4.31	4.52	4.43	4.59	1.91	4.64	4.14	4.48
Subjective Norm								
My friends think safe driving is important My colleagues find it important to drive safely	3.65	3.91	3.72	3.95	2.84	4.00	3.77	4.12

For all countries the results were significant for the 'perceived knowledge' category. This shows that through exposure to the technology, participants felt that they better knew the benefits of safe driving, and perhaps more importantly, they better knew what is needed to drive safely. Results for 'subjective norm' were also significant for all countries, which suggests exposure to the technology positively impacted this particular socio-cognitive disposition, though it's not fully clear why this would be. Furthermore, for Greece and the UK, the before-after differences were statistically significant for every question category.

3.3 Process Evaluation

3.3.1 App usage

The process evaluation results in this section include all drivers who engaged with the i-DREAMS app; therefore, the sample size is different from other reported results. The app was available to participants in Phase 3 and Phase 4 of the study, with additional functions activated during Phase 4, as described in section 2.1.

Table 28 shows the number of drivers for each country who used the app, and the total number of app visits. Figure 15 further shows the total app visits on each day, to show how app use varied throughout the trial. As the app was not active for the first eight weeks. app use started on day 57, and it is reminded here that Phase 4 had a longer duration than Phase 3 (also note the days were adjusted for Greece, who did not have a Phase 2).

Table 28: Total app users and visits per country (cars)

	Belgium	Germany	Greece[2]	UK
Total # app users	49	23	71	51
Total # app visits	2768	342	1412	3594

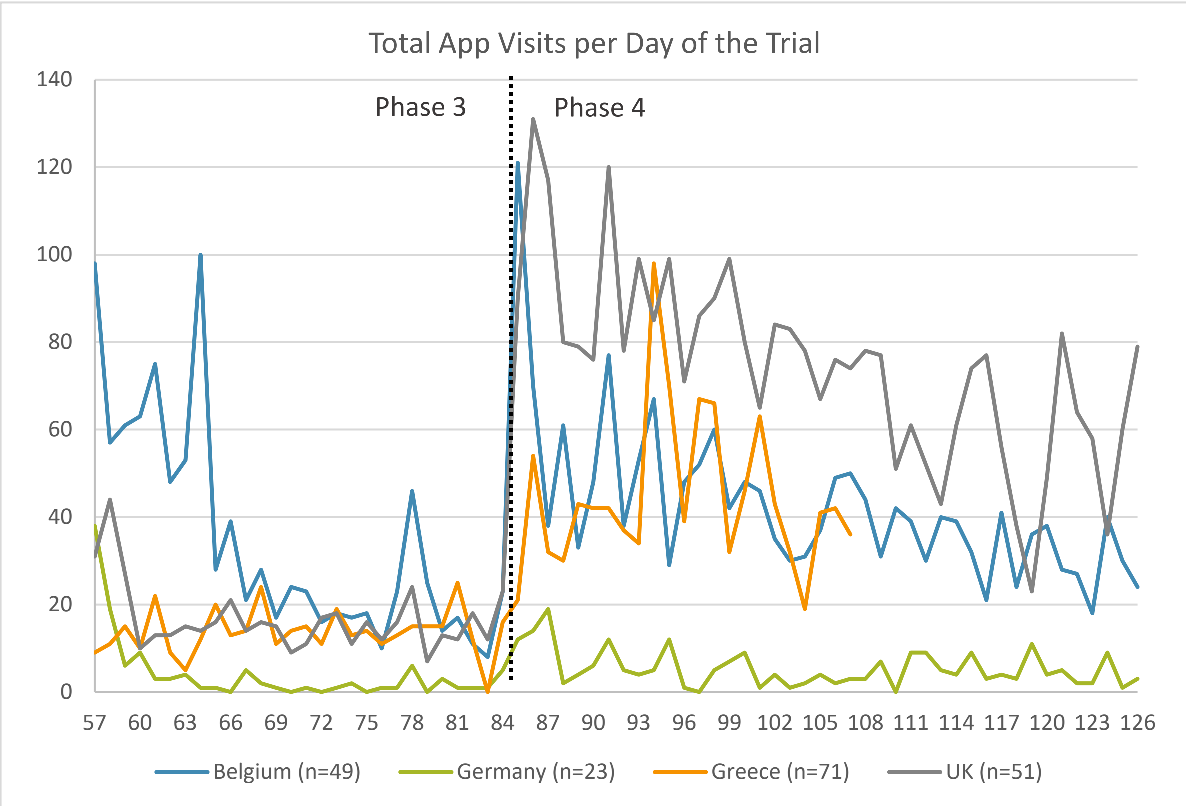


Figure 15: Total app visits per country, per day of trial (cars)

In all countries, drivers showed more app use in Phase 4 of the trial compared with Phase 3, after the introduction of the gamification features. However, the level of engagement with the app varied between countries; in Phase 3, the highest usage was by Belgian drivers, and in Phase 4 the UK drivers had the most usage. In Belgium and the UK there was a peak in the usage at the start of each phase, with app use gradually decreasing throughout the phase, however usage by Greek drivers fluctuated more. It is noted that the total number of visits by

[2] Due to time constraints, the data for Greece was exported before the end of the trial, therefore the total number of app visits is not comparable to other countries, as less days are counted.

Greek drivers was relatively low, considering their higher number of drivers; Table 29 below shows an average of 28 visits per day, which is lower than Belgium and the UK. Even accounting for having fewer drivers, German drivers showed particularly low app use (an average of 5 visits per day).

Additional results are given in Annex 1 (Figure 37 and Figure 38), which show the number of users who visited the app each day, and the average number of visits made by these users. As with total visits, more users engaged with the app during Phase 4 than Phase 3. In most countries, the average visits per user was also higher in Phase 4, however for Belgian drivers the average visits per user was more comparable between Phases. Table 29 shows that Belgium and the UK had on average more users per day than Germany and Greece, though it's interesting to note the average visits per user is similar for Belgium and Greece.

Table 29: Average daily app users and visits per country (cars)

	Belgium	Germany	Greece	UK
Total # users	49	23	71	51
Av. # visits per day	39.5	4.9	27.7	51.3
Av. # users per day	18.4	2.5	12.7	17.6
Av. # visits per user per day	2.1	1.5	2.0	2.6

Figure 16 and Figure 17 next show the proportion of visits per day of the week and per time of the day for each country. Generally, app use was consistent throughout the week, with little differences in the Greek and UK data, though Belgian and German drivers had a peak on Tuesdays, and German drivers had lower use at the weekend. In Belgium and the UK, the time of day shows three distinct peaks in use: 7AM, midday, 9PM. This was when the participants received push notifications that were aimed at increasing app use, so cannot be interpreted as times when app use is more likely, but rather demonstrates the effectiveness of push notifications. It is interesting that the German and Greek data doesn't show the same peaks, though the German sample is quite small to draw conclusions from.

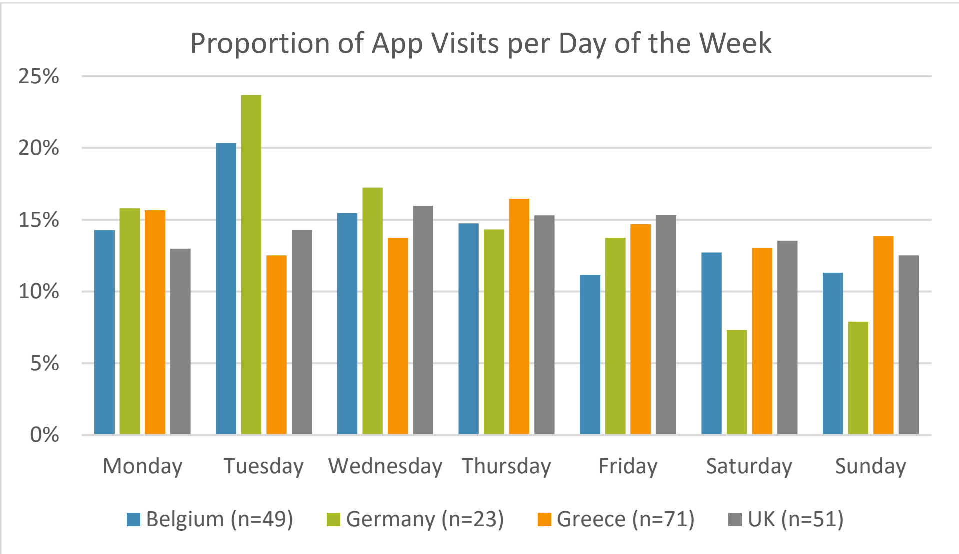


Figure 16: Proportion of app visits per country, per day of the week (cars)

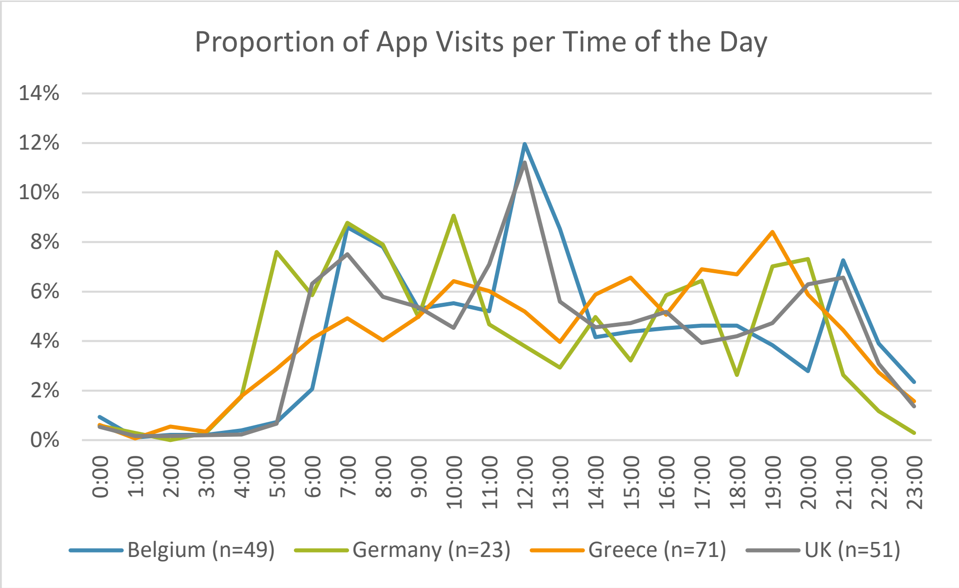


Figure 17: Proportion of app visits per country, per time of the day (cars)

Finally, the app use data were analysed to determine how frequently different app functions were used by drivers. These results are given below in Table 30, and the three most popular functions for each country are highlighted (note that the total in this table is higher than the total visits stated above, as a user could visit multiple areas of the app within the same 'visit').

Table 30: App functionalities used, per country (cars)

App functionality	**Belgium (n=49)**		**Germany (n=23)**		**Greece (n=71)**		**UK (n=51)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Open the trend menu	106	1.8%	20	2.5%	76	1.7%	123	1.7%
Open the goal menu	579	9.6%	98	12.5%	248	5.4%	817	11.3%
Join a goal	227	3.8%	31	3.9%	52	1.1%	277	3.8%
Open the con menu	257	4.3%	38	4.8%	136	3.0%	250	3.5%
Dislike a con	3	0.0%	0	0.0%	0	0.0%	2	0.0%
Like a con	20	0.3%	1	0.1%	1	0.0%	9	0.1%
Open the fact menu	568	9.4%	58	7.4%	207	4.5%	730	10.1%
Like a fact	22	0.4%	1	0.1%	0	0.0%	9	0.1%
Open the pro menu	285	4.7%	43	5.5%	173	3.8%	280	3.9%
Like a pro	40	0.7%	3	0.4%	1	0.0%	17	0.2%
Dislike a pro	0	0.0%	2	0.3%	0	0.0%	0	0.0%
Open the tip menu	364	6.0%	48	6.1%	187	4.1%	417	5.8%
Like a tip	56	0.9%	1	0.1%	0	0.0%	15	0.2%
Open the leader board menu	520	8.6%	35	4.4%	384	8.3%	1296	18.0%
Open the message menu	428	7.1%	95	12.1%	1389	30.2%	422	5.9%
Open the scores menu	831	13.8%	120	15.2%	553	12.0%	829	11.5%
Open the trip menu	1711	28.4%	193	24.5%	1192	25.9%	1715	23.8%
Total	**6017**		**787**		**4599**		**7208**	

For all countries, we see that the ‘trip’ and ‘scores’ menus were among the most visited. For Belgium and Germany, the next most visited area was the ‘goals’ section, however in Greece the message menu was most visited, and in the UK the leader board was among the most popular areas. It’s interesting to note that trips and scores were available in both Phase 3 and Phase 4, though the previous data suggests they were used more frequently in Phase 4, despite being available earlier. Whereas the goals and leader board were only active for Phase 4, which further explains the higher app engagement seen in that Phase.

3.3.2 Use of technology and user acceptance

At the end of the trial, to assess **system fidelity**, participants were asked seven questions relating to their use of the in-vehicle system (questions UX1-UX7 in Annex 2). Participants were asked to what extent they agreed with the statements, and responded using a 5-point scale, ranging from ‘strongly disagree’ to ‘strongly agree’.

Figure 18 shows the proportion of drivers for each country who responded ‘agree’ or ‘totally agree’ to each question, for Belgium and the UK (these data are not available for Germany, were not applicable in Greece as they relate to the in-vehicle system, and it is also noted a smaller number of Belgian drivers completed this questionnaire). The full tables of responses for each country are given in Annex 1 (Table 63, Table 64).

Results were fairly consistent between the two countries. Generally, participants felt the system was easy to use and the warnings were clear. However, there were lower scores regarding the correctness of the alerts: less than half of UK drivers and only a quarter of Belgian drivers agreed that the alerts correctly reflected the situation. Further, over half of the UK drivers felt that the alerts were sometimes distracting.

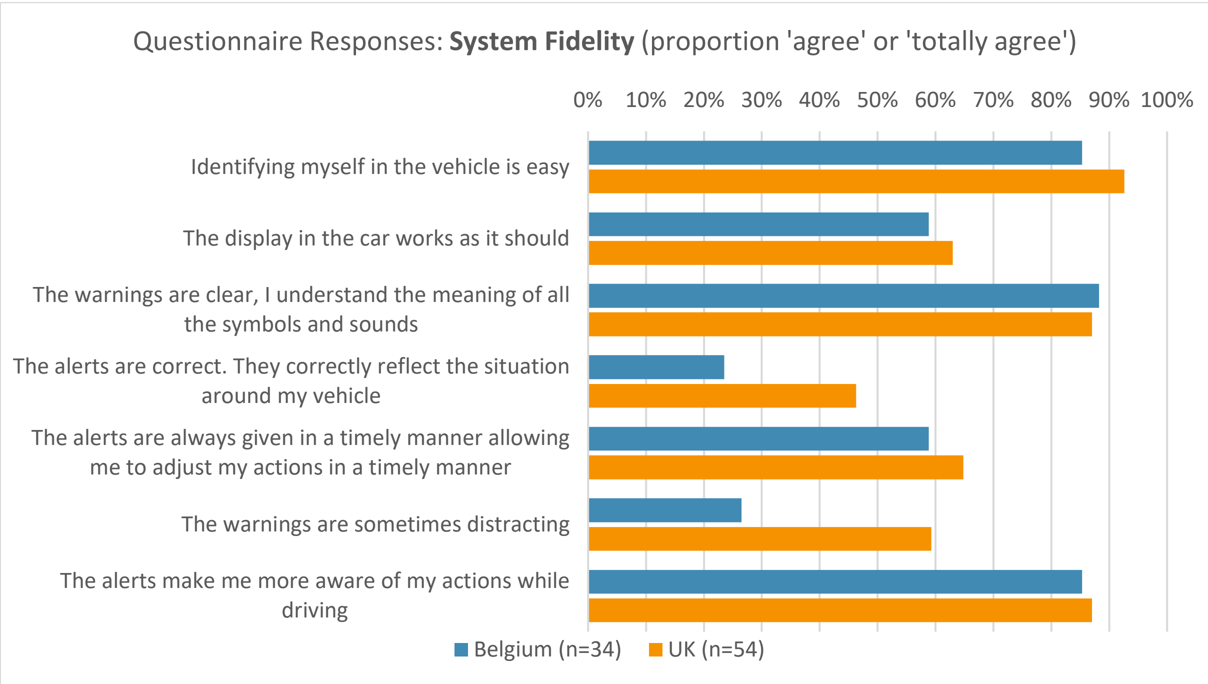


Figure 18: System fidelity questionnaire data - proportion agree / totally agree, per country (cars)

Drivers were also asked to rate how clear they found the overall system, visual symbols and sounds (Figure 21). Belgian and UK drivers rated the system clearer than German drivers, and across all countries drivers responded that the sounds were less clear than the visual symbols.

In Greece, the question was rephrased to relate to the app instead (Table 31). Generally, drivers felt the app was clear, though nearly a quarter felt the scores / rankings weren't clear.

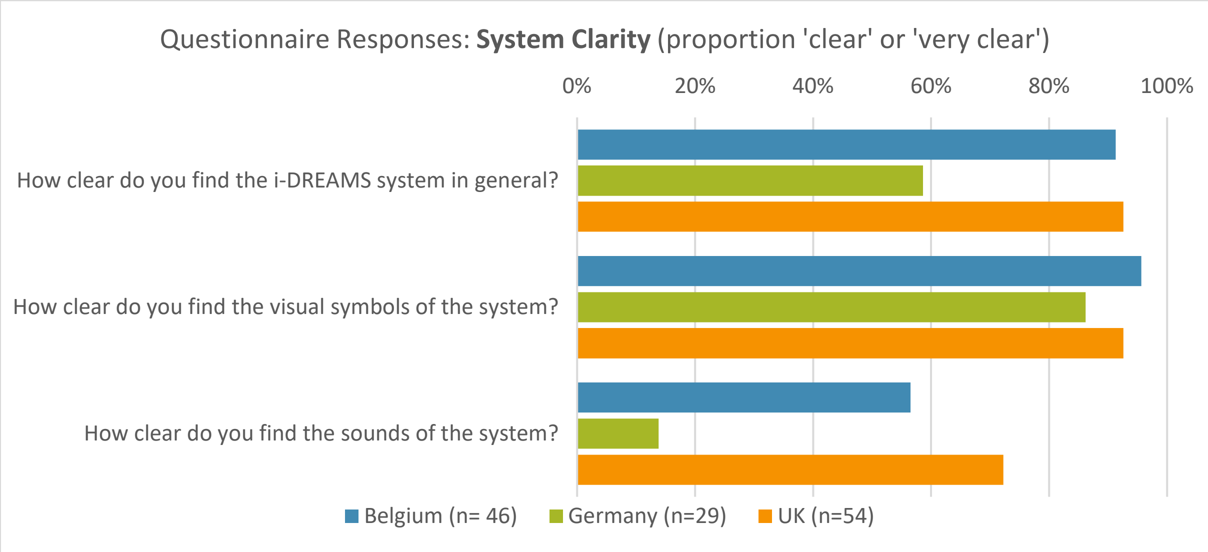


Figure 19: System clarity questionnaire data - proportion replied clear / very clear, per country (cars)

Table 31: App clarity questionnaire data – Greece (n=44)

Question	**Very clear**	**Clear**	**Neutral**	**Unclear**	**Very unclear**
How clear do you find the i-DREAMS app in general?	20.5%	54.5%	18.2%	6.8%	0%
How clear are the app scores and rankings?	15.9%	40.9%	20.5%	20.5%	2.3%
How clear is the feedback provided by the app?	11.4%	45.5%	27.3%	9.1%	6.8%

To assess **user acceptance**, a series of 20 questions were asked of drivers at the end of the trial, to assess their acceptance of the technology (question EX1 in Annex 2). Participants were asked to what extent they agreed with the statements about the i-DREAMS system, and responded using a 5-point scale, ranging from 'strongly disagree' to 'strongly agree'.

Figure 20 shows the proportion of drivers for each country who responded 'agree' or 'totally agree' to each question. The full tables of responses for each country are given in Annex 1 (Table 65, Table 66, Table 67).

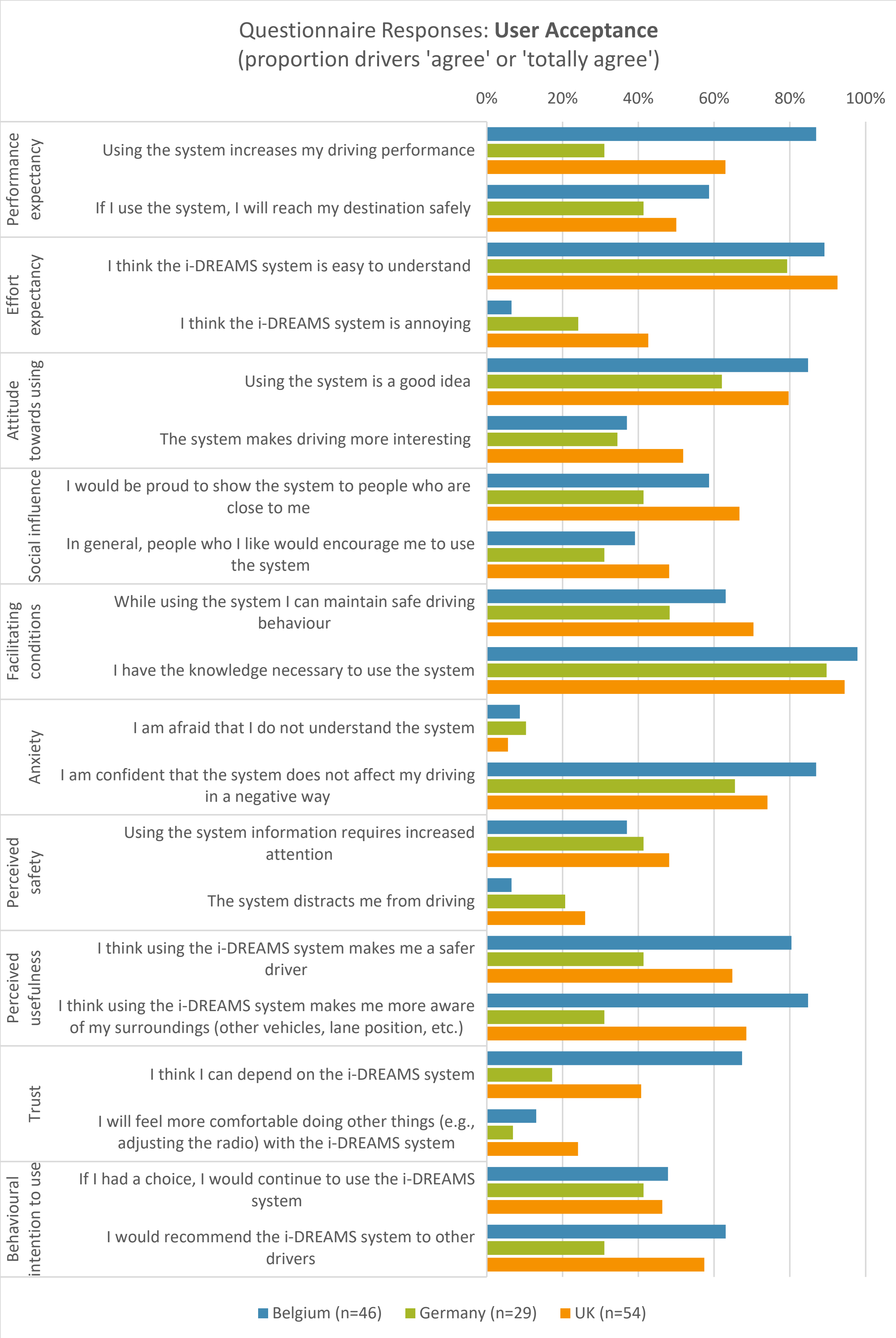


Figure 20: User acceptance questionnaire data - proportion agree / totally agree, per country (cars)

Responses varied for the different countries, with Belgian drivers generally being the most accepting and seeing the benefits, and German drivers being the least accepting. Questions were asked in 10 category areas, the key results are discussed below:

- **Performance expectancy:** Over three quarters of Belgian drivers (87%) and nearly two thirds of UK drivers (63%) felt that the system improved their driving performance. In contrast, less than a third of German drivers (31%) said it improved their performance, but 41% said it would help them reach their destination safely.
- **Effort expectancy (ease of use):** Most drivers (79%-93%) felt that the system was easy to use. Very few Belgian drivers (7%) found the system annoying, however, this increased to a quarter of German drivers (24%) and nearly half of UK drivers (43%). The UK had the highest number of events, and Germany had particularly high numbers of 'speeding' events, so there could be a correlation.
- **Facilitating conditions:** Less than half of German drivers (48%) felt that the system helped them maintain safe driving behaviour, compared with around two thirds of Belgian (63%) and UK (70%) drivers. However, the majority of drivers (90-98%) felt they had the knowledge to use it, re-enforcing the view that it was easy to use.
- **Perceived safety:** Between a third to half of drivers (37% to 48%) felt that using the system required increased attention, and furthermore, a quarter of UK drivers (26%) and a fifth of German drivers (21%) said it was distracting. Only 7% of Belgian drivers felt that the system was distracting. This suggests that, although drivers generally found the system easy to use and understand, some level of demand is imposed on users.
- **Trust:** Trust varied greatly between the different countries: two thirds of Belgian drivers (67%) felt they could depend on the system, but less than half of UK drivers (41%) and less than a quarter of German drivers (17%) said the same. For all countries, less than a quarter (7%-24%) said the system helped them feel more comfortable doing other secondary tasks.
- **Behavioural intention to use:** 40%-50% drivers said they would continue to use the system if they had a choice. This is a bit in contrast with the other results that suggest drivers felt the system improved their performance and was a good idea, but could be partly due to the low trust for some drivers and the feeling that the alerts were not always correct (Figure 18).

Of the 20 questions relating to user acceptance of the i-DREAMS system, eight of them map onto questions asked during the entry questionnaire regarding acceptance of ADAS systems in general (question EQ3 in Annex 2). Converting the responses to a numerical scale (where 1 = strongly disagree and 5 = strongly agree), we can see the before and after averages for each country. (Note that data for Greece are not included, as the questions mostly relate to the in-vehicle technology).

Table 32 gives the results and indicates which changes were significant in the before-after statistical analyses. For most questions, an increase in average score would show improved acceptance, however for questions h and k a decrease in score is an improvement.

Data were not normally distributed, therefore either the Wilcoxon sign-rank test or the paired-sample sign test was used (depending on whether or not the distribution of difference was symmetrical). Results highlighted in blue show where the 'after' value is statistically significantly different to the 'before' value at the α=0.05 level (one further result highlighted in yellow was significant at α=0.1). Complete statistical results (including which test was used for each measure) are given in Annex 1 (Table 68).

Most of the significant changes indicated decreased acceptance of ADAS systems, with all countries having significantly less trust after experiencing the i-DREAMS system. However, for the question about ADAS being clear and understandable, there were significant positive changes in Germany and the UK, and in Belgium there was a significant positive change regarding ADAS being distracting.

Table 32: ADAS acceptance questionnaire data – statistical significance, per country (cars)

Question	BE (n=45)		DE (n=26)		UK (n=54)	
	Before	After	Before	After	Before	After
b. Using ADAS increases my driving performance	4.18	4.09	3.54	2.58	3.48	3.61
c. My interaction with ADAS is clear and understandable	3.91	4.24	3.58	3.96	3.46	4.31
e. Using ADAS is a good idea	4.38	4.00	4.15	3.58	3.76	3.94
f. I can maintain safe driving behaviour while using ADAS	4.33	3.60	3.69	3.35	3.61	3.81
g. I will feel more comfortable doing other things (e.g., adjusting the radio) with ADAS	3.64	2.47	3.00	1.85	3.15	2.87
h. Using ADAS information requires increased attention	3.09	2.89	2.92	2.96	3.28	3.28
j. I trust the information I receive from ADAS	4.04	3.49	3.54	2.15	3.41	3.07
k. ADAS distract me while driving	3.07	2.09	2.69	2.42	3.00	2.78

Finally, drivers were asked to rate their overall experience of the trial (question EX9 in Annex 2). Drivers were asked to rate their experience on a scale from 'very unfavourable' to 'very favourable', the results of which are shown in Figure 21. Drivers in the UK rated their experience better on average, though in all countries over half rated it 'somewhat favourable' or 'very favourable'. A relatively higher proportion of Greek drivers responded neutrally, and a relatively higher proportion of German drivers responded negatively.

In Belgium, the responses were instead on a scale from 'not interesting at all' to 'very interesting'. Results here (Table 33) were very positive, with 93% rating the experience as 'interesting' or 'very interesting'. However, it should be noted that these results are not directly comparable with the other countries data, and that an experience can still be interesting even if it is not enjoyable or favourable.

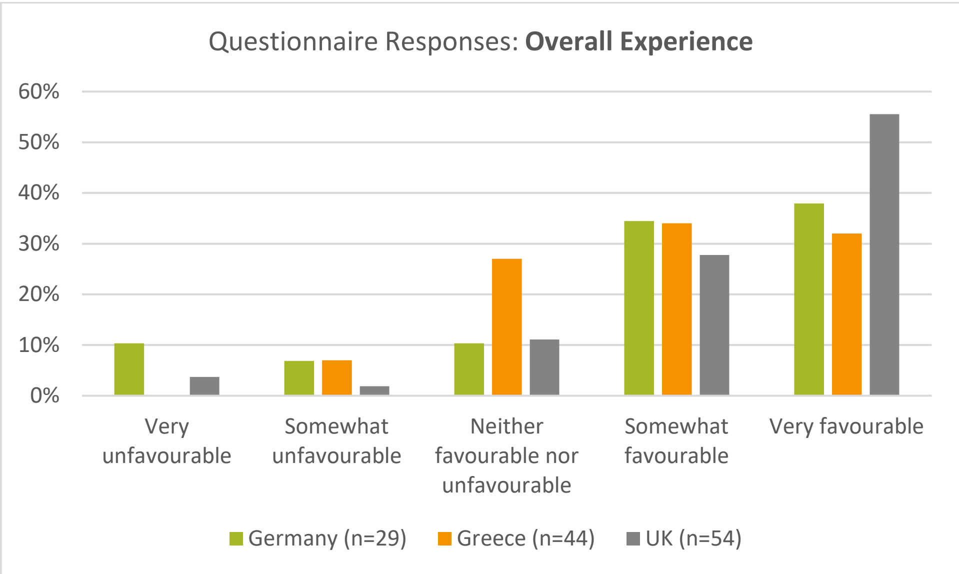


Figure 21: Questionnaire data – overall experience, per country (DE & UK cars)

Table 33: Questionnaire data – overall experience (BE cars)

How would you rate your experiences participating in this study? (BE: n=46)			
Very Interesting	Interesting	No Opinion	*No Response*
52%	41%	4%	2%

3.4 Differences Between Drivers

The previous sections presented results for all drivers per country analysed together. However, there is additional value in analysing differences between drivers, as the previous results showed a large variance in the data, which could indicate that the i-DREAMS technology had a varied effect on different drivers within the sample.

3.4.1 Outcome evaluation

For events data, the average number of total events per phase was calculated for each driver, to determine which drivers showed improvement after exposure to the i-DREAMS technology, and which did not. The change in events from Phase 1 to Phase 4 was used to determine this, i.e., overall change during the trial, rather than each step change between phases. Drivers could then be categorised into two 'change types': 'Type A' – the number of events/100km decreased (i.e., outcome improved), and 'Type B' – the number of events/100km increased (i.e., outcome did not improve).

Table 34 gives the number and proportion of drivers in each change type for each country, and the average event change for these drivers. It also shows this change as the percentage increase / decrease in events from Phase 1 to Phase 4, to give a fairer comparison between countries given the varying SPGs included in the total and varying number of events.

It is noted that Germany has a smaller sample, which is further reduced when split into groups. Therefore, results give an indication of differences but should be interpreted with caution.

Table 34: Total event change per change type, per country (cars)

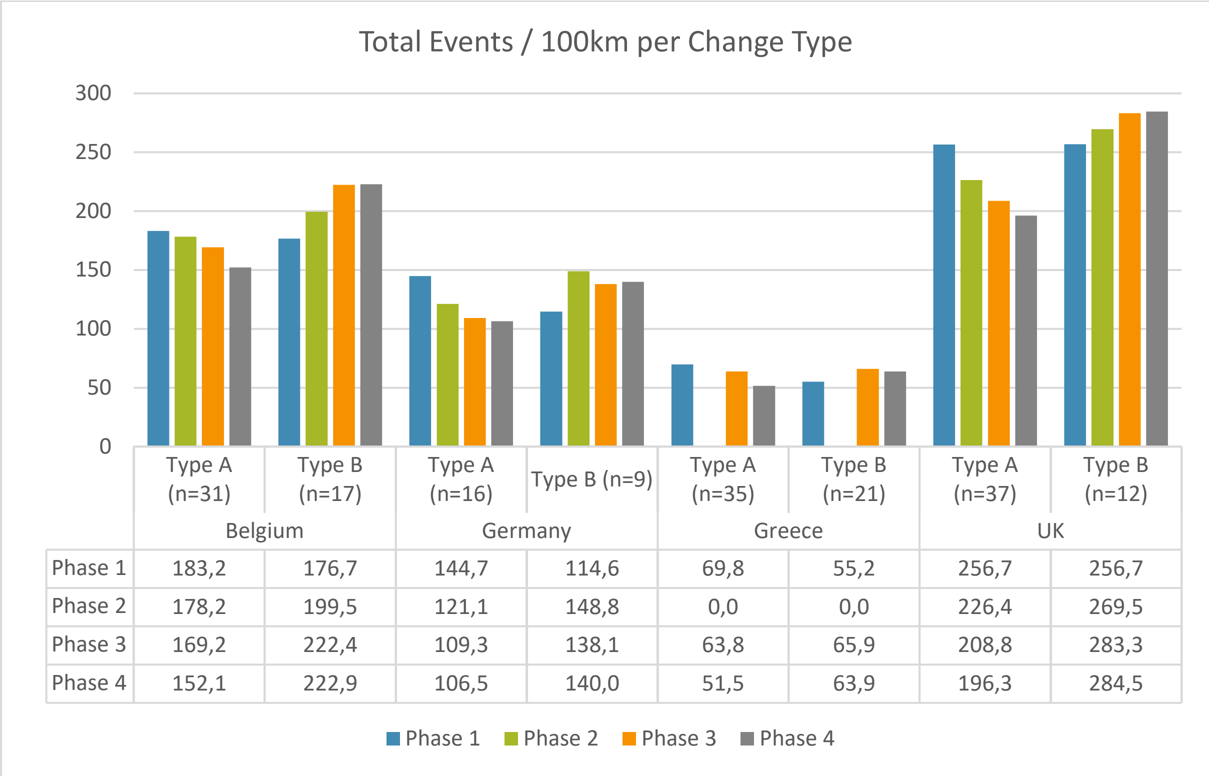
Overall Change in TOTAL Events (Phase 1 – Phase 4)						
Country	Type A - Events/100km Decreased			Type B - Events/100km Increased		
	# Drivers	Average Decrease	Percentage Decrease	# Drivers	Average Increase	Percentage Increase
Belgium	31 (65%)	-31.1	-17.0%	17 (35%)	46.2	26.1%
Germany	16 (64%)	-38.1	-26.4%	9 (36%)	25.4	22.2%
Greece	35 (63%)	-18.3	-26.2%	21 (38%)	8.7	15.8%
UK	37 (76%)	-60.4	-23.5%	12 (24%)	27.8	10.8%

For Belgium, Germany and Greece, around two thirds of drivers showed improved outcomes after exposure to the technology, but in the UK this figure increased to three quarters. Furthermore, UK Type A drivers had the greatest decrease in events, however they also had the largest number of events to begin with; when the percentage decrease is examined, the largest reduction was seen in Germany and Greece.

In Germany, Greece and the UK, the event decrease for Type A drivers was greater than the increase for Type B drivers, i.e., outcomes improved to a greater scale than they worsened. However, the opposite is seen with the Belgian drivers, which may partly explain the results seen for Belgium when all drivers are included (an overall small change in total events).

Figure 22 shows the events / 100km per phase for each group of drivers for each country. For all countries, we can see that where drivers showed improved outcomes (Type A), there was a consistent decrease in events in each successive phase. For Germany and the UK, the biggest reduction appears to be between Phase 1 and Phase 2, whereas in Belgium and Greece it is from Phase 3 to Phase 4. For Type B drivers, there is more fluctuation / plateauing across phases, though trends vary between countries.

It is also interesting to note that in Belgium and the UK, the baseline event numbers were similar for each type of drivers, i.e., they started at a similar level, but some improved and some worsened. In Germany and Greece, Type A drivers actually had more events than Type B in the baseline phase, but then improved and had less in the final phase.



	Belgium Type A (n=31)	Belgium Type B (n=17)	Germany Type A (n=16)	Germany Type B (n=9)	Greece Type A (n=35)	Greece Type B (n=21)	UK Type A (n=37)	UK Type B (n=12)
Phase 1	183,2	176,7	144,7	114,6	69,8	55,2	256,7	256,7
Phase 2	178,2	199,5	121,1	148,8	0,0	0,0	226,4	269,5
Phase 3	169,2	222,4	109,3	138,1	63,8	65,9	208,8	283,3
Phase 4	152,1	222,9	106,5	140,0	51,5	63,9	196,3	284,5

Figure 22: Events/100km per change type, per country per phase (cars)

Finally, to further investigate the differences between Type A and Type B drivers, the demographic factors of each group were compared. These are shown in Annex 2 (Table 69, Table 70, Table 71, Table 72) and summarised here.

- Gender: In Germany and the UK, Type B had a higher proportion of male drivers compared to Type A. Little gender differences in Belgium and Greece.
- Age: In Germany and Greece, Type B drivers were on average older than Type A drivers, whereas in the UK they were younger. Little age differences in Belgium.
- Driving experience: In Greece Type B drivers were more experienced than Type A drivers, whereas in the UK they were less experienced. Little differences in Belgium and Germany.
- Accident history: In Belgium, Greece and the UK, Type B drivers reported less recent accidents than Type A drivers, whereas it was the opposite in Germany. It is also noted that all the participants who had multiple recent accidents were in Type B.
- Offence history: In Germany, Greece and the UK, Type B drivers reported less recent offences than Type A drivers, whereas it was the opposite in Belgium.
- Driving style: In Germany, Type B drivers were both more hesitant and more risk-taking (Type A were more discreet or sportive). In Greece, Type B driver were also more risk-taking, but more discreet (Type A were more sportive). In the UK, Type B drivers were again more risk-taking, and also more sportive (Type A were more discreet and more hesitant). There were little differences in Belgium.
- Driving confidence: In Belgium, Germany and the UK, Type B drivers were typically more confident, and Type A drivers were more neutral (and in the UK Type A were also more insecure). In Greece the opposite is seen.

It is finally noted that for Belgium 'wave 1' drivers, 46% were classified as Type A and 54% as Type B. When looking at 'wave 2', 86% were Type A and only 14% were Type B. This is very different from the UK, which showed a nearly identical split of Type A / Type B drivers in each wave (roughly 75%/25%). In this report, each wave was not analysed separately as theoretically there should have been no difference between them: both waves of participants experienced the same trial design, the multiple waves were only to collect the most data with a limited amount of equipment. However, in regard to the Belgian data, their wave 1 participants experienced the most changes in COVID-19 restrictions, in particular an easing of restrictions as they progressed through the trial, with an associated increase in traffic density. Also, the Belgium wave 1 was the first group to undertake the trial, and the earlier participants of this wave experienced some delays moving between phases as technology issues were identified and resolved. Specifically, these early drivers spent a longer time in Phase 3 than other drivers, and, although the data were cleaned to only include trips for the correct number of days for each phase, these delays could still have had an impact on driver behaviour and could partly explain the results for Belgian drivers.

3.4.2 Process evaluation

To assess how app use varied between drivers, the number of visits per user was identified, and this number was divided by 70[3] (the number of days participants were able to access the app) to give the average number of visits per day per user. Although in reality app use varied across time, this method gives an approximation of daily use. It was then possible to categorise drivers into 'high' (n/70 = greater than or equal to 1), 'average' (n/70 = between 0.5-0.99) or 'low' (n/70 = less than 0.5) users.

Table 35: App usage per usage level, per country (cars)

Usage Level	Belgium (n=49)		Germany (n=23)		Greece (n=71)		UK (n=51)	
	N	Visits	N	Visits	N	Visits	N	Visits
High	16 (33%)	1683 (61%)	0 (0%)	0 (0%)	9 (13%)	804 (55%)	14 (27%)	2662 (74%)
Average	14 (29%)	762 (28%)	3 (13%)	112 (33%)	8 (11%)	266 (18%)	10 (20%)	523 (15%)
Low	19 (39%)	323 (12%)	20 (87%)	230 (67%)	54 (76%)	393 (27%)	27 (53%)	409 (11%)

It can be seen that, although the UK had most app visits overall, just over half the drivers were in the 'low' use category. Furthermore, just over a quarter of the drivers (27% 'high' users) accounted for nearly three quarters of the total visits. Regarding Belgian drivers, there was a more even split between the usage categories, though a slightly higher proportion of 'low' users than 'average' users. In Greece, the vast majority (75%) were 'low' users, and just 13% of users ('high' users) accounted for half of the visits. For Germany there were no 'high' users, and the great majority were 'low' users, which is to be expected given the overall low app usage for German drivers.

[3] For Greece, the number of visits was divided by 51 instead, as data had to be exported before the end of the trial

Finally, a brief analysis was carried out to see if there was any link between app usage and whether or not the driver's outcome improved. (Note that the sample size differs here as to be counted drivers must have been included in both the outcome and process analyses).

Table 36: Driver app usage level vs improvement type, per country (cars)

Change in Total Events/100km		**App Usage Level**		
		Low (N)	Average (N)	High (N)
BE	Type A (# decreased)	10	6	15
	Type B (# increased)	8	7	1
DE	Type A (# decreased)	11	1	0
	Type B (# increased)	6	2	0
EL	Type A (# decreased)	10	6	15
	Type B (# increased)	8	7	1
UK	Type A (# decreased)	15	8	13
	Type B (# increased)	9	1	1

It can be seen that, for drivers in Belgium, Greece and in the UK, the great majority of drivers in Type B (outcome did not improve) had 'low' or 'average' app usage. Indeed, for each country only one Type B driver had 'high' app usage. Furthermore, almost all 'high' app users had improved outcomes. Results for Germany are less clear, however, given the overall low app usage it is difficult to draw conclusions for these drivers.

4 Results – Trucks

Section 4 presents the analysis results for truck drivers. In the following analyses the sample size may vary, as some drivers are excluded from certain analyses (for example if they had no trips in one or more data collection phases, missing questionnaire data, etc.). All analyses will state the sample used.

4.1 Data Sample

Truck data were analysed for Belgium (BE) and included drivers from five companies. The participant sample is described below. All drivers were male, though this is not unusual.

Table 37: Trucks sample

		Belgium
Number of participants (drivers)		40
Participant gender	Male	40 (100%)
	Female	0 (0%)
Participants mean age (years)		44.9
Standard deviation of age (years)		11.5
Years driving experience (range, average)		2 - 47, 20.2

Before taking part in the trial, drivers were asked a number of questions, including about their previous accidents and offences, and how they would describe their driving style and confidence. The responses to these are shown in Table 38 below. Most drivers were 'confident' or 'very confident', and all described themselves as either 'average' or 'sportive'. Over a third of drivers had been involved in an accident in the last three years, and nearly three quarters had a recent traffic offence, mostly speeding offences. However, it is likely that they also have greater exposure compared to the average car driver.

Table 38: Driver accident and offence history, and confidence (trucks)

Question / Response Option		BE (n=40)
In the last three years, have you been involved in an accident, which was self-inflicted?	No	60%
	Yes, once	35%
	Yes, twice	5%
Within the last three years, have you been fined for a traffic offence while driving your truck? (Excluding parking offences)	No	28%
	Once - speeding	43%
	Once - using phone	3%
	Twice - speeding + using phone	10%
	Twice - speeding + illegal overtaking	5%
	Twice - speeding + running a red light	3%
	Twice - speeding + tailgating	3%
	Three times – (NS, speeding + using phone + illegal overtaking / running stop sign)	9%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	0%
	Discreet, average	75%
	Sportive, ambitioned	28%
	Risk-taking, offensive	0%
How confident you are concerning your own driving skills?	Insecure	0%
	Neutral	18%
	Confident	63%
	Very confident	20%

Drivers were also asked how often they believed they engaged in certain risk-taking behaviours. Participants were asked to estimate how often they engaged in these behaviours over the previous year and responded using a 5-point scale of: ‘almost always’, ‘on a regular basis’, ‘often’, ‘sometimes’ and ‘never’.

Truck drivers reported relatively less risk-taking behaviour when compared with car drivers. Numbers were highest for ‘driving faster than the limit’ and ‘use a hand-held mobile phone while driving’, which correlates with the reported offences.

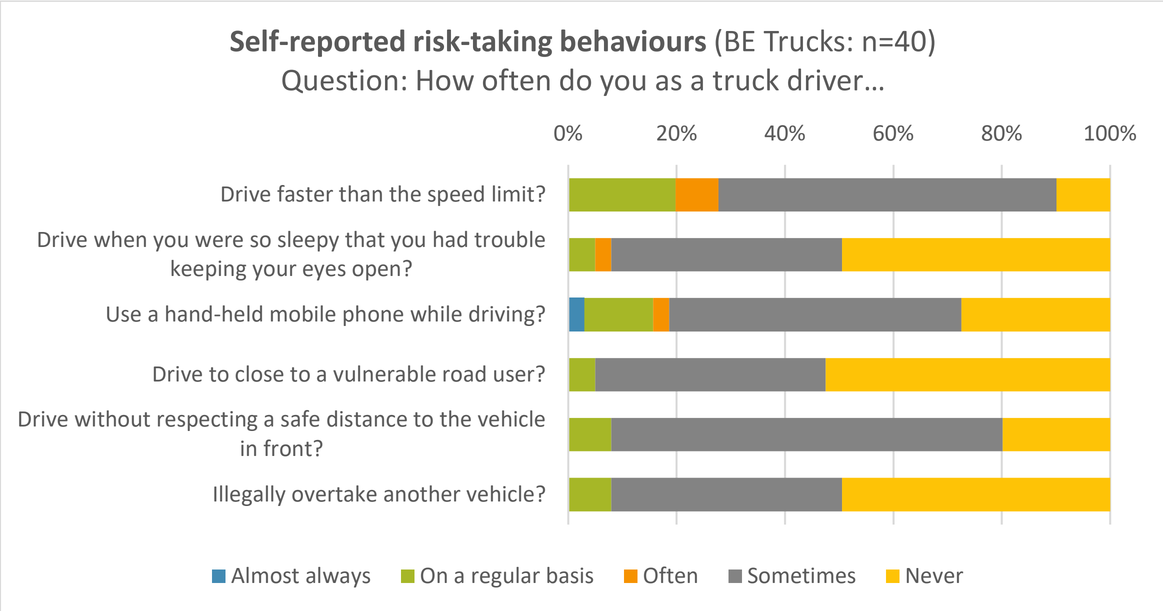


Figure 23: Questionnaire data – self-reported risk-taking behaviours, (Belgian trucks)

4.1.1 Results overview

The table below describes the valid trip data available for analysis of Belgian truck data, and also gives overview results for each data collection phase. Participants were excluded if trips were not present for all phases, and outlier trips were removed as described in section 2.3.1.

Table 39: Results overview – trucks

Data Collection Phase	**Belgium Trucks (n=37 drivers)**					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	2,680	199,463	90.77	41.44	88.96	4.39
Phase 2	3,180	206,183	92.02	42.10	89.23	4.11
Phase 3	3,185	228,774	92.14	39.25	88.43	4.09
Phase 4	3,716	265,735	90.25	35.41	87.98	4.58
TOTAL	12,761	900,155	91.30		88.65	

It is noted that the total number of trips is similar to the car data from Belgium, whereas the distance travelled is substantially higher. This would be expected as truck drivers typically undertake longer journeys.

Regarding the number of events / 100km, the total events is roughly half the number seen for Belgian car drivers, and there is a much smaller standard deviation. However, the results show very little change throughout the trial, with event numbers only rising slightly then falling back to baseline (Phase 1) levels.

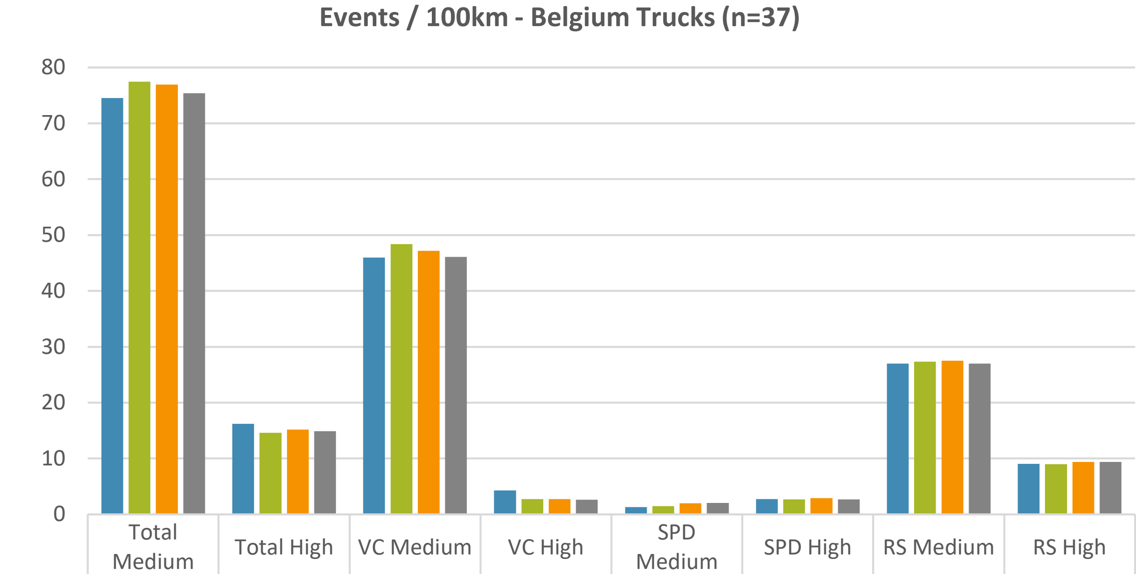
4.2 Outcome Evaluation

4.2.1 Events and scores analysis

A more detailed breakdown of the events and scores for truck drivers is shown below, providing results for 'total', 'vehicle control' (VC), 'speeding' (SPD) and 'road sharing' (RS) events. Figure 24 further shows the split of 'medium' and 'high' severity events.

Table 40: Events / 100km and scores per event type and per phase (trucks)

Phase	Belgium Trucks (n=37)							
	Total Events / 100km	Overall Scores	VC Events / 100km	VC Scores	SPD Events / 100km	SPD Scores	RS Events / 100km	RS Scores
Phase 1	90.77	88.96	50.23	75.40	4.08	94.12	36.05	94.49
Phase 2	92.02	89.23	51.13	75.25	4.15	94.01	36.33	94.79
Phase 3	92.14	88.43	49.96	75.49	4.90	90.92	36.93	94.44
Phase 4	90.25	87.98	48.74	75.17	4.78	89.81	36.39	94.47



	Total Medium	Total High	VC Medium	VC High	SPD Medium	SPD High	RS Medium	RS High
Phase 1	74,54	16,22	45,95	4,28	1,33	2,75	27,02	9,02
Phase 2	77,43	14,59	48,39	2,74	1,47	2,69	27,32	9,01
Phase 3	76,93	15,21	47,20	2,76	1,98	2,92	27,51	9,41
Phase 4	75,38	14,87	46,10	2,63	2,07	2,71	27,00	9,39

Figure 24: Medium and high events / 100km per event type and per phase (trucks)

Changes were generally very small, and patterns varied between the difference SPGs. There was an overall decrease in 'vehicle control' events, and an overall increase in 'speeding' and 'road sharing' events, but the significance of these changes is discussed further below. Again, it is noted that event numbers were substantially lower than for Belgian car drivers, particularly for 'high' severity 'speeding' events.

Table 41 gives a summary of the change in events for each event type and level of severity, and between each phase. The table first gives the significance value for the overall test (Friedman test or repeated measures ANOVA depending on normality of data). Then, pairwise comparisons between each phase show the direction (if the number of events/100km decreased ↓ or increased ↑), as well as the significance value from the relevant statistical test (Wilcoxon signed rank / paired t-test).

Results highlighted in grey indicate normally distributed data, therefore parametric tests were used. Results highlighted in blue were significant at the α=0.1 level, which has been chosen due to the small sample size.

Table 41: Event change summary and statistical significance, per event type & severity and per phase – trucks

Change in Number of Events: Belgium Trucks (n=37 drivers)											
Event Type		**Friedman / ANOVA test significance**	**Overall Change (Phase 1 – Phase 4)**		**Change Between Phases**						
					P1 - P2		**P2 - P3**		**P3 - P4**		
Total	**Medium**	p = 0.812	↑	0.572	↑	0.667	↓	0.492	↓	0.723	
	High	p = 0.254	↓	0.551	↓	0.021	↑	0.163	↓	0.354	
	All	p = 0.901	↓	1.000	↑	1.000	↑	1.000	↓	1.000	
Vehicle Control	**Medium**	p = 0.965	↑	0.712	↑	0.678	↓	0.839	↓	0.910	
	High	p = 0.554	↓	0.551	↓	0.202	↑	0.898	↓	0.910	
	All	p = 0.981	↓	0.982	↑	0.803	↓	0.874	↓	0.994	
Speeding	**Medium**	p = 0.074	↑	0.015	↑	0.823	↑	0.054	↑	0.719	
	High	p = 0.787	↓	0.780	↓	0.718	↑	0.815	↓	0.769	
	All	p = 0.446	↑	0.108	↑	0.753	↑	0.394	↓	0.827	
Road Sharing	**Medium**	p = 0.930	↓	1.000	↑	1.000	↑	1.000	↓	1.000	
	High	p = 0.330	↑	0.660	↓	0.826	↑	0.096	↓	0.950	
	All	p = 0.429	↑	0.637	↑	0.765	↑	0.245	↓	0.470	

It can be seen that results were mixed, especially for the overall change (Phase 1 to Phase 4), and few changes were statistically significant. The increase in 'medium' severity 'speeding' events was significant, both overall and from Phase 2 to Phase 3, as was the increase in 'high road sharing' events from Phase 2 to Phase 3. The only decrease that was statistically significant was for 'total high' events from Phase 1 to Phase 2.

Table 42 shows the **Fatigue** events for Belgian truck drivers. More events were recorded than for car drivers, however numbers are still very low, and the data shows a consistent decrease across the data collection phases. None of the changes here were statistically significant.

Table 42: Fatigue events / 100km and scores per phase (trucks, n=37)

Phase	**Fatigue Events / 100km (All Severities)**	**Fatigue scores**	**Medium Fatigue Events / 100km**	**High Fatigue Events / 100km**
Phase 1	0.4118	91.82	0.2491	0.1626
Phase 2	0.4034	92.88	0.2600	0.1434
Phase 3	0.3602	92.86	0.2399	0.1204
Phase 4	0.3511	92.45	0.2078	0.1432

Valid **distraction** data were only available for two truck drivers; therefore, no results are presented here.

4.2.2 Questionnaire analysis

A set of 12 questions were asked identically at both trial entry and trial exit (respectively EQ11 and EX3 in Annex 2), to allow analysis of before and after responses. These questions assess the change objectives level of the logic model of change, and relate to the areas of perceived knowledge', 'self-efficacy', 'attitude', 'personal norm', and 'subjective norm', which are described in more detail in section 2.2.

Participants were asked to respond using a 5-point Likert scale, where 1 = strongly disagree and 5 = strongly agree. The average score for each question, both before and after, is given in the table below.

Table 43: Questionnaire data – outcome evaluation (trucks)

Evaluation Measure		BE Trucks (n=9)	
		Before	After
Perceived Knowledge	I know the benefits of safe driving	3.89	4.11
	I know what is needed to drive safely	4.00	4.11
Self-Efficacy	I have the skills to drive safely	3.78	3.78
	I feel competent enough to drive safely	4.00	3.89
	I control whether I drive safely or not	3.56	3.33
	For me, safe driving is easy to do	3.67	3.67
Attitude	Safe driving is important to avoid crashes	4.22	4.22
	Safe driving makes me feel comfortable	4.00	4.00
Personal Norm	For me personally, safe driving is important	4.11	4.11
	Safe driving should be a personal obligation	4.11	4.22
Subjective Norm	My friends think safe driving is important	3.56	3.67
	My colleagues find it important to drive safely	3.56	3.78

It is noted that the sample here is extremely small; exit questionnaire data were only available for nine truck drivers. The results are presented for completeness and to give a tentative indication of opinion change with respect to socio-cognitive dispositions towards safety.

4.3 Process Evaluation

The process evaluation results in this section include all truck drivers who engaged with the i-DREAMS app. Therefore, the sample size is different from the results above. The app was available to participants in Phase 3 and Phase 4 of the study, with additional functions activated during Phase 4, as described in section 2.

Unfortunately, very few truck drivers completed the exit questionnaires, therefore no results are presented in relation to use of technology and user acceptance.

4.3.1 App usage

For Belgian trucks, 24 drivers engaged with the i-DREAMS app, with a total of 1,255 visits. It is noted that one of the companies did not allow the app to be installed on drivers' work phones, which is why the sample is lower. Although there are fewer users, the average visits per driver is roughly similar to that seen for Belgian car drivers.

Figure 25 shows the total app visits on each day, to show how app use varied throughout the trial (remembering that for the first eight weeks the app was not active, therefore app use started on day 57, and Phase 4 had a longer duration than Phase 3).

As with car drivers, app usage was higher in Phase 4 compared to Phase 3. However, for truck drivers the usage throughout Phase 4 did not decrease to the same extent as with car drivers. Although app usage fluctuates, on average it is more consistent.

Figure 26 further shows the total visits per day of the week and per time of the day. For truck drivers, usage was highest on Wednesday and noticeably lower at the weekends. Regarding time of day, the data shows the expected peaks in the morning and at midday, coinciding with the push notifications, however the evening peak is not present. This could be due to the nature of shift working with truck drivers, and perhaps in the evening they were either driving or were sleeping for an early shift.

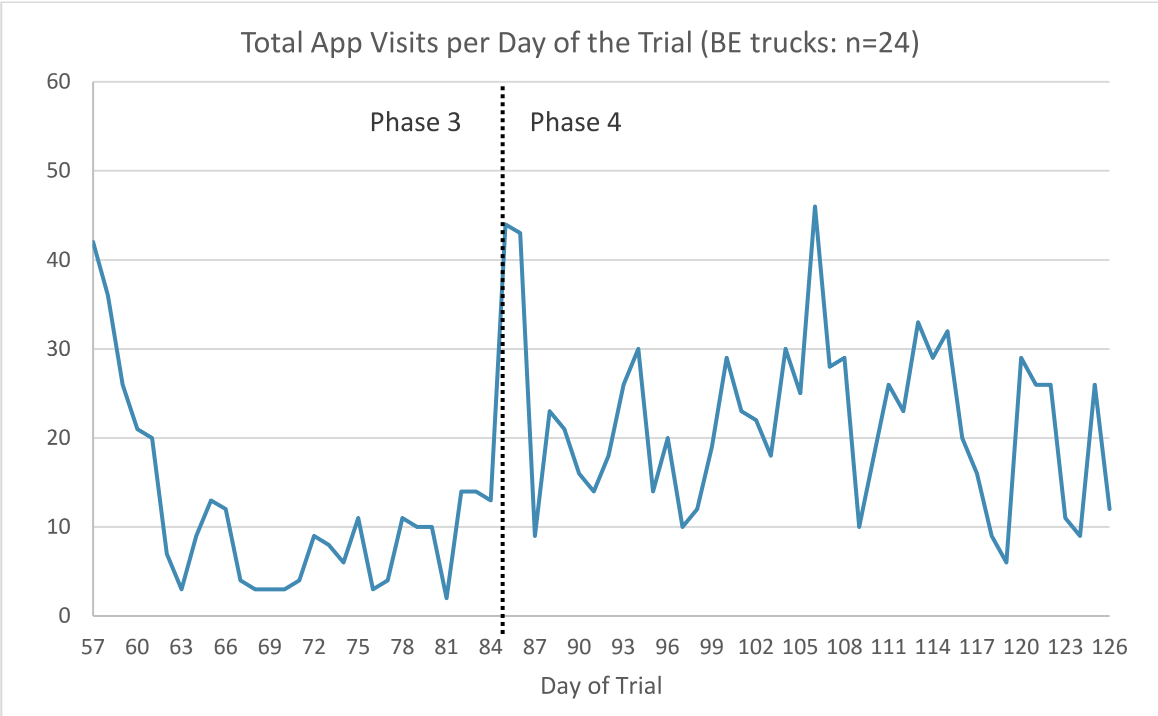


Figure 25: Total app visits per country, per day of trial (trucks)

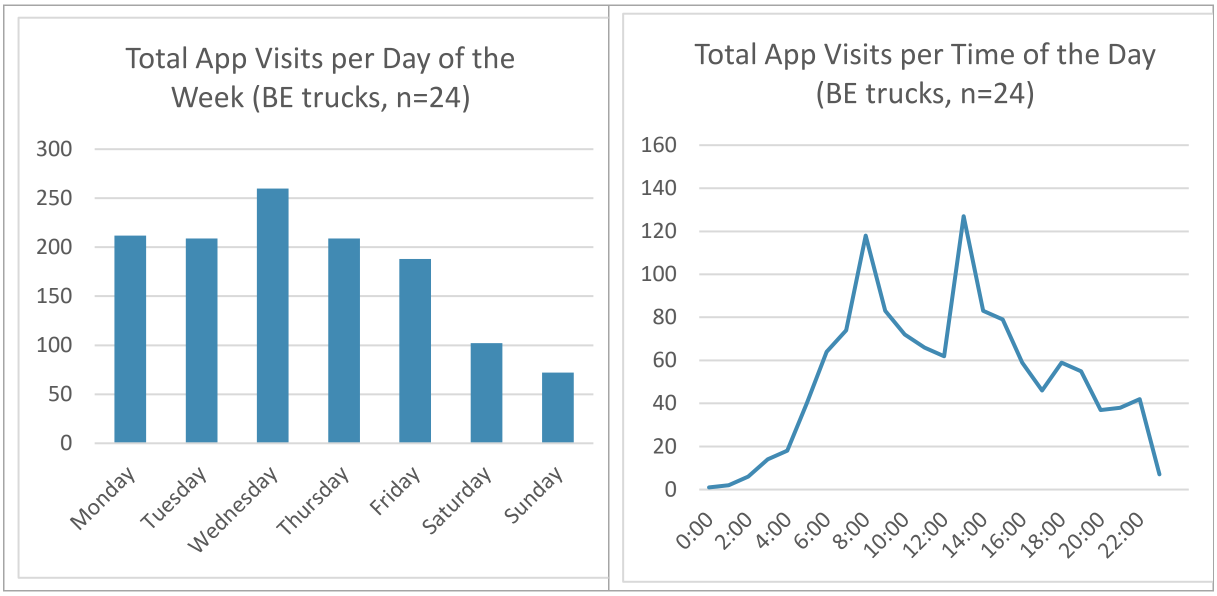


Figure 26: Total app visits per day of the week and time of day (trucks)

Finally, the app use data were analysed to determine how frequently drivers used different app functions. These results are given below in Table 44, and the three most popular functions are highlighted (note that the total in this table is higher than the total visits stated above, since a user could visit multiple areas of the app within the same 'visit'). The results show that the 'leader board' and 'trip' menus were most popular with Truck drivers, which is similar to results seen for car drivers. However, in contrast, the 'fact' menu was the third most popular here, and the 'goal' and 'scores' menus were less visited.

Table 44: App functionalities used (trucks)

App functionality	Belgium (n=24)	
	N	%
Open the trend menu	14	0.7%
Open the goal menu	136	6.8%
Join a goal	30	1.5%
Open the con menu	115	5.8%
Dislike a con	1	0.1%
Like a con	3	0.2%
Open the fact menu	211	10.6%
Like a fact	2	0.1%
Open the pro menu	121	6.1%
Like a pro	17	0.9%
Open the tip menu	161	8.1%
Like a tip	40	2.0%
Open the leader board menu	415	20.8%
Open the message menu	135	6.8%
Open the scores menu	185	9.3%
Open the trip menu	409	20.5%
Total	**1995**	

4.4 Differences Between Drivers

The overall results for truck drivers showed a lower standard deviation, however differences between drivers are still analysed here to help better understand the overall results.

4.4.1 Outcome evaluation

For events data, the average number of total events per phase was calculated for each driver, to determine which drivers showed improvement after exposure to the i-DREAMS technology, and which did not. The change in events from Phase 1 to Phase 4 was used to determine this, i.e., overall change during the trial, rather than each step change between phases. Drivers could then be categorised into two 'change types': 'Type A' – the number of events/100km decreased (i.e., outcome improved), and 'Type B' – the number of events/100km increased (i.e., outcome did not improve).

Table 45 gives the number and proportion of drivers in each change type, the average event change for these drivers, and this change as the percentage increase / decrease in events from Phase 1 to Phase 4. Then, for each type of group of drivers, the number of total events in each phase is given.

For truck drivers, just under half of drivers showed improved outcomes (Type A), which is a lower proportion than was seen for car drivers. The percentage increase / decrease was similar for Type A and Type B, i.e., outcomes improved on a similar scale as they worsened.

Where drivers showed improved outcomes, there was a consistent decrease in events in each successive phase, although the change was small between Phase 2 and Phase 3.

It is interesting to note that the Type A drivers had substantially more events in the baseline phase (Phase 1) compared with Type B drivers. Furthermore, although Type B drivers had an increase in events, they still had fewer events than Type A drivers in Phase 4).

©i-DREAMS, 2019

Table 45: Event change per change type, per phase – BE trucks

Event Change – TOTAL Events (Belgium Trucks)						
Overall Change (Phase 1 – Phase 4)	**Type A - Events/100km Decreased**			**Type B - Events/100km Increased**		
	# Drivers	**Average Decrease**	**Percentage Decrease**	**# Drivers**	**Average Increase**	**Percentage Increase**
	16 (43%)	-17.5	-15.0%	21 (57%)	12.4	17.5%

Number of Events per Phase		
Phase	**Type A (n=16)**	**Type B (n=21)**
1	116.5	71.1
2	111.1	77.4
3	110.9	77.8
4	99.0	83.6

4.4.2 Process evaluation

To assess how app use varied between drivers, the number of visits per user was identified, and this number was divided by 70 (the number of days participants were able to access the app) to give the average number of visits per day per user. Although in reality app use varied across time, this method gives an approximation of daily use. It was then possible to categorise drivers into 'high' (n/70 = greater than or equal to 1), 'average' (n/70 = between 0.5-0.99) or 'low' (n/70 = less than 0.5) users. Also, a brief analysis was carried out to see if there was any link between app usage and whether or not the driver's outcome improved.

Table 46: App usage per usage level and driver change type (trucks)

Usage Level	BE Trucks (n=24)				App Usage Level by Driver Change Type (N)	
	N	**%**	**Visits**	**%**	**Type A (# Decreased)**	**Type B (# Increased)**
High	5	21%	811	65%	3	2
Average	6	25%	288	23%	2	3
Low	13	54%	156	12%	2	7

Just over half the drivers were in the 'low' usage category. Furthermore, just a fifth of the drivers were responsible for two thirds of the app visits. In terms of the link between app use and outcome, results are not as clear as they were for car drivers, though proportionally few of the Type B drivers (whose outcomes worsened) had 'high' app use.

Further investigation could be done to understand why app use was low, for example, is it related to working conditions, companies' safety climate, etc.?

5 Results – Buses

Section 5 presents the analysis results for bus drivers. In the following analyses the sample size may vary, as some drivers are excluded from certain analyses. All analyses will state the sample used.

As bus drivers did not use the i-DREAMS app, results are presented only for outcome evaluation, for Phase 1 and Phase 2. Furthermore, at the time of writing the questionnaire data for bus drivers were not available, therefore no questionnaire results or driver demographics included here.

5.1 Outcome Evaluation

5.1.1 Results overview

Bus data were analysed for Portugal (PT). The table below describes the valid trip data available for analysis, and also gives overview results for each data collection phase. Participants were excluded if trips were not present for both phases, and outlier trips were removed as described in section 2.3.1. A sample of 21 drivers remained, which is relatively small, therefore care should be taken when interpreting results.

Table 47: Results overview – buses

Data Collection Phase	**Portugal Buses (n= 21 drivers)**					
	Number of Trips	Distance Travelled (km)	Number of Events / 100km	Standard Deviation of Events / 100km	Overall Average Score	Standard Deviation of Average Score
Phase 1	1,252	61,221	295.00	166.45	75.44	8.72
Phase 2	922	49,391	334.41	184.61	73.66	8.76
TOTAL	**2,174**	**110,613**	**311.72**		**74.76**	

The number of trips per phase was lower than Truck drivers, but similar to German car drivers (which also had a smaller sample). The average distance per trip was higher than seen for cars however, which is to be expected from professional drivers.

The number of events / 100km is higher than all the other transport modes, and three times higher than for truck drivers, which is surprising. There was also increase in events from Phase 1 to Phase 2.

5.1.2 Events and scores analysis

A more detailed breakdown of the events and scores for bus drivers is shown below, providing results for 'total', 'vehicle control' (VC), 'speeding' (SPD) and 'road sharing' (RS) events. Figure 24 further shows the split of 'medium' and 'high' severity events.

Table 48: Events / 100km and scores per event type and per phase (trucks)

Phase	**Portugal Buses (n= 21 drivers)**							
	Total Events / 100km	**Overall Scores**	**VC Events / 100km**	**VC Scores**	**SPD Events / 100km**	**SPD Scores**	**RS Events / 100km**	**RS Scores**
Phase 1	295.00	75.44	214.75	43.90	18.99	85.91	41.60	77.56
Phase 2	334.33	73.66	250.75	41.08	16.78	85.28	45.74	75.24

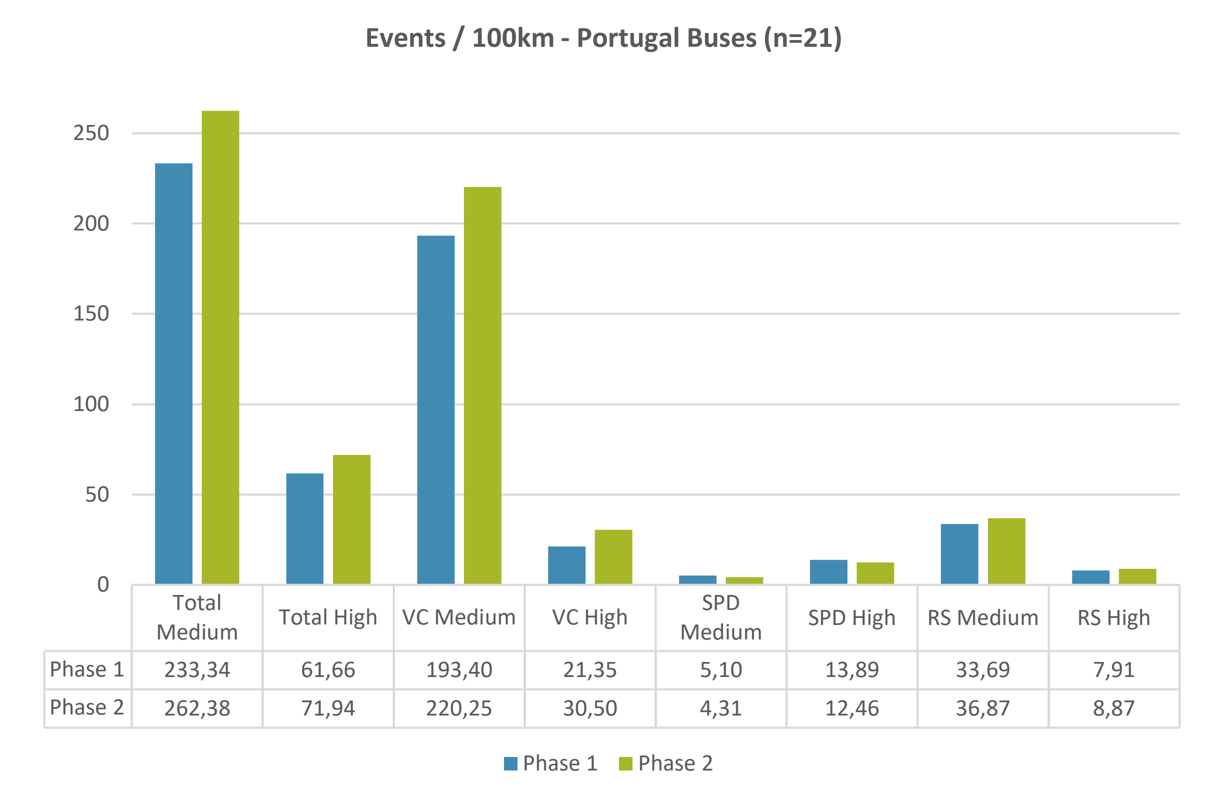


Figure 27: Medium and high events / 100km per event type and per phase (buses)

The vast majority of events were related to 'vehicle control', and again there is a greater number of these compared to the other transport modes. The number of 'speeding' events is similar to the data for Belgian and UK drivers, which is again surprising, given the expectation that professional drivers conform more to traffic rules. However, the number of 'road sharing' events is lower compared to car drivers and is more similar to Truck drivers' numbers.

Overall, there was an increase in 'vehicle control' events, which is the main contributor the increase in 'total' events. There was an overall decrease in 'speeding' events, and overall increase in 'road sharing' events, however these changes were small - their significance is discussed below.

Table 49: Event change summary and statistical significance, per event type & severity and per phase – buses

Change in Number of Events: Portugal Buses (n=21)			
Event Type		**Overall Change (Phase 1 – Phase 2)**	
Total	**Medium**	↑	0.280
	High	↑	0.830
	All	↑	0.140
Vehicle Control	**Medium**	↑	0.130
	High	↑	0.190
	All	↑	0.089
Speeding	**Medium**	↓	0.430
	High	↓	0.950
	All	↓	0.950
Road Sharing	**Medium**	↑	0.240
	High	↑	0.200
	All	↑	0.230

Table 49 gives a summary of the change in events for each event type and level of severity. As only two phases are analysed, Wilcoxon signed rank tests were used to compare data from each phase, and the table shows the direction (if the number of events/100km decreased ↓ or increased ↑), as well as the significance value. Results highlighted in blue were significant at the α=0.1 level, which has been chosen due to the small sample size.

The only statistically significant result for bus drivers was the increase in 'vehicle control' events when both severities are included.

Table 42 shows the **Fatigue** events for bus drivers. More fatigue events were recorded compared to other transport modes, though numbers are still very low. There was an overall decrease in fatigue events, however it was not statistically significant.

Table 50: Fatigue events / 100km and scores per phase (buses, n=21)

Phase	Fatigue Events / 100km (All Severities)	Fatigue scores	Medium Fatigue Events / 100km	High Fatigue Events / 100km
Phase 1	1.37	94.38	1.15	0.22
Phase 2	1.21	93.05	0.95	0.26

As the bus drivers did not use the i-DREAMS app, distraction data are not available.

5.1.3 Differences Between Drivers

The average number of total events per phase was calculated for each driver, to determine which drivers showed improvement after exposure to the i-DREAMS technology, and which did not. As bus drivers did not use the app, the change in events from Phase 1 to Phase 2 was used to determine this. Drivers could then be categorised into two 'change types': 'Type A' – the number of events/100km decreased (i.e., outcome improved), and 'Type B' – the number of events/100km increased (i.e., outcome did not improve).

Table 45 gives the number and proportion of drivers in each change type, the average event change for these drivers, and this change as the percentage increase / decrease in events from Phase 1 to Phase 2.

Table 51: Event change per change type, per phase – PT buses

<table>
<tr><th colspan="7">Event Change – TOTAL Events (Portugal Buses)</th></tr>
<tr><th rowspan="3">Overall Change (Phase 1 – Phase 2)</th><th colspan="3">Type A - Events/100km Decreased</th><th colspan="3">Type B - Events/100km Increased</th></tr>
<tr><th># Drivers</th><th>Average Decrease</th><th>Percentage Decrease</th><th># Drivers</th><th>Average Increase</th><th>Percentage Increase</th></tr>
<tr><td>7 (33%)</td><td>-22.5</td><td>-8.1%</td><td>14 (67%)</td><td>55.0</td><td>18.0%</td></tr>
<tr><th colspan="7">Number of Events per Phase</th></tr>
<tr><th>Phase</th><th colspan="3">Type A (n=16)</th><th colspan="3">Type B (n=21)</th></tr>
<tr><td>1</td><td colspan="3">277.3</td><td colspan="3">305.0</td></tr>
<tr><td>2</td><td colspan="3">254.9</td><td colspan="3">359.9</td></tr>
</table>

For bus drivers, only a third of drivers showed improved outcomes, which is a lower proportion than for car and truck drivers. Furthermore, the increase in events for Type B drivers was greater than the decrease in events for Type A drivers.

It is noted here that the Portuguese sample was small, and when split into groups the sample is even smaller, so these results are provided more to give an indication of the sample.

6 Method and Results – Rail

The rail mode, incorporating both trains (heavy rail) and trams (light rail) was included in i-DREAMS to broaden the application of the i-DREAMS platform which was originally designed for use in road vehicles. The i-DREAMS platform could not be directly applied to trains due to the differences in operation. For example, train drivers to not employ line of sight driving, instead signals are used to manage crossings and intersections. The train mode has therefore been studied within the context of the transferability of the i-DREAMS platform to other modes and information can be found in the Deliverable 8.1 (Lourenco et al., 2023).

In contrast, trams operate within a mixed-traffic environment, driving on both segregated track, and shared, multi-user road. Therefore, aspects of the i-DREAMS platform can be applied to trams and may be beneficial to tram driving safety and risk mitigation. Two main studies were carried out to assess the use of the i-DREAMS platform in trams. The first was a simulator study to test the real-time element of the platform and the second was a focus group study to assess the potential use of the post-trip feedback app in the tram context.

6.1 UK Tram Simulator Trial – Effectiveness of Real Time Interventions

The tram simulator trial took place on the premises of a UK tram operator using their training simulator during July and August 2021. The training simulator had the routes normally driven by the tram drivers including the signal locations, tramstops and intersecting roads pre-programmed along with a number of events designed to test the drivers' competencies. The study design had to take into account these constraints as it was not possible to add to these pre-set programs. However, the benefit of this was that it was possible to test the i-DREAMS platform within a realistic as possible simulated environment. The i-DREAMS interventions that were the focus of the study were Vulnerable Road User (VRU) detection and speeding, and how the i-DREAMS equipment might help with these aspects of driving performance. Fatigue[4] was also featured in the study, through use of the i-DREAMS wearable, questionnaire items, and discussions with drivers. The following sections produce a brief overview of the data collection and analysis methods employed. A more detailed description of these is included in Annex 3.

6.1.1 Method

The trial had three elements. There was a traditional simulator study where a familiarisation drive was followed by a baseline and then an intervention drive. This was supplemented by a 'manipulation drive' in which the driver was instructed to drive in a particular way to trigger the i-DREAMS technology warning. This element of the study was intended to expose the driver to warnings they might not have experienced during the main simulator trial. The drivers were then asked to express their opinion about the warnings – including their style and delivery, timing and purpose. In addition to this, there was a questionnaire assessing the driver's opinion about technology already fitted to the tram and all drivers were asked two questions about their driving style during the break between the baseline and intervention phase.

The design of the simulator study was repeated measures, with the tram drivers operating the simulator for two main drives: one as a baseline drive with no interventions from the i-DREAMS technology, and the other main drive as an intervention drive, with the i-DREAMS equipment switched on. The same route was used for both drives, including both suburban and urban sections of track, reflective of typical tram driving (see Annex 3 for more information).

The simulator sessions were conducted at a UK tram company depot during the participants' scheduled work time, ranging in start time from 08:45h to 20:00h. The tram drivers who

[4] Fatigue in i-DREAMS is a composite of sleepiness and length of driving measures. The tram study primarily focused on the sleepiness elements and the nature of the i-DREAMS fatigue warning.

participated in the study were relieved for part of their rostered duty and placed on standby (either before or after a break). Following the study, depending on the driver's roster, they either finished and went home, or had their break and returned for the second half of their shift.

The equipment used in the tram simulator consisted of:

- Tram simulator (including tram company pre-programmed events and pedestrian/passenger movement) which provided speed and speed limit data.
- Questionnaires (including entry and exit questionnaire, after baseline drive questions, manipulation drive questions, self-report scales)
- A wearable device to measure heart rate and derive sleepiness.
- i-DREAMS equipment (including a display screen to show alerts that was attached to the simulator screen and a Mobileye camera, which used an additional screen to the side of the simulator to detect pedestrians and cyclists)

The drivers completed a series of four drives (familiarisation, baseline, intervention, manipulation) and a series of questionnaires and discussions about their drives, the equipment, and their experiences. More detail about the four drives and the manipulated events is included in Annex 3.

At the end of the session, the drivers were de-briefed, and asked to complete an exit questionnaire including feedback of their experiences of the i-DREAMS system. The simulator session was audio recorded to aid with transcription and analysis of the after-baseline questionnaire and manipulation drive questions.

The drivers were also asked to report their Karolinska Sleepiness Scale[5] (KSS) (Åkerstedt & Gillberg, 1990) score before and after the baseline and intervention drive, as well as twice during each drive, at the start and end of the urban area. The level of sleepiness inferred from the wearables was also reported as KSS scores, but in three bands (KSS 1-5, 6-7, 8-9).

Participants

Participants of the tram simulator study were all employees of the tram company and based at the main depot in the UK. In total, 30 participants were recruited for the simulator study. All participants were certified to operate and drive the trams, with a mixture of tram drivers and driver trainers. Participants were recruited through posters and volunteered to take part during their rostered work time. Session slots and driver allocation were organised by the Operations Director and the control room. The eligibility criteria stipulated at least six months tram driving experience. Drivers who participated were provided with a £10 Amazon voucher at the end of the session.

Recorded Variables

The following variables were recorded during the simulator study:

- Questionnaire data
  - Participant demographics
  - Experience (tram driving, sleepiness, safety systems)
  - Use and opinion of safety systems
- Simulator data
  - Time
  - Distance travelled
  - Speed

[5] The KSS is a subjective measure of sleepiness, consisting of a 9-point scale ranging from 1, 'extremely alert' to 9, 'very sleepy, great effort to keep awake, fighting sleep'. Drivers were familiarised with this as part of the introduction to the study.

    - Speed limit
    - Signal status
- Subjective sleepiness (KSS scores)
- Heart rate data and inferred sleepiness level (wearable)
- Triggering of i-DREAMS system – speed and VRU detection (Mobileye and gateway)

6.1.2 Data analyses

Analysis of i-DREAMS alerts – speeding, VRU detection and fatigue

Data from the simulator was combined with data collected with the i-DREAMS technology to form a database that recorded when speeding occurred, a VRU was detected, or a fatigue warning threshold was met. Similar to the cars and trucks, these 'events' are classified as 'medium' (STZ level 2, dangerous driving) and 'high' (STZ level 3 avoidable accident). This allowed comparisons of events triggered in the baseline and interventions to be made.

Qualitative data from participants

Participants were asked a number of interview style questions both during and after the study. These sections of the study were audio recorded and transcriptions of these were produced. Thematic analyses were conducted using the software NVivo. Two researchers were involved in the analysis. Once data from the first three participants were analysed these researchers agreed on a set of themes to code the remaining participating drivers' data.

6.1.3 Results

Participant Demographics

Thirty participants were recruited for the study; two of these withdrew towards the beginning of their session due to feeling simulator sickness which leaves a total of 28. The average age of participants was 47 years. The oldest participant was aged 66 years and the youngest was aged 24 years at the time of the study. There was just one female participant, with the remaining 27 being male.

The average period that participants had held a licence was 10 years, with the longest period being 23 years and the shortest being one year. Twenty-three of the participating drivers worked full-time with the other five stating their hours were part-time.

Real time intervention and driving behaviour

For the tram simulator study, the i-DREAMS system could provide two real-time warnings, either for speeding, or if the system detected a VRU. These alerts were activated during the intervention drive of the study; however, the system also recorded the times the warnings would have alerted during the baseline drive. The following results detail the number of times the warnings were triggered for speeding and for VRU warnings. Fatigue data was collected and is discussed in the following sections, but real-time warnings were not available to the drivers during the simulator trials.

Due to technical issues, the simulator results are based on 26 sets of data. Of these 26 data sets, 14 participants' baseline drives, and 16 intervention drives finished early due to drivers failing to negotiate pre-programmed events[6]. Therefore, in relation to the number of i-DREAMS

[6] As the simulator used in the trial was designed for the purpose of training, if the system judged the tram to have come into contact with a pre-programmed parked vehicle, cyclist or pedestrian, the simulation immediately ended (as the drive would in real life). It should also be noted that it is more difficult for a driver to judge proximity to other vehicles and VRU in the simulated environment and some events required the driver to sound the horn or bell at a specific moment to successfully navigate.

warnings, some participants drove for a shorter distance and the number of alerts is likely to be lower than it would be if all the drivers had completed the full drives.

Speeding warnings

For the baseline drive, the total number of speeding alerts triggered was 70; 41 at STZ level 2 and 29 at STZ level 3. The number of triggered alerts ranged from 0 – 6 for STZ level 2 with an average of 1.6, and from 0 – 5 with an average of 1.1 for STZ level 3.

The total number of triggered speeding warnings for the intervention drive was 76; 51 at STZ level 2 and 25 at STZ level 3. The number of alerts ranged from 0 – 7 for level STZ level 2 with an average of 2.0, and from 0 – 3 with an average of 1.0 for STZ level 3. The results are displayed below in Figure 28 (baseline) and Figure 29 (intervention).

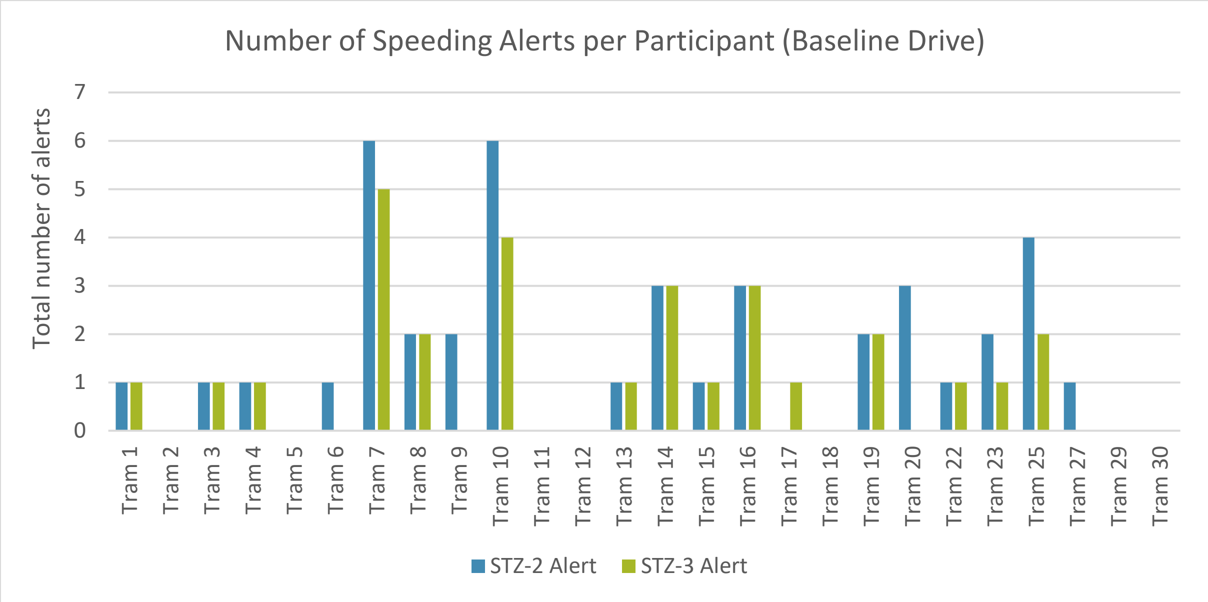


Figure 28: Total number of speeding alerts per participant during the baseline drive

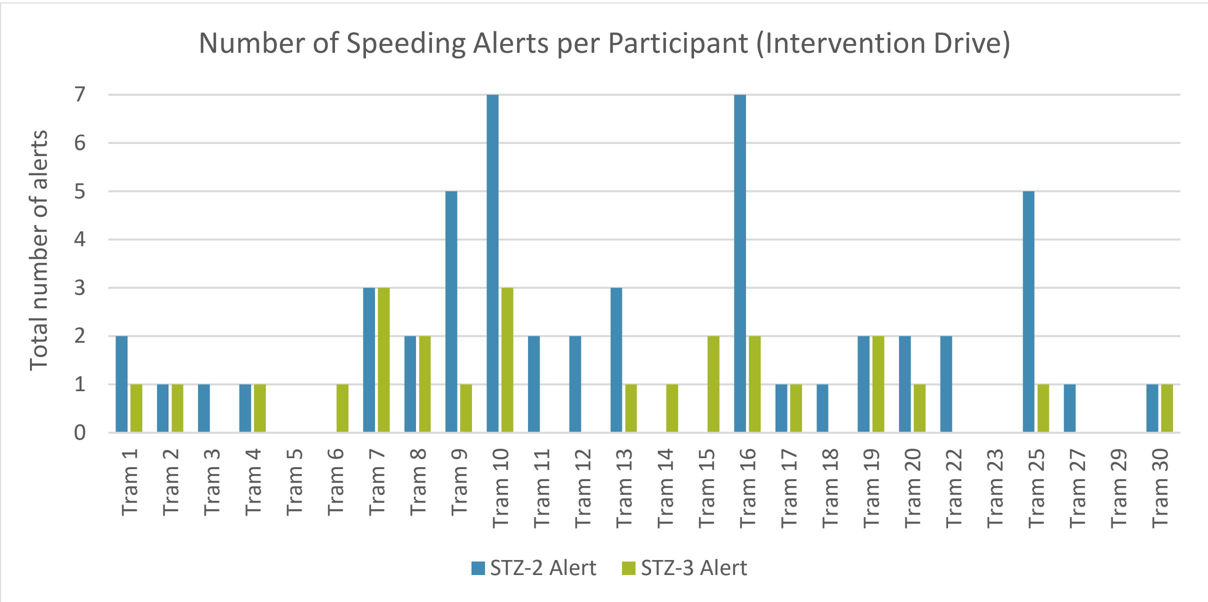


Figure 29: Total number of speeding alerts per participant during the intervention drive

The results show that the i-DREAMS speeding warnings did trigger during the simulator study, both during the baseline and the intervention drive. For both drives, the participants were driving the same route, but to avoid learning effects the events were different for each of the main two drives. Therefore, the participants always completed their baseline before the intervention drive and the program of events during the drive were alternated between participants (as described in Annex 3).

The participants of the simulator study were tram drivers and as such are highly trained to avoid going over the speed limit. Therefore, instances of speeding were unexpected. However, it should be recognised that simulator driving is different from real life driving, and it was reported by the participants during the study that the Traction Brake Controller[7] (TBC) was slightly different in the simulator and less responsive, which possibly contributed to the speeding events. A simulated environment also contains reduced risk, which will have likely impacted the participants' driving.

VRU warnings

The results show that the i-DREAMS VRU warnings did trigger during the simulator study – in total 57 times for the baseline drive (no alerts to the participant) and 48 times for the intervention drive. However, when exploring the results in more detail through observations made during the simulator session, it was clear that they did not appear to alert when simulated pedestrians were in front of the tram or crossing the tram's path. Instead, the warnings were more likely to trigger when pedestrians were standing at the edge of the platform, therefore resulting in false positives. As previously mentioned, for the baseline and intervention drives, the same route was followed, but with different events. However, the pedestrians were randomly pre-programmed into the simulator and therefore did not appear in the same place every drive. It was clear from the observations that the warnings did occur in similar locations on both drives (before tramstop 2, leaving tramstop 3 and before tramstop 4, and after tramstop 7). However, it is not clear whether the false alerts were due to either the equipment (Mobileye) having issues detecting simulated pedestrians or the design of the simulator study and environment, and therefore requires further development.

Overall, the results of the testing indicate that the system and the VRU warnings did not work in this tram simulation context, producing false positive alerts and not warning for pedestrians running across the tram's path. It is possible with further development that the system could improve in terms of reliability of warnings, for example, altering the algorithm of Mobileye and the level of detection to account for track versus tramstop.

Fatigue warnings

Due to the nature of the study (a short simulator study where the drivers had to remain fit to work), measurements of fatigue were taken via the wearable, but warnings were not triggered during the baseline or intervention drives. However, drivers were able to experience the fatigue warning during the manipulation drive, during which the i-DREAMS technology was programmed to provide a fatigue warning. Driver opinion on this is discussed below. KSS were asked for at points within both the baseline and intervention drives (Start of drive, start of urban section, end of urban section, end of drive). The majority (range 17-28) of participants were in the 'not sleepy' range of KSS 1-5 with only one driver reporting a KSS 6-7 (sleepy but no effort to stay awake). On 16 occasions the end of drive KSS score was not collected (this was often due to the drive having been terminated before the finish point).

Driver interpretation of i-DREAMS warnings (Speeding, VRU and fatigue)

Speeding

The speed warning was not triggered by all the participants during their drive so the research team asked the drivers to exceed the speed limit during the manipulation drive. The drivers were therefore generally aware of the appearance and sound of the warning, although four said that they thought that drivers would be distracted by looking at the display and five others

[7] The means of controlling the acceleration and braking of a tram

said the chime was too quiet. It is difficult to be sure how the drivers would respond to the warning during real-time driving since seven thought it was unnecessary because they do not generally speed and six others because drivers are well aware of all speed limits. They therefore could not see the usefulness of the speed alert in its current state. Eleven drivers suggested it would be more helpful as a warning prior to the occurrence of speeding whilst 23 believed it would be more effective as a constant in-cab reminder of the current speed limit.

VRU detection

All of the participants experienced the VRU event warning, albeit as a result of false positives, usually because of pedestrians at the side of the tram at tramstops or along routes. The research team explained the limitations of the VRU detection technology in the simulator to the participating drivers. As a result, their views were not based on practical experience of the alert but were largely hypothetical in nature. Notably, two drivers claimed not to have "*connected"* with the alert and five of the drivers said that they had not known what it signified when they first heard it, with one of these mistaking it for a speed warning. After further discussion and explanation from the research team, all 28 participants recognised how the warning would sound, look and work.

When asked how they would react to the VRU event warning, two of the participating drivers said they would slow down to a safe speed and two others claimed they would apply the hazard brake. It should be noted here that 12 of the 28 participants felt that looking at the screen would be a distraction from the driving task, taking attention away from the road in front of the driver; this might affect their reaction to the alert. For example, as one participant said: "*So I think this is useful but it shouldn't obviously detract from what the driver is meant to be doing … Instead of thinking, oh, I need to look at these screens, when your attention should be focused in front of you*."

Fatigue

Twenty one of the 28 participants recognised that the visual warning of a coffee cup[8] relates to needing to take a break or being tired, linking the symbol with needing some caffeine. One driver noted, "*… my car gives you that signal. It gives you the same image*." The remainder (seven) believed that it was not obvious or intuitive and did not recognise what it signifies, for example, "*… is it saying take some fluid or something? Is it?".* It was suggested by one driver that the cup alone might not be sufficient and that a red triangle or other symbol might be added to signify a warning.

With regard to their potential reaction to the fatigue warning, three participants recognised that they should report the alert to the tram control room and stop driving. There was acknowledgement from five of the participants of the potential difficulties caused by a driver becoming fatigued whilst driving the tram route. Two of these suggested that they believed that the company does not have the facility to take a driver off the tram and replace them with another person and also thought that the company would be unlikely to introduce a procedure for this occurrence[9]. It was also suggested that the fatigue warning would be triggered rather often, which the managers might dislike. Two drivers highlighted the fact that they cannot just stop the tram for a break, as they might when driving a car, with three participants noting that the company already has a relevant policy in place. These drivers referred to the fatigue policy as being the solution to the problem of fatigue: "*With us, with the fatigue policy It's kind of – there's no excuse anyway, if you're tired you come off. Whether people are scared or not, that's up to them, but there's a policy now*".

[8] The display screen showed a symbol of a coffee cup and an audio warning at the beginning of the manipulation drive.

[9] NB. A fatigue policy which allowed the tram driver to be taken off shift (either before or during) was in place at the time of the study

With regard to alerting the company to fatigue, two of the drivers said that there could be reluctance amongst drivers to call up and say they are fatigued in case managers think they are trying to avoid driving, but one participant also said the system would be helpful if it confirms fatigue when it is claimed. Conversely one participant noted that drivers might deny being tired and others might abuse the system. Another concern of one of the drivers was how the control room would be alerted by the system, and indeed whether that should happen because the fatigue warning could be used to discipline drivers *("I can't help feeling that something like that could be used as a stick")*.

Comparison with wearable data

The main aim of collecting both verbal KSS data and the wearable data was to compare the reported subjective scores with the i-DREAMS sleepiness score. The i-DREAMS sleepiness score used the Inter-beat Interval[10] (IBI) as measured by the wearable to infer KSS equivalent scores. Algorithms applied to the data produced output reported as KSS scores, but in three bands, KSS 1-5, 6-7, 8-9. The analysis aimed to investigate whether the output of the wearable wristband correlates with reported KSS scores and whether any correlation (e.g. stronger/weaker) was dependent on the environment being driven (urban/suburban).

In total, the study resulted in 20 data sets of both verbal KSS and IBI wearable data: 19 sets of data for the baseline drive and 18 sets of data for the intervention drive. The sampling rate of the wristband was approximately every 2-3 minutes, and therefore not the same as the verbal KSS data collection which was obtained at four specific time points in the drive. However, due to connectivity issues and simulator drives finishing early, the wearable data was not obtained every 2-3 minutes for every drive (numbers are shown in Table 52). For the baseline drive, the range of IBI reported scores was 7-18, with three drivers having less than 10 scores from the wearable. For the intervention drive, the number of IBI wearable scores ranged from 5-20, with five drivers having less than 10 scores from the wearable.

Table 52: The number of wearable scores obtained per drive

	Up to 5 wearable scores	6-10 wearable scores	11-15 wearable scores	16-20 wearable scores
Baseline drive (n=19)	0	3	12	4
Intervention drive (n=18)	1	4	7	6

According to the IBI wearable data, of the 20 data sets (19 baseline, 18 intervention), 20 drivers were KSS 6-7 at least once during either baseline or intervention (18 drivers during baseline and 18 during intervention) and 18 drivers were KSS 8-9 at least once during either the baseline (18 drivers) or intervention (13 drivers) drives. This does not correlate with the previously reported KSS scores with most drivers verbally scoring their KSS 5 or below and only one driver reporting a KSS of more than 5.

As the sampling rate of the wearable and the verbally reported KSS are different, and because of the inconsistent number of reports, it is difficult to draw direct comparisons between the verbal KSS and the wearable output. Figure 30 shows a comparison of verbal KSS reports and the IBI wearable data for the first and last KSS scores obtained (grouped into the same bands as the wearable output), and the first and last IBI wearable data output.

[10] Relating to heart rate

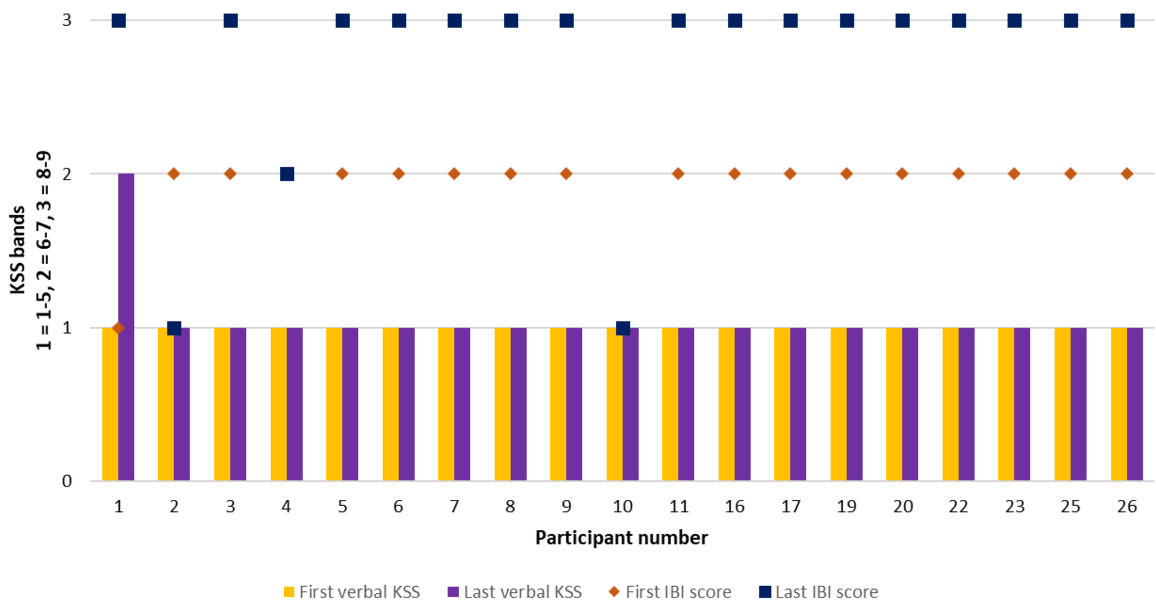


Figure 30: A comparison of verbal KSS scores and IBI wearable data

For most participants, the verbal KSS scores and IBI wearable scores for these time points do not match. The figure indicates that the difference between the verbal KSS and the IBI wearable scores was greater during the intervention drive, with a greater proportion of inferred KSS 8-9 from the wearable output. KSS 8 and KSS 9 are indicators of severe sleepiness (KSS 8 = sleepy, but some effort to keep awake, KSS 9 = very sleepy, great effort to keep awake, fighting sleep). Therefore, in well-rested individuals and without induced sleepiness (either through sleep restriction or conducting the study during the night), KSS scores of 8 and 9 are unlikely, albeit not impossible.

The results show that the IBI wearable data output and the verbally reported KSS scores did not correlate. Although the drivers received information relating to the KSS before the study, had the scale explained to them on the day, and had a chance to practice, it is possible that some drivers did not understand the scale or reported lower KSS scores due to the influence of the work environment and not wanting to appear sleepy at work. However, due to the time of day of the study and the fact drivers were participating during their duty time, minimal sleepiness was to be expected. There were connectivity issues with the wearables which influenced the number of data points obtained and due to technical issues and issues with the pre-programmed events during the simulator drive, there were a varying number of verbal KSS reports per participant, both of which limit the analysis. Different sampling rates and time points of KSS data collection also limited the comparison analysis.

As the analysis found that the IBI wearable output was too high compared to the verbally reported KSS scores, the data obtained as part of this study has been used to improve the IBI/KSS algorithms. The verbal KSS scores reported by the drivers were used to extend the database and train the sleepiness model, improving the classification of sleepiness in a wider population range and mitigating some of the difficulties inherent to inter-subject variability. The data was also separated and used as a train and test data set, before retraining, to ensure generalisation of the results. This has resulted in the updated model producing results more in line with the verbally reported KSS data. Therefore, the data from the tram simulator study, which took place before the on-road trials, was used as part of the i-DREAMS refinement process.

6.2 UK Tram Simulator Trial - Acceptability/Usability of Technology

Drivers were asked to complete questionnaires at the beginning and end of the tram simulator trial sessions. The first questionnaire assessed drivers' views about the importance and acceptance of safety systems already fitted on the trams. They were also asked to identify any warnings/systems that would be useful additions to the driving task. The second questionnaire, completed at the end of the simulator trial, included questions aimed at assessing the drivers' opinions about the i-DREAMS system. Results from these questionnaires will be discussed in the following sections. Interview style data gathered during the simulator trials was also used to inform the section examining the usefulness and acceptance of the i-DREAMS system.

6.2.1 Subjective importance/acceptance of safety systems in general

Relationship with safety systems

In order to investigate the relationship between the participating drivers and the safety systems already existing in the tram cab, they were asked to answer the following question: 'How important do you think the following safety systems are for safe tram driving?'. The possible responses were as follows:

1. **Not very important** – it never contributes to my safe driving.
2. **Not important** – it rarely contributes to my safe driving.
3. **Important** – it sometimes contributes to my safe driving.
4. **Very important** – it actively contributes to my safe driving.
5. **Not applicable**

The analysis showed that there was very little variation between the ratings from the participants associated with each safety system. As can be seen from Table 53, ratings were very much concentrated around 3-4 ('important' and 'very important'). It was therefore not possible to investigate any relationships between ratings and any characteristics of the participants (for example age, driving experience, enjoyment of driving etc.). It should be noted that the two participants who experienced simulator sickness are included in this analysis, but the one participant who rated all systems at 5 ('not applicable') has been discounted.

Table 53: Average ratings from participating drivers for safety systems (n=29)

	Drivers Safety Device (DSD)	Fatigue monitoring device	Track-based speeding intervention	Device preventing wrong-side door opening	Emergency Stop Button	Emergency PAN (pantograph) Down button
Average	3.7	3.4	3.7	3.8	3.6	3.5
Median	4	4	4	4	4	4
Mode	4	4	4	4	4	4

The participants were asked to provide up to three words to describe their relationship with each of the safety systems present in the tram. The participants were not asked to rank the terms. Table 54 shows the five most used terms in relation to each of the safety systems in the tram cab. It can be seen from the Table that 'useful' appears in all of the lists and is the most often chosen term for all systems except for the Fatigue monitoring device where it is equal first with 'distracting'. 'Necessary' also appears in five of the six lists, as does 'important'. It is notable that the examples 'reliable', 'useful' and 'distracting' were offered to the participants, but it can be argued that they only picked those terms with which they agreed.

Table 54: The most used terms to describe the driver's relationship with current safety systems

DSD		Fatigue monitoring device		Track based speeding intervention	
Useful	15	Distracting	9	Useful	8
Reliable	10	Useful	9	Necessary	5
Necessary	3	Annoying	4	Reliable	5
Essential	2	Necessary	3	Needed	4
Important	2	Needed	3	Important	3
Device preventing wrong-side door opening		**Emergency Stop Button**		**Emergency PAN Down Button**	
Useful	11	Useful	5	Useful	10
Important	7	Useful if needed	5	Needed	5
Needed	5	Important	4	Necessary	4
Helpful	3	Essential	3	Good	3
Essential	2	Necessary	3	Important	3

For most of the emergency systems the five most commonly used terms were positive suggesting that the systems both well accepted. The exception to this is the Fatigue monitoring device. The primary purpose of the system is to detect fatigue events by providing a warning when the system thinks eyes are closed (or not visible). Both distracting and annoying appear in the list for Fatigue monitoring device, suggesting this system is less well accepted. When asked in the questionnaire to comment on their choices, one of the drivers claimed that the system is 'distracting when activated in error'. This suggests that false positives might play a role in why the system is less accepted. This is supported by the interview style data collected during the simulator trial. Five drivers stated that they thought the Fatigue monitoring device is overly sensitive and is tripped too often thereby introducing too many alerts to the driver.

Additional warnings or safety systems

In order to capture any other opinions held by the participating drivers not covered by the existing questions, they were asked the following question: If warnings or safety systems could be added within the tram cab related to your driving, what would you find the most useful to be alerted about? In response, comments were received from 17 of the 28 participants, with some offering more than one idea.

One participant was generally in favour of safety systems, noting that they thought 'Anything to alert drivers/control of any issues that are safety related would be welcome'. Another participating driver noted that they would like an internal auto braking system to be added to the tram cab; it is noted here that such braking is important in rail more broadly, so in theory i-DREAMS would be required to integrate it to be suitable for rail. The remainder of the comments are presented according to those suggestions that are currently addressed by i-DREAMS, those that could be integrated into the system and those which are out of the scope of the functionality of the system.

Currently addressed by i-DREAMS

In order to integrate these safety issues, the i-DREAMS system might need refining, but it already measures them to some extent.

- Alarm for overspeed everywhere: the system currently monitors the tram's speed compared to the speed limit. Participants suggested, i-DREAMS warning could indicate current speed limits, but this would need further consideration and the warning sound should be louder than it is currently.
- Stress alert: the system currently provides an indicator for fatigue (based on heart rate); this could be extended to alerts for stress.

- Collision alert: the system currently provides warnings related to headway and VRU events. The technology would need calibrating for the tram context but has previously successfully been fitted on trams in Portugal.
- Proximity: two drivers commented on the potential for a proximity alert relating to the driver's cab area. The i-DREAMS system currently provides VRU event warnings which could be focused around the cab.
- Alarm for vehicles close behind you: the system currently provides warnings related to headway which could be modified.
- Pedestrians in close proximity of the tram when moving off: the system currently provides a VRU event warning which could potentially be modified for this purpose.

Potential for integration into the i-DREAMS system

The following issues have some possibility of being integrated into the i-DREAMS system by, for example, the addition of a monitoring tool or sensor.

- Swept path[11]: comments from two drivers suggested that they would find it useful to include an alert to detect obstructions within the swept path: it is suggested that an additional sensor might achieve this.
- Signal aspects: a device to give a warning about signals which are on stop as you approach them was suggested. In common with this comment, another participant noted that approach signal state would be useful. Lastly, one driver said that if a device could be applied to signals to show when they are about to time out, they think it could prevent some Signal Passed At Stop (SPAS) events.
- An alert to warn of upcoming curves and signals could be an effective addition; given current knowledge of the tram route, such an alert could be added to the system.

Not within the scope of the i-DREAMS system

The following are warnings which the participating drivers would like to be added to the cab, but which are out of the scope of the functionality of the i-DREAMS system:

- Louder alarm for DSD on Stadler Trams: this is a system external to i-DREAMS over which the developers would have no jurisdiction.
- Comfortable seats: the system is not concerned with the ergonomics of the train cab.
- Large decrease in speed zones or limits: the management of the tram network is subject to the authority of established organisations and bodies not allied to the i-DREAMS system.
- Overheating of driver: this relates to driver comfort and is therefore not related to safety.

Acceptance and usefulness of *i*-DREAMS interventions

This section presents the analysis of the interview style data gathered after the intervention drive and during and after the manipulation drive. It will focus on the main factors addressed by the i-DREAMS platform, speeding, VRU detection and fatigue.

Speeding

When questioned about the speed warning, seven of the participants doubted the usefulness of the system by commenting that the tram drivers do not generally speed because they are taught to remember the speed limits across the tram network. For example, "*I don't think there's a huge amount of speeding anyway on the system … because most people know the*

[11] The term used to describe the area beyond the rails needed for a tram to successfully navigate a curve.

speed limits". In addition, nine drivers noted that they already have a system which alerts them through sound to the 70 kph limit already. It was stressed by the drivers that this is operational only in the 70 kph zones. Given the assertion that not a lot of speeding occurs, it was feared by participants that the system might become just another noise in the tram cab and therefore distracting.

Six of the drivers noted that the i-DREAMS speed warning alert would be unnecessary because drivers are trained to be well aware of the speed limits across the whole route: they should know where the speed zones are as part of their normal routine driving. Five drivers believed that the speed warning would be counterproductive and is not needed because the trams have the track based speeding intervention system which fulfils the tasks that would be done by it. The track based speeding intervention system is a track-based speed detection and countermeasure system. At designated high risk locations along the track, and in the event a tram is travelling above the speed limit, the track based speeding intervention system will automatically apply the service brakes and bring the moving tram to a controlled stop. As stated by one driver: "*And you've got a [track based speeding intervention] in the dangerous – where it is dangerous, so personally, I think there's enough now*."

An alternative view on the track based speeding intervention from 11 drivers was that the i-DREAMS system might be helpful in avoiding the activation of the track based speeding intervention ("*I think it would be good in a [track based speeding intervention] area to give you that extra warning so you didn't trigger the [track based speeding intervention]*"). These drivers suggested that a generally preventative technology would be effective, letting drivers know whilst driving that they are approaching the top speed, or the limit is about to change, for example, "*… kind of a reminder to say you're under the limit or you're approaching the limit, you might want to do something about it, kind of thing. That could be helpful*". This was therefore thought to be of benefit in avoiding the activation of the track based speeding intervention (which can lead to disciplinary action) especially if set at a low threshold.

Six of the participants could see a reminder of the speed limit being useful, with one of these noting that it can be easy to speed by just a small margin without noticing. The constant speed reference which is displayed by the i-DREAMS system was commented on positively with 23 of the 28 drivers suggesting it would be effective as an aide memoire, but only if it was as accurate as possible. As stated by one participant, *"… when you've been driving the system all the time you know the speed limits. But there's nothing wrong with having it there, it wouldn't harm. It's like a sat nav in your car it tells you – in case you forget*."

With regard to the way in which the speed alert is delivered to the drivers, three participants said that the visual display could be useful, with two others commenting that visual warnings inside the cab are excessively distracting. With regard to the sound alert, five participants found this difficult to hear and suggested that it should be louder. Conversely, four drivers stated that if the alert was to be made louder, it could be a distraction and an overload of audible warnings may be annoying and become so commonplace they eventually come to be ignored: "*And it will probably just become another noise that you just – it's so difficult*". It should be noted that the current version of the i-DREAMS system presents a relatively quiet sound for the speed warning. This was not easy to hear within the simulator room due to the noise of the simulator itself, the numerous associated computers and noise from outside the room. This may in fact reflect the reality of the noise encountered in the real tram driving context. If a display is used for the alert, four drivers thought that it should be placed in the driver's line of sight, although carefully positioned so as not to be a distraction.

VRU detection

Eleven of the participants thought that the VRU warning system could be beneficial and may keep people alert and safe, whilst nine others did not see its usefulness, for example, "*It's* *useful information but it's information the driver should know anyway to be honest*". Only one participating driver was neutral about the warning system. Seven drivers mentioned that if the technology was overly sensitive it would be triggered too often and ineffectively ("*If it's going to be beeping a lot, no, it would be unhelpful. Yeah, that'd be distracting, I think, for us, yeah*"). They suggested that a system which is making a noise too often will be distracting and will ultimately become background noise and too many alerts will be confusing to the drivers and a distraction while they work out what it signifies.

There was disagreement about the usefulness of the system at tram stops, with six participants believing that the system could be useful when arriving or leaving the stop. For example, one participant said "*… there's obviously a lot of people in close proximity to the trams, coming into a tram stop sometimes … it's absolutely heaving, you're coming in at 5k because you know somebody could push from the back and before you know it there's a problem …*". However, three others noted that the system would not be needed at stops because drivers should be slowing on approach to stops anyway and because passengers always stand near the edge drivers already anticipate that eventuality. Indeed, drivers noted there is a lot of close proximity to the tram by VRUs so this might confound the system and make it less useful. It was also suggested by four of the 28 participants that the warning might be more effective on the segregated (out-of-town) sections of the tram route due to there being fewer pedestrians there.

Nine of the drivers who disagreed with the use of the VRU event warning said that there should be no need for this system: "*I think that if we're doing our jobs properly, we won't need a system to tell us when people are too close or not close enough*". Therefore, if drivers stay alert during the day the system would be unnecessary, providing useful information but that which the driver should know already.

Six participating drivers suggested that the VRU event warning would be most useful if it covered VRU movement along the side of the tram because drivers can see quite well already at the front[12] *("Not from the front because that you can see, but down the side"*). One of the drivers also thought to be of potential use if concentrating on tram stops or signals where people might be situated in their peripheral vision and therefore hard to see or seeming to appear out of nowhere. Another participant said that the VRU event warning might additionally be an effective aid when drivers are experiencing two potential events at one time, and it is hard to address both simultaneously.

The chime and symbol associated with the VRU event warning were said by four drivers to be not too annoying. It was stressed by one participant that the audible warning should not be similar to the Fatigue monitoring device, which they said can shock drivers when it is activated. One participant noted that any sound would need to be distinctive because the tram cab can be loud at times: "*so at least you can obviously recognise that that's what it means. Yeah, it's not a bad sound to be honest … because there are some really annoying sounds*". However, there was disagreement over the suitability of the current VRU event warning sound, with two drivers suggesting it was insufficiently loud (and the cab can be noisy) and four others finding it acceptable. A further four drivers thought the warning might be a distraction and would likely be set off too often so perhaps a visual warning might be better than an audible one.

Drivers also demonstrated a lack of trust in the technology in relation to the VRU event warning. One participant said that *"… it takes 208 metres to stop a tram at about 75/80k and the human eye would see it before anything else. If you had an electronic device, it would need to see so much that it would be giving an alarm persistently*". Another suggested that drivers have such

[12] The Mobileye system used for the road trials and tram simulator trial only had a forward facing camera however a version of Mobileye that has additional 'blind spot' cameras that cover the area along the sides of the tram could be used in a real-world application.

important experience of the driving task that they are able to see far in advance and selectively. In addition, one participant said their experience of body language would not be reproduced by technology, another said that driver experience is the best so this would be counterproductive. Another stated they believe technology would not be sufficiently sensitive to different kinds of pedestrians (people at a foot crossing compared to on the tramway). Two participants therefore suggested that the technology could be more helpful if adjusted appropriately for use in trams.

Fatigue

With regard to whether the fatigue warning would be appropriate in the tram, twelve of the 28 participants thought that it would be useful. Three of them suggested it could be used as a check for the drivers who tend not to realise they are becoming fatigued. As a result, they suggested it would be useful for advice, a visible warning prompting drivers that they are fatigued and/or should be taking their allotted break. This could act as a signal to consider their state of alertness. Two drivers expressed the opinion that people know better for themselves than technology how they feel when it comes to fatigue. Two of the participants were less positive about the fatigue warning, suggesting that it would be unlikely to impact on driver safety because it would be just another system within the cab and could in fact act as a distraction.

Seven of the drivers noted that the existing Fatigue monitoring device is sufficient for their needs, thereby negating the requirement for the i-DREAMS system. However, two drivers suggested that the new technology could be used as a forewarning before the Fatigue monitoring device comes into play, that is, "*I would say yes* [it would be useful] *if it can detect the onset of fatigue. Because at the moment the [Fatigue monitoring device] only detects the end (laughs) of fatigue*".

With regard to the chime which is associated with the fatigue warning, one participating driver thought an audible warning would be preferable to a visual one; conversely four further participants suggested that a visual warning would be more effective, due to the myriad other tones evident in the tram cab. For example,

"*There's a lot going on with the [Fatigue monitoring device]. If you look to your left for too long, a buzzer goes off. Then you've got all the bells and that going off. If it'd be something inaudible, that fatigue alert warning, that would be a bit better than having an audible warning*".

Five drivers expressed concern that there were too many false alerts from the Fatigue monitoring device. It was therefore suggested by two drivers that the i-DREAMS system should aim to be less sensitive.

A key issue for the drivers was a lack of trust in the technology related to the fatigue warning; five of them said that they believed it would not be reliable or accurate and would therefore be triggered too often. Reasons for lack of trust were misgivings about the method used to calculate fatigue[13] ("*Well, I mean, what's the accuracy of that? Is it going from your body … is it really reliable?*") and comparisons with the Fatigue monitoring device; drivers who find the Fatigue monitoring device untrustworthy thought the i-DREAMS system might be similar and they would only accept it if it is more reliable.

[13] The drivers were told that the i-DREAMS wearables used heart rate-based data to infer fatigue

6.3 UK Tram Focus Group Study – Post-trip App Evaluation

The simulator trial did not explore the post-trip feedback/App elements of the i-DREAMS platform. Instead, a focus group study was conducted to explore the views and observations of tram drivers about the post-trip feedback functionalities of the i-DREAMS system and how it might be employed in tram cabs.

The qualitative method of focus groups with tram drivers was used to carry out the investigation. The benefits of qualitative data (understanding attitudes, providing insights specific to the industry, detailed data which can explain a complex issue) were thought to apply in this study since it can help to understand the reasons behind the issues. A discussion of the use of qualitative research methods in the study of drivers can be found in Stiegemeier et al. (2022) who suggest that qualitative methods can provide a more holistic understanding and might add to the current body of literature through allowing the identification of constructs beyond the commonly known.

6.3.1 Method

Participants

Six focus groups were conducted online with drivers and driver trainers attending for refresher training at the depot. They were therefore contributing to the focus groups as part of their scheduled working day. These usually involve up to four tram drivers and a driver trainer; all those present at the course were invited to participate, so some of the groups included driver trainers (who also work as drivers for a certain period during a month). Since the focus groups were organised by the participants' employer it was stressed to all of them before beginning the discussion that they were under no obligation to participate and could terminate their participation at any time without giving a reason. This was clearly understood since five of the groups which were approached by the operator declined to participate in the study; the researchers estimate this to be around 50% of those who were invited.

A total of 20 participants were involved in the focus groups. As can be seen in Table 55, the average age of the participants was 43 years, and all were male. The time working in their current role as reported by the participants ranged from 0.9 to 21 years, with an average of 5.6 years. The time spent working for the present employer ranged from 0.9 to 21 years with an average of 6.7 years.

Table 55: Demographic data provided by participating drivers and driver trainers

Age & Gender of Participants	
Average age	43 years (SD 9.4: Range 29 – 59)
Male	100% (n = 20)
Participants' service as a train driver/instructor	
Average period spent working in current role	6.1 years (SD 5.6; Range 0.9 – 21.0)
Average period spent working for current employer	6.7 years (SD 5.5; Range 0.9 – 21.0)
Average period spent working in the rail industry	6.8 years (SD 5.8; Range 0.9 – 23.0)

Procedure

The research received full ethical approval from Loughborough University. Prior to the focus group, participants were sent an information sheet and completed an informed consent form which included details about the recording of the discussion. They also provided basic anonymous demographic data. Participants were encouraged to treat the session as an informal forum in which to discuss their views, and to express their honest opinions. Participants were encouraged to talk to each other and to the researchers during the session to explore the shared experience between drivers. Participants were informed that any information they provided would be kept confidential, and that no individuals or operating companies would be identified in any reports.

A focus group discussion guide of questions and prompts was developed iteratively by the research team to ensure each group followed the same format and to ensure the reliability of the data collection. The guide was influenced by the previous research carried out with tram drivers (from the same operator) involving a simulator and the prior knowledge and expertise of the researchers. The focus group was intended to introduce a new safety system to the participants and a PowerPoint presentation (together with a print-out of colour images from the App) was presented to them to describe and explain the concept. This was broken down into sections (see Table 56 for details) in which some of the presentation was given and followed by questions related to the content presented. Table 56 shows an overview of the topics, content, and example questions which featured in the focus group. At the end of the questions from the researchers, participants were invited to make additional comments or pose further questions.

Table 56: Demographic data provided by participating drivers and driver trainers

Topic	Example questions
Presentation explaining these functionalities of the App: Real time data collection; App home screen; Trip Info; Scores per safety area/overall; Examples of use	What is your first impression of what you have seen so far? What do you like? What do you dislike? As a driver would you be likely to use this App? Do you think it would improve safety?
Presentation explaining these functionalities of the App: Safety information and coping tips; Examples of use	What would be the effect of providing this type of information to drivers? Would there be benefits to you? Who would benefit and who would not? Do you think it would improve safety?
Presentation explaining the gamification aspects of the App: Leader board; Goals and badges; Examples of use	Do you think you and other drivers would make use of this aspect of the App? Do you think the operator would be interested in implementing this element? Would there be any barriers to its use? Do you think this would motivate drivers to change?
Sharing App information with the operator	What do you think would be the benefits of sharing information with the operator? What would be the disadvantages? Would there be any barriers to this application of the App? If so, what would they be? How do you think drivers would feel about this?

Data collection took place online (via MS Teams) in April and May 2022 during the usual working hours of the participants. The focus groups involved between three and five participants, lasted 50-60 minutes, and were held in a private room where the discussion could not be overheard by anyone outside. No managers were present. Two researchers facilitated the focus groups, one gave the presentation explaining the functionalities of the App and the other asked the questions. The latter was always the same person. this was an attempt to ensure the reliability of the data. Both were able to answer any questions from the participants.

6.3.2 Analysis

The focus groups found that drivers had mixed opinions about the potential introduction of the app, with both positive and negative views being discussed.

Basic app – trip maps, events and performance scores, improvement tips

In terms of the advantages of the app, the most frequently cited was that having access to data and videos could help to provide evidence to management that issues drivers frequently complain about (such as shift work, pedestrians acting recklessly etc.) are in fact a real problem. The participating drivers believed that the data from the app could provide evidence

©i-DREAMS, 2019

of issues on the tram route and for the service in general, for example, hazards in the town centre, effects of certain shift patterns, certain problematic areas, locations where drivers often speed and so on. As noted by one participant,

"… that could be seen as evidence … lots of these drivers have had a day off but gone from a late shift to a dead early shift, and then they've been fatigued on the early. So that's obviously a problem, something we should be looking at, from a management perspective, to address."

Similarly, areas of track that have repeat issues shown on the system could be flagged using the data and the videos. This could offer insight into the nature of such issues and could allow possible solutions to be devised which could benefit all drivers. The ability to share best practice, new policy changes and reminders about safe driving was also seen as advantageous by some in the hints and tips sections.

The ability to monitor their own driving performance and reflect on how they could change their behaviour to be safer or more efficient after the fact was also appreciated by many. In particular, two of the participants were in favour of the features related to sleepiness and fatigue monitoring. This was in terms of the support such features could offer to the drivers rather than their supervisors. For example,

"That's a useful tool to help you regulate yourself and give you an opportunity to say, "You know what? I am feeling tired. It's my job and duty to call the controller and say, 'I think it's time for me to be relieved' … I think that should be driver-only unless an incident occurs. In which case, retrospectively you could say, let's try and find out why this happened. But I think that's a good thing."

One of the participants noted that drivers are not routinely rewarded for their safe driving,

"*We don't get rewarded for collisions we avoid, the lives we save, smooth rides. We don't get anything like that",* and the feedback could confirm when exemplary driving happens.

Another positive aspect identified by the participants was the ability of the app to recognise weaknesses in an individual's behaviour which can then be considered and improved. The individual aspect of this positive was clear in many of the responses, for example*, "I like the self-improvement side of it. So yes … that side of it would be good."* The participants believed that the feedback could also be used to suggest tips on improving the driving of others. For example, *"I think it's good because I think, what that does is it identifies where they need to brush up a little bit sort of thing … it's come up with suggestions of how to do a certain thing better."*

The ability to review trips and events during particular drives was also seen as an advantage of the app; this might relate both to good drives and to those which a driver thinks could have been safer: *"… you'd be able to look back and say, I felt that I did a good drive there, what was it that I did?"*

The point was also made that the app data would permit drivers to compare scores with each other and, perhaps more importantly, to track their own performance across time (in terms of scores): *"And I like the, you can also compare yourself over time with the scores over time, so you can actually look back to see if you're getting better, so I think that's quite a good one as well."* In connection with this idea, it was suggested that the feedback could be used as part of regular training or performance review and to overcome complacency and over-confidence which might be present in some drivers.

A more inclusive approach was seen in the suggestions that drivers might share tips with each other, for example, "*when it is hot, how do you sleep better and during the day how much water should you drink without needing to go to the toilet."* In addition, the feedback might encourage and support drivers in sharing updates on policy and refreshing their knowledge: "*if the training department were able to put things on there, that'd be quite useful rather than having to ask".*

App data and operator 'management'

There was, however, scepticism towards the system described to the participants which was evident from the more negative views about it. Two issues often mentioned included the fear of being watched, and very closely related to this, the fear of how management would handle any data (such as poor performance, low leader board standing etc.) or video evidence of unsafe behaviour. There was much distrust towards management and a fear that the data would be used in a punitive manner rather than to reward and praise. Connected to this, there was a wealth of discussion in the focus groups about who would have access to the data and the leader board and even given assurances that the app was being introduced to motivate and improve driving there was consensus that management would be using it in a disciplinary manner.

> *"I think it's a good idea, but … it depends how management use this … They may say, oh we can use this for disciplinary, or we can use this to get you on certain areas, or they can use it as a company tool rather than a safety tool. So it depends where the company's going to use it."*

The participating drivers were concerned that some of the metrics they would be scored on in the app could be out of their control. For example, that harsh braking is sometimes unavoidable: "*… what pings on a recorded system doesn't show in real life what happened, what event made you harsh brake or what made you pull away."*

Gamification elements

Participants suggested that the leader board might change driver behaviour, but with the risk that it would not necessarily be for the better. If drivers became too score focused they might change their driving behaviour making it difficult to gauge whether the driver is demonstrating optimal vehicle control. There was doubt that certain drivers would be prompted by the system to change their driving behaviour: *"… a lot of the older drivers – not older drivers – more experienced, set in their ways, know the system inside out, it won't affect them in any way whatsoever."*

The participating tram drivers also doubted the leader board would perform its designated function, that is, to motivate engagement with the app and encourage safer driving behaviour. They suggested that the leader board would instead demotivate those who frequently found themselves towards the bottom.

Post-feedback app – perceived usefulness

The belief that the perceived features of the App would be of little use or are already covered by other systems or training was quite frequently discussed. There were five specific key issues related to this:

- Previous awareness of fatigue and alcohol best practice: *"We're all very aware of that side of things. Fatigue wise, lifestyle wise, it is something that is repeated fairly regularly within the business. I'm not sure how much use having that would be, or how much people would look at it."*
- Participants were happy with their own current driving performance: *"I understand benefits to it, but I don't know. It might be useful for trainees and stuff, but from the point of view of an experienced driver, I don't really think you would need it to be honest."*
- App features already exist: *"I think those tools, you've already touched on the fact that our trips are already recorded by CCTV. We already have devices that monitor our speed … we've got speed monitors on the trams as well that beep for high speed and you get an alert going on. So we've got a lot of these things anyway, so my interest in this as a driver is what system can they bring us that we don't have?"*

- Less necessary in a low demand environment*: "… last night I was driving for two hours empty, picking up one or two people. So for me … the system may not be relevant as much as a late driver and it probably would be more useful, an early driver or a middle driver, because they go, they see a lot more in the system. There's a lot more … things can go wrong during the day than at night."*
- Track based speeding intervention offers speed warnings in the current trams: *"But the speed management I don't think that's going to be really any good to us, because like I say, we've already got certain things already here, which looks at the speed anyway, like [track based speeding intervention]".*

Recommendations for implementation

The focus groups also involved asking participating drivers for recommendations that could be taken on board when implementing the app. Most of these focused around ensuring drivers were consulted when designing the app to ensure it meets their needs and is designed and introduced with the user in mind. For example, one driver said "*… it depends how it's brought in and I think the drivers need to be involved rather than, here comes the [track based speeding intervention] system, here comes the [Fatigue monitoring device]."*

There were discussions of issues suggesting that the data and output of the app should be reliable and valid, and relevant to the driving task. Some of the participants had worries that the system would misjudge their driving as hazardous, for example, "*All of a sudden you could slide and then you'd have to act quickly to bring it back. But I get the feeling that this would then class that as a harsh brake, but that's just us bringing it back into control."* This could ensure that the measurements are equitable and would not penalise drivers on certain shifts, for example, some shifts have a greater number of stops to make, traffic, hazards and so on, whereas during the later shifts there tends to be less traffic and fewer passengers boarding.

Individual concerns were uppermost, with participants wanting access to any data from the app being limited to the tram drivers to whom it relates, and any tips and advice offered to individuals being tailored to their own particular needs.

The suggestion was made that there would need to be regular updating of the app to avoid users quickly losing interest. In reference to the hints and tips one driver said: "*I can't see the longevity of it. I think it might be a case of you read one of two of them and … I think once you've read them, they're probably read once or twice more and then probably not much after tha*t."

Recommendations for the leader board were also that it should be opt in, so those who do not believe they would be motivated by it could choose whether to be included. It should also allow users to choose who they are compared with (grouped with close friends, for example) so that only healthy rivalries would ensue. Finally, the implementation of the leader board would require assurances at the outset regarding the role of company processes and procedures for those exhibiting lower scores or frequently placing lower down on the board, for example, *"Then management tells us, well this is what we're going to be using it for and this is what's going to happen if you get a bad score on say three days in a row."* It was also suggested by participants that the leader board should include very limited information.

6.4 Learnings for Future *i*-DREAMS Development

This section of the deliverable summarises the learnings from the simulator study and issues which should be considered for the future. It is concluded that the i-DREAMS system cannot be directly translated into trams in its current format and so some adaptation is required. The points below are therefore some suggestions on methods and issues which can guide this adaptation.

General learnings

- Further developments need to be made to the i-DREAMS system based on the results of this study and then tested both within a tram simulator context and then as a series of on-rail studies. The restrictions and limitations of the present tram simulator study should be addressed in future studies.
- Any integration into trams would require wide-ranging modification of the system.
- Any new system would likely need to be able to work in conjunction with other existing safety systems, in particular the Fatigue monitoring device. Currently installed in all of the trams driven by the participants, the Fatigue monitoring device vibrates when it detects sleepiness, alerting the driver and helping them to become more alert. It is also used in regulating the speed of the trams.
- The position of the i-DREAMS display would be crucial and would need to be in the line of sight of the driver. It should also be made bigger in size.
- Tram drivers continuously make use of their extensive route knowledge and experience of driving the route for safety; this could be incorporated into the system.
- Drivers are trained to expect events and to follow body language of VRU and drivers.
- Weather conditions may have an effect on the ease of driving in trams.
- The participating tram drivers experience higher workload in built-up areas, which often leads to them choosing to drive more defensively.
- In less built-up areas participants claimed to relax more although they still need to be vigilant particularly in terms of events further away from the tram. They recognise that their relaxation might make them more hazardous in terms of driving.
- There was a general lack of trust in the technology in terms of all of the warnings; this was sometimes due to prior experience of other technology proving unreliable. In addition, there was a belief that technology is not as effective as human experience (especially in professional driving).
- The issue of how to present warnings needs further consideration; some participants noted already having many audible alarms, so their preference is for a visual method. Audible alarms which trigger too often become less trusted and eventually are ignored.
- The drivers want to be involved early on in the process of adopting new technology. They do not want it imposed on them and want to have an opportunity to express concerns before implementation.

Fatigue warning

- At the time of the study, the algorithms used for the wearables did not correlate with the KSS scores and were inflated, i.e., they were higher (sleepier) than the verbally reported KSS scores and did not fit with the context of the study design (e.g., time of day, no sleep restriction). The study data has since been used in the refinement process to further improve the algorithms, extend the database and train the sleepiness model. This should be further tested, in both alert and sleep restricted individuals.
- The fatigue warning symbol would benefit from some embellishment to make its meaning more obvious.
- The fatigue warning element of the system would require careful integration into the culture of the company; fatigue management would be a part of such a process.
- Many of the drivers would like the fatigue warning to act as a reminder to them that they should check their sleepiness level or to prove fatigue to their managers. The fatigue warning would not be welcome amongst the drivers if used as a reprimand.

Speed warning

- The speed warnings did trigger in the simulator study. However, the simulator environment is different from real life driving, with reduced risk, reduced workload and reduced sensitivity in terms of braking and acceleration.
- The speed warning was generally thought by the participating drivers to be unnecessary since they claimed that tram drivers rarely speed. Such a warning could, however, be usefully employed as a forewarning of speeding or as a reminder on the occasions when the speed is just above the limit.
- Tram drivers are trained (and expected) to learn all the speed limits so some drivers suggested a warning should not be needed but it was recognised that a warning system could be effective for new drivers and when Temporary Speed Restrictions (TSRs) are in operation.

VRU event warning

- The VRU warnings did not work in the simulator environment, resulting in false positive warnings or the system missing pedestrians passing in front of the tram.
- The i-DREAMS system's ability to detect pedestrians in a tram environment needs to be further developed and explored in future studies. One consideration may be the height of the tramstops compared to the height of the track.
- Tram drivers habitually use their peripheral vision when driving and some suggested that the VRU warning would be most useful for monitoring the sides of the tram.
- VRUs around the tram tend to be unpredictable and the participating drivers did not envisage that there could be a technology-based warning system which would be of help in perceiving them in a useful way.

Post-feedback app

- The data used for monitoring needs to be tailored for the tram context and to be accurate.
- Information provided by the app should be updated regularly to promote engagement.
- Further research into whether gamification elements are appropriate for the tram mode is needed.
- Drivers feel that they are experts on the driving task and take safety very seriously. They are sceptical of the introduction of technology due to past experience and want to be involved and listened to within the process of introduction of new technology.

7 Discussion and Conclusions

Section 7 discusses the results and gives some key conclusions that can be drawn from the data. As the results varied between transport modes, conclusions are presented separately for private and professional drivers.

7.1 Private Drivers (Cars)

Outcome Evaluation

When data is combined for Belgium, Germany and the UK, who experienced the full system, there was a statistically significant decrease in events from Phase 1 to Phase 4. This was for both medium and high severities, for ‘total’, ‘vehicle control’, ‘speeding’ and ‘road sharing’ events. This suggests that the i-DREAMS system had a positive impact on the measured safety outcomes and succeeded in keeping drivers in the first level of the STZ for more of their journey. In Greece, where only the app was used, there was also a statistically significant decrease in events overall, for all severities of ‘total’ events, medium severity ‘vehicle control’ events and high severity ‘speeding’ events. Therefore, although results are still very good, it could be argued that the app alone was not as effective as the full i-DREAMS system. When individual phase changes are considered, in both the combined and Greek data the most significant results were seen from Phase 3 to Phase 4. This suggests that the addition of the gamification elements had a significant impact on safety outcomes, and further supports the conclusion that the full system provides the most effective results. There were also positive results regarding the change objectives, with a statistically significant improvement in drivers self-reported knowledge of the benefits and importance of safe driving, and of what is needed to drive safely.

Looking at the different safety promoting goals, the interventions appeared to have the greatest and most consistent impact on ‘road sharing’ events. However, these data were only available for Belgium and the UK, so it would be useful to collect further data for other countries to support this finding. ‘Vehicle control’ events were least significantly impacted, which could be due to the fact that there are no real-time warnings in relation to this SPG. When ‘speeding’ events are analysed for the combined countries data, the most significant change was from Phase 1 to Phase 2, when the real-time warnings were introduced. Due to issues with data collection, it is difficult to form robust conclusions regarding ‘driver fitness’ events (fatigue and distraction). Valid distraction data were only really available for Greek drivers, and we did see a statistically significant decrease in events. Further work is needed to refine the i-DREAMS system to better capture data relating to driver fitness for the car mode, which could then be used to understand how these specific behaviours can be targeted.

Differences were found when each country was analysed individually, which were statistically significant, though there is not a clear reason why this would be so (apart from the Greek data, where drivers did not experience the full system). In the UK, the greatest reduction in events was found, and for all SPGs/severities there was a statistically significant decrease in events from Phase 1 to Phase 4 (apart from ‘driver fitness’ events, which were considered separately). The UK drivers had the largest number of events, therefore one suggestion could be that the technology has the greatest impact on more ‘risky’ drivers. In the questionnaire data, UK drivers did describe themselves as more confident and risk-taking relative to other countries. Belgian drivers had fewer significant results, and had the only statistically significant increase (‘speeding’ events overall from Phase 1 to Phase 4). Furthermore, Belgian drivers generally had fewer events than other countries, but in the questionnaires reported that they engage in risk-taking behaviours far more frequently, and also had more recent speeding offences. One explanation could be that because drivers in Belgium believed themselves to engage in risk-taking behaviours quite frequently, if the technology then showed otherwise (i.e., less events /

higher scores than they expected), they did not feel such a need to change their behaviour. Belgian drivers had the most significant changes from Phase 3 to Phase 4, when gamification features were introduced. As this was only for the last part of the trial, this could be that they did not show as much reduction overall. Drivers in Germany and Greece had a much greater number of 'speeding' events compared to Belgium and the UK. German and Greek drivers were also typically younger and had less driving experience relative to the other countries, so there could be a correlation. Despite the higher number of events, both countries had statistically significant decreases, again suggesting the technology has greater impact on more 'risky' drivers.

Finally, differences were also found between drivers within countries. In each country, between two thirds to three quarters of drivers showed improved outcomes (i.e., a reduction in events), but the remainder had worse outcomes (an increase in events). It's not clear from the data why some individuals responded positively to the technology and others did not. There were little demographic differences between the two types of drivers, and where there were differences, they generally were not consistent across countries. There is some data to suggest the drivers who worsened were more confident relative to the drivers who improved, so it is possible they had less desire to change their behaviours, though this cannot be concluded for sure. Further work is needed to understand why the system has such varied effects on different drivers.

Process Evaluation

For all countries, drivers engaged more with the app in Phase 4 of the trial compared with Phase 3, after the introduction of the gamification features. Although the 'trips' and 'scores' menu were the functions most used by drivers (functions that were available in both phases), the data suggests that the gamification functions were more engaging and held attention more consistently. One suggestion is that the gamification functions prompted more regular visits, as for example a drivers position on the leader board and their progress towards goals would change daily. However, trips and scores, while clearly of interest, may be something drivers only review every few days, or if they felt a recent trip was particularly eventful. As the gamification elements also had a significant positive impact on the outcome evaluation, it is clear they add value, and it is recommended that future development of the i-DREAMS system or similar interventions make sure to focus on the use of gamification to promote sustained behaviour change.

The generic information in the app (hints, tips etc.) was less appealing to users. They found more interest in personalised feedback such as their trip information, goals, and position on the leader board. The data also suggests that push notifications appeared to be an effective method for increasing engagement, though this was mostly seen in Belgium and the UK.

Results were generally more consistent between countries regarding the app (compared to the outcome evaluation results), however there was particularly low engagement from the German drivers. It is not fully clear why this is, but they were also generally less accepting of the technology overall, and had more technical difficulties during the trial, which may partly explain the results.

With respect to user acceptance, the clear majority of participants felt the i-DREAMS system was easy to use and easy to understand. In Belgium and the UK, most drivers also felt the system improved their performance and helped them maintain safe driving behaviour, though scores were lower for the German drivers. However, only around half of drivers said they would continue to use the system, which is unexpected given the positive feedback. The most likely explanation for this is that the data also showed fewer positive responses in regard to trust, in fact there was a statistically significant decrease in drivers trust of ADAS after using the i-DREAMS system. In particular, drivers felt that the alerts did not always accurately reflect the situation (i.e., they received too many false alerts). In order to increase user acceptance of the system, further development should be done to reduce false alerts and ensure warnings are accurate and timely.

Finally, the data also suggests a link between app usage and performance outcome; nearly all the drivers who used the app heavily (on average one every day or more), showed improved outcomes. It would be interesting to investigate this further to determine whether there is a causal effect between these results, or if the high app usage was simply a factor of an overall higher engagement with the study and a greater willingness to change behaviour.

7.2 Professional Drivers (Trucks and Buses)

For private drivers, it is first noted that sample size was relatively small for each of the Truck and Bus groups, therefore results are indicative. Further testing should be done to form more robust conclusions.

Generally, the i-DREAMS system showed less positive impact with professional drivers compared to private drivers. Specifically, a lower proportion of the professional drivers showed improved outcomes, and little significant change was seen in terms of safety outcomes. Where there were significant results, these were most often increases in events, i.e., worse outcome. Again, it is not obvious why this result is observed. Possible reasons could be lower driver engagement, more delays and technical difficulties during the trial, less clear communication between the project and the drivers (typically communication was via the company, not directly), or the underlying safety culture and climate at the companies. However, these are just theories, there is no data to clearly explain why professional drivers were less impacted, and it is recommended to explore this further, in order to identify how the technology could be improved to be more beneficial for professional drivers.

The only statistically significant improved outcome was for truck drivers, which was for ‘total’ high severity events specifically between Phase 1 and Phase 2. Therefore, it can tentatively be concluded that the system had a positive impact on the most severe events.

Limited data were available for bus drivers, who did not use the app, only the in-vehicle system. It was found that bus drivers had a higher number of events relative to car drivers, whereas truck drivers had a lot fewer events. One possible explanation is that buses more often drive in cluttered environments (i.e., more interactions with other vehicles and vulnerable road users, more stopping and starting), or it could be due to driver behaviour and road environment differences between countries. There is no data to support conclusions on why these are differences were found, therefore it is again recommended to research further with respect to professional drivers.

Process evaluation results were only available for Truck drivers, but showed similar results to private drivers, with more app engagement in Phase 4 compared to Phase 3, after the introduction of gamification features. This further supports the value of gamification features, though further work is needed to understand why they had less impact on outcome performance compared to private drivers (decreases were seen from Phase 3 to Phase 4 in nearly every event category, however they were not statistically significant). Furthermore, the Truck data does not show the same correlation between high app use and improved outcome, so it would be interesting to explore this further.

7.3 Rail Transferability

The conclusions of the tram studies are given here. Readers can also refer back to section 6.3 for learnings that can be applied to future development of the i-DREAMS system.

Value for fatigue, VRU and speed warning in trams

Due to the mixed traffic, multi-user environment that tram drivers operate in, a system to help improve safety and mitigate risk has the potential to be useful. The tram simulator study suggests that the i-DREAMS system and associated warnings offer several benefits for tram driving operations. Firstly, as instances of speeding are rare, the speed alert would be more helpful as a warning before the occurrence of speeding, alerting the drivers they are approaching the limit, or more effective as a constant in-cab reminder of the current speed limit. The concept of a VRU warning could be beneficial to tram drivers operating in mixed traffic environments encountering VRUs regularly. However, it was clear that the VRU warning needs to be developed to take into account specific aspects of tram driving and there is a concern from drivers about it being triggered too often. The warning possibly has the most value in terms of approaching and leaving tramstops, detecting pedestrians on the segregated sections of track, or detection along the side of the tram/in the driver’s peripheral vision. The fatigue warning could potentially be beneficial as a warning before the Fatigue monitoring device alerts, as a prompt to drivers to consider their alertness or take a break. While the time on task fatigue element may not be as useful due to the management of driving time through shifts, it could support drivers in reporting instances of fatigue based on physiological data, if accurate. Finally, visual warnings may be useful for the drivers as the tram cab can be loud and audio warnings can be missed or difficult to distinguish between. However, this needs to not be distracting and could be difficult to distinguish if multiple alerts are being triggered.

Value of the post-trip feedback app

Tram drivers suggested that the app would be most useful in identifying issues that were common to drivers and as a self-evaluation tool. They were more sceptical about the gamification elements, in particular the leader board, and expressed views that competition could have a negative impact on safety and is therefore not desired. There were also mixed views on sharing the data with ‘management’. This would be acceptable to identify issues and to be used as a way to improve safety generally, but the fear was that data on individual drivers or incidents would be used in a disciplinary way.

7.4 Ranking of interventions

The intention was to use the results to inform the ranking of interventions and provide an assessment of which intervention schemes are most effective. However, given the varied results between countries and transport modes, it is difficult to conclude a definitive ranking of the different interventions.

The results indicate that the full system (real-time warnings plus app feedback plus gamification features in the app) provides the most significant positive impact on driver outcome.

For private drivers, the analysis of combined data showed that most significant positive change was seen in Phase 4 of the trial, i.e., the gamification features, however it cannot be said that those alone were the most effective, as they were tested in combination with the other interventions. However, the data does suggest that app feedback on its own is less effective than when the app also includes gamification features.

For truck drivers, we can tentatively conclude that the real-time interventions (introduced in Phase 2) had the most impact, however more data is needed to support this.

8 References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Brijs, K., Brijs, T., Ross, V., Donders, E., Yves Vanrompay, Y., Geert Wets, G., . . . Gaspar, C. (2020). Toolbox of recommended interventions to assist drivers in maintaining a safety tolerance zone. Deliverable 3.3 of the EC H2020 project i-DREAMS.

Fuller, R (2000) The task-capability interface model of the driving process, Recherche - Transports - Sécurité,Volume 66, 47-57. https://doi.org/10.1016/S0761-8980(00)90006-2.

Fuller, R. (2005). Towards a general theory of driver behaviour. Accident Analysis and Prevention,37, 461-472

Fuller, R. (2011) Driver Control Theory: From Task Difficulty Homeostasis to Risk Allostasis. In Porter, B. E. (Ed) Handbook of Traffic Psychology, 13-26, Academic Press: London.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. American Journal of Public Health, 89(9), 1322–1327. https://doi.org/10.2105/ajph.89.9.1322

Hancox, G. et al. (2020). Organisation of the driving simulator and on-road trial experiments in i-DREAMS Safe Tolerance Zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions. Deliverable 5.1 of the EC H2020 project i-DREAMS

Hancox, G. et al. (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

Katrakazas, C., Michelaraki, E., Yannis, G., Kaiser, S., Senitschnig, N., Ross, V., . . . Taveira, R. (2020a). 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.

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

Lourenço, A. Carreiras, C., Velez D., et al. (2020). A flexible driver-machine interface for real-time warning interventions, Deliverable 4.4 of the EC H2020 project i-DREAMS

Lourenço, A. Carreiras, C., Velez D., Gaspar. C, Brijs, T. (2021). A set of flexible modules for sensor data collection, integration and real-time processing, Deliverable 4.1 of the EC H2020 project i-DREAMS.

Lourenço, A. Carreiras, C., et al., (2023). Toolkit for vehicle operator safety. Deliverable 8.1 of the EC H2020 project i-DREAMS.

Pilkington-Cheney, F., et al., (2020). Experimental protocol. Deliverable 3.4 of the EC H2020 project i-DREAMS.

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.

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

Papazikou, E., Talbot, R., Brown, L., 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.

Stiegemeier, D., Bringeland, S., Kraus, J. & Baumann, M. (2022) "Do I really need it?" An explorative study of acceptance and usage of in-vehicle technology. Transportation Research Part F: Psychology and Behaviour, 84, 65-82. DOI: 10.1016/j.trf.2021.11.011

Talbot, R., Pilkington-Cheney, F., Hancox, G., Filtness, A., Brijs, K., Brijs, T., Ross, V., Katrakazas, C., Al Haddad, C., Taveria, R., Kaiser, S., Mateus, T. (2020) Framework for operational design of experimental work in i-DREAMS. Deliverable 3.1 of the EC H2020 project i-DREAMS.

project i-DREAMS

Vanrompay, Y., Donders, E.,Fortsakis, P.,Brijs, T., Brijs K.& Wets, G. (2020).A smartphone app (Android) for personalized driving behavioural feedback. Deliverable 4.5 of the EC H2020 project i-DREAMS

Annex 1: Additional Results

Cars – Data Sample

Question: How often do you as a car driver... (BE: n=50, DE: n=29, EL: n=80, UK: n=54)

Table 57: Questionnaire data – self-reported risk-taking behaviours, per country (cars)

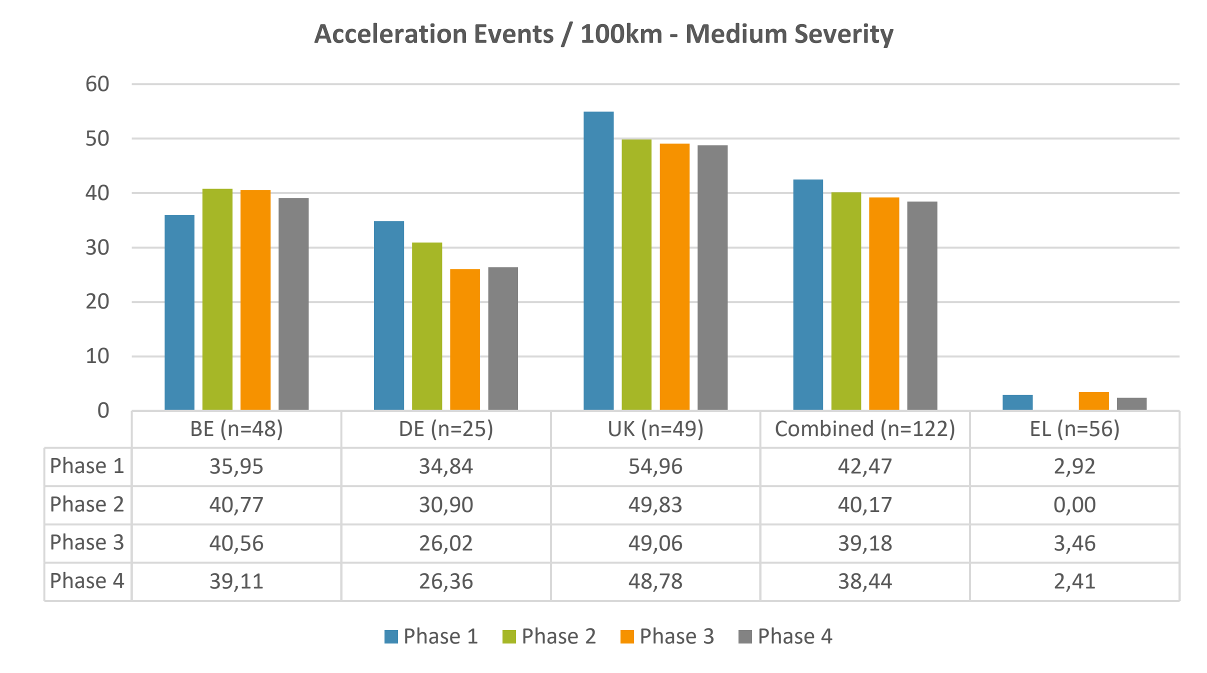
Question / Country Responding		Almost always	Usually	About half the time	Seldom	Never
…drive faster than the speed limit?	BE	26%	6%	62%	0%	6%
	DE	7%	10%	28%	48%	7%
	EL	6%	28%	43%	24%	0%
	UK	4%	17%	35%	44%	0%
…drive when you were so sleepy that you had trouble keeping your eyes open?	BE	10%	8%	52%	0%	30%
	DE	0%	0%	3%	31%	66%
	EL	0%	3%	6%	38%	54%
	UK	2%	0%	4%	48%	46%
…use a hand-held mobile phone while driving?	BE	4%	0%	36%	0%	60%
	DE	0%	3%	10%	59%	28%
	EL	6%	20%	36%	25%	13%
	UK	0%	2%	9%	48%	41%
…drive to close to a vulnerable road user?	BE	2%	0%	42%	0%	56%
	DE	0%	0%	3%	38%	59%
	EL	1%	8%	28%	46%	18%
	UK	0%	0%	9%	70%	20%
…drive without respecting a safe distance to the vehicle in front?	BE	8%	0%	56%	0%	36%
	DE	0%	3%	17%	45%	34%
	EL	1%	6%	25%	51%	16%
	UK	0%	2%	19%	59%	20%
…illegally overtake another vehicle?	BE	4%	0%	52%	0%	44%
	DE	0%	0%	3%	24%	72%
	EL	0%	5%	18%	44%	34%
	UK	0%	0%	4%	37%	59%

Cars - Outcome - Events & Scores Results

Table 58: Acceleration, deceleration and steering events / 100km and scores per country and per phase (cars)

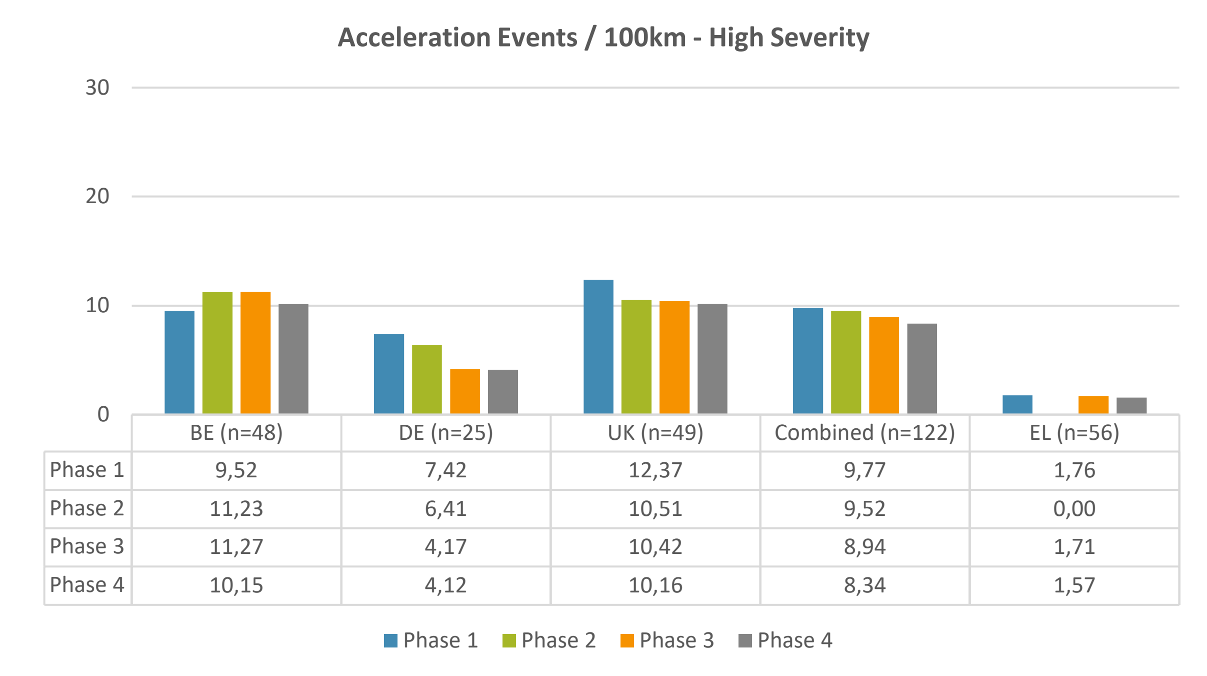
Phase	Belgium (n=48)		Germany (n=25)		UK (n=49)		Combined (n=122)	Greece (n=56)	
	Acc Events / 100km	Acc scores	Acc Events / 100km	Acc scores	Acc Events / 100km	Acc scores	Acc Events / 100km	Acc Events / 100km	Acc scores
Phase 1	45.47	48.58	42.26	57.11	67.34	42.27	52.24	4.69	91.24
Phase 2	52.00	47.42	37.31	53.05	60.34	42.90	49.69	NA	NA
Phase 3	51.83	45.67	30.19	62.49	59.47	44.12	48.12	5.17	90.94
Phase 4	49.26	46.45	30.48	61.80	58.94	43.42	46.78	3.98	92.31
	Dec Events / 100km	**Dec scores**	**Dec Events / 100km**	**Dec scores**	**Dec Events / 100km**	**Dec scores**	**Dec Events / 100km**	**Dec Events / 100km**	**Dec scores**
Phase 1	5.45	88.77	4.25	92.67	7.21	89.35	5.70	8.87	82.79
Phase 2	6.32	87.21	4.01	92.59	6.94	89.59	5.89	NA	NA

Phase 3	8.04	85.99	4.16	92.39	7.00	89.03	6.31	8.84	82.55
Phase 4	7.12	87.17	4.99	91.12	6.76	89.81	5.86	7.93	83.34
	Steer Events / 100km	**Steer scores**	**Steer Events / 100km**	**Steer scores**	**Steer Events / 100km**	**Steer scores**	**Steer Events / 100km**	**Steer Events / 100km**	**Steer scores**
Phase 1	50.60	47.37	50.30	45.34	62.15	38.98	51.04	NA	NA
Phase 2	49.64	47.57	52.76	40.59	64.39	37.61	52.09	NA	NA
Phase 3	50.07	47.73	55.10	41.70	64.21	37.22	52.21	NA	NA
Phase 4	46.36	50.68	61.78	40.26	64.90	35.33	49.85	NA	NA



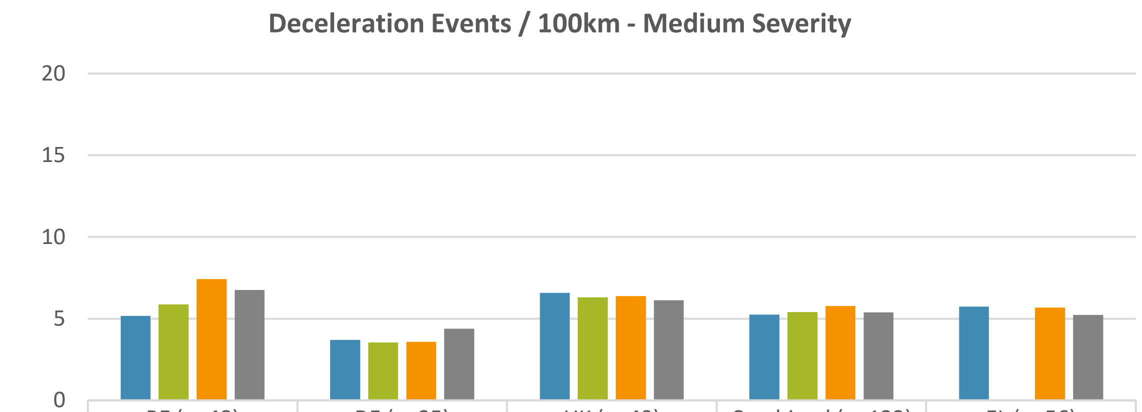
	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	35,95	34,84	54,96	42,47	2,92
Phase 2	40,77	30,90	49,83	40,17	0,00
Phase 3	40,56	26,02	49,06	39,18	3,46
Phase 4	39,11	26,36	48,78	38,44	2,41

Figure 31: Medium acceleration events / 100km per country and per phase (cars)



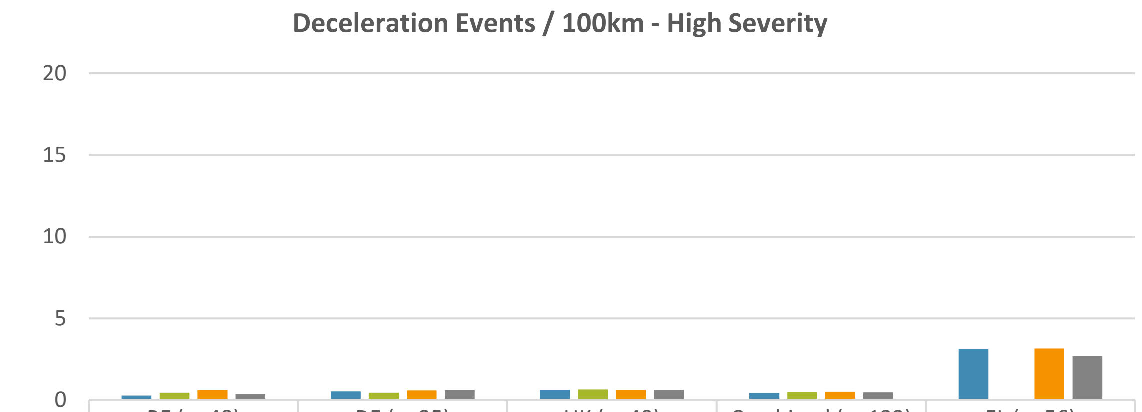
	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	9,52	7,42	12,37	9,77	1,76
Phase 2	11,23	6,41	10,51	9,52	0,00
Phase 3	11,27	4,17	10,42	8,94	1,71
Phase 4	10,15	4,12	10,16	8,34	1,57

Figure 32: High acceleration events / 100km per country and per phase (cars)



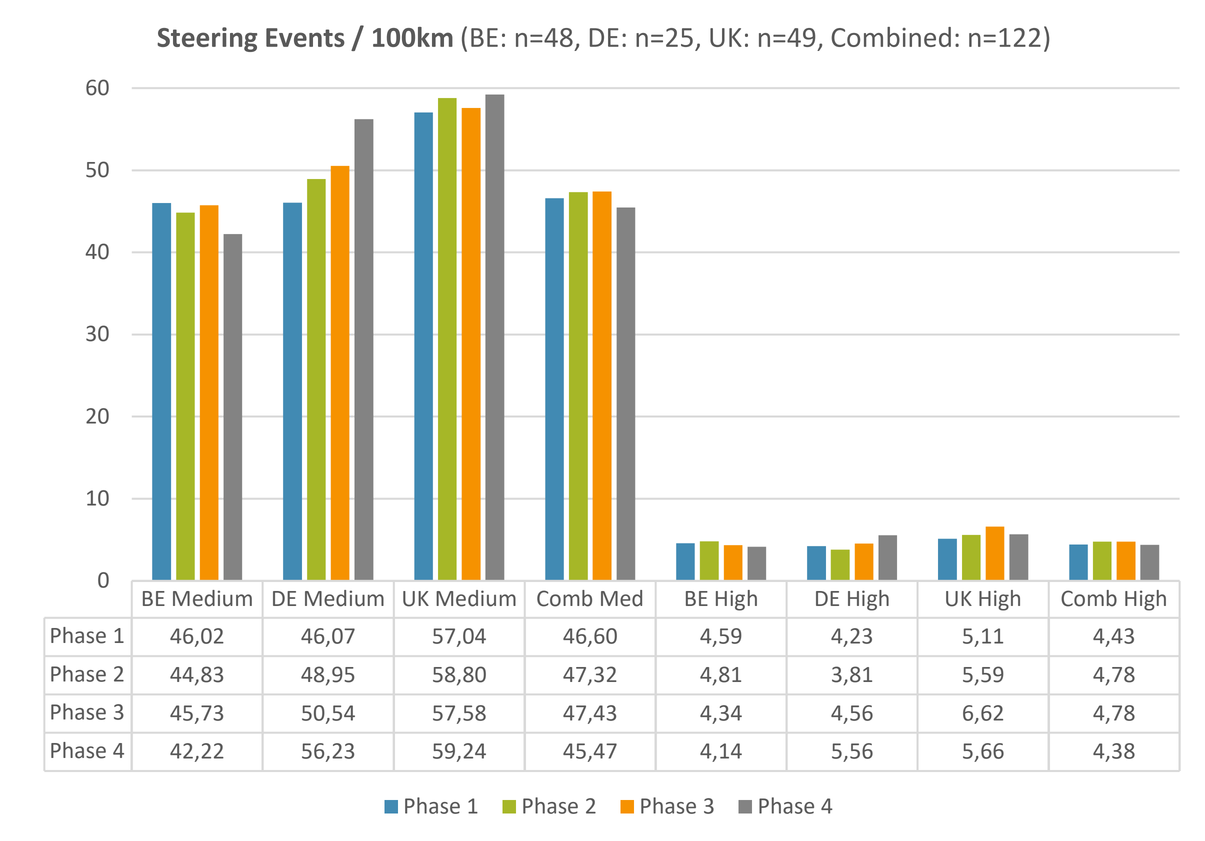
	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	5,18	3,71	6,58	5,26	5,73
Phase 2	5,87	3,55	6,30	5,40	0,00
Phase 3	7,43	3,58	6,38	5,79	5,68
Phase 4	6,75	4,40	6,14	5,39	5,24

Figure 33: Medium deceleration events / 100km per country and per phase (cars)



	BE (n=48)	DE (n=25)	UK (n=49)	Combined (n=122)	EL (n=56)
Phase 1	0,28	0,54	0,63	0,44	3,13
Phase 2	0,45	0,46	0,64	0,49	0,00
Phase 3	0,61	0,58	0,62	0,52	3,16
Phase 4	0,37	0,60	0,62	0,47	2,69

Figure 34: High deceleration events / 100km per country and per phase (cars)



	BE Medium	DE Medium	UK Medium	Comb Med	BE High	DE High	UK High	Comb High
Phase 1	46,02	46,07	57,04	46,60	4,59	4,23	5,11	4,43
Phase 2	44,83	48,95	58,80	47,32	4,81	3,81	5,59	4,78
Phase 3	45,73	50,54	57,58	47,43	4,34	4,56	6,62	4,78
Phase 4	42,22	56,23	59,24	45,47	4,14	5,56	5,66	4,38

Figure 35: Medium and High steering events / 100km per country and per phase (cars)

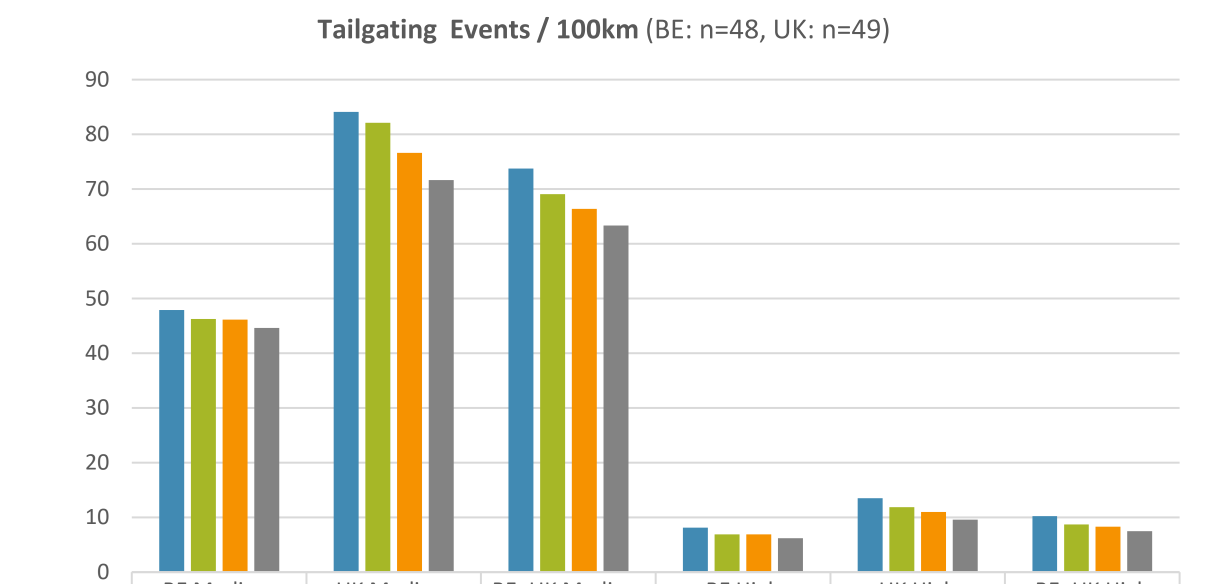
Table 59: Tailgating, lane departure, illegal overtaking, forward collision avoidance and vulnerable road user collision avoidance events / 100km and scores per country and per phase (cars)

Phase	Belgium (n=48)		UK (n=49)		BE+UK (n=97)
	Tailgating Events / 100km	**Tailgating scores**	**Tailgating Events / 100km**	**Tailgating scores**	**Tailgating Events / 100km**
Phase 1	56.05	79.94	97.62	79.92	73.76
Phase 2	53.17	82.22	94.01	83.42	69.06
Phase 3	53.05	81.75	87.61	83.96	66.40
Phase 4	50.85	81.86	81.28	85.87	63.37
	Lane Departure Events / 100km	**Lane Departure scores**	**Lane Departure Events / 100km**	**Lane Departure scores**	**Lane Departure Events / 100km**
Phase 1	5.39	89.93	7.78	85.95	6.41
Phase 2	5.32	90.07	6.29	88.71	5.55
Phase 3	5.36	90.38	6.45	88.82	5.70
Phase 4	4.63	91.74	4.94	90.94	4.86
	Illegal Overtaking Events / 100km	**Illegal Overtaking scores**	**Illegal Overtaking Events / 100km**	**Illegal Overtaking scores**	**Illegal Overtaking Events / 100km**
Phase 1	0.015	99.56	0.000	100.00	0.008
Phase 2	0.029	99.34	0.014	99.97	0.019
Phase 3	0.043	99.45	0.012	99.95	0.023
Phase 4	0.039	99.62	0.000	100.00	0.019
	FCA Events / 100km	**FCA scores**	**FCA Events / 100km**	**FCA scores**	**FCA Events / 100km**

Phase 1	3.53	87.20	13.72	71.29	7.63
Phase 2	3.59	88.74	12.92	72.71	6.96
Phase 3	3.16	89.25	11.59	75.72	6.45
Phase 4	3.58	89.53	9.65	77.96	5.94
	VRU CA Events / 100km	**VRU CA scores**	**VRU CA Events / 100km**	**VRU CA scores**	**VRU CA Events / 100km**
Phase 1	0.41	97.44	0.59	97.13	0.47
Phase 2	0.19	98.31	0.55	97.31	0.37
Phase 3	0.27	98.51	0.39	97.83	0.32
Phase 4	0.27	98.08	0.36	97.95	0.37

Table 60: Medium and high overtaking events / 100km per country and per phase (cars)

Overtaking Events / 100km	BE Medium	UK Medium	BE+UK Medium	BE High	UK High	BE+UK High
Phase 1	0.0153	0.0000	0.0076	0.0001	0.0000	0.0001
Phase 2	0.0290	0.0140	0.0187	0.0000	0.0000	0.0000
Phase 3	0.0409	0.0116	0.0225	0.0020	0.0000	0.0010
Phase 4	0.0389	0.0000	0.0192	0.0000	0.0001	0.0001



	BE Medium	UK Medium	BE+UK Medium	BE High	UK High	BE+UK High
Phase 1	47,93	84,09	73,76	8,12	13,53	10,22
Phase 2	46,28	82,12	69,06	6,88	11,88	8,72
Phase 3	46,13	76,61	66,40	6,92	11,00	8,32
Phase 4	44,65	71,66	63,37	6,19	9,62	7,50

Figure 36: Medium and high tailgating events / 100km per country and per phase (cars)

Cars - Outcome - Questionnaire Analysis

Table 61: Questionnaire data – outcome evaluation, results per question, per country (cars)

Evaluation Measure		BE (n=45)		DE (n=29)		UK (n=54)	
		Before	After	Before	After	Before	After
Perceived Knowledge	I know the benefits of safe driving	4.24	4.47	4.45	4.79	4.30	4.80
	I know what is needed to drive safely	4.11	4.33	4.21	4.52	4.13	4.69
Self-Efficacy	I have the skills to drive safely	4.13	4.18	4.55	4.76	4.07	4.46
	I feel competent enough to drive safely	4.15	4.20	4.31	4.48	4.20	4.54
	I control whether I drive safely or not	3.69	3.80	4.21	4.38	4.46	4.61
	For me, safe driving is easy to do	3.93	4.04	4.14	4.45	3.85	4.11
Attitude	Safe driving is important to avoid crashes	4.58	4.71	4.03	4.00	4.46	4.85
	Safe driving makes me feel comfortable	4.17	4.40	3.83	3.76	4.09	4.56
Personal Norm	For me personally, safe driving is important	4.38	4.49	4.52	4.66	4.33	4.61
	Safe driving should be a personal obligation	4.24	4.56	4.34	4.52	3.94	4.35
Subjective Norm	My friends think safe driving is important	3.55	3.91	3.83	3.93	3.80	4.13
	My colleagues find it important to drive safely	3.75	3.91	3.62	3.97	3.74	4.11

Table 62: Questionnaire data – outcome evaluation, stats results, per country (cars)

Evaluation Measure	BE (n=45)		DE (n=29)		EL (n=44)		UK (n=54)	
	Test	**sig**	**Test**	**sig**	**Test**	**sig**	**Test**	**sig**
Perceived Knowledge	Sign	0.007	Sign	<0.001	Wilcoxon	<0.001	Wilcoxon	<0.001
Self-Efficacy	Sign	0.115	Sign	0.134	Sign	<0.001	Wilcoxon	<0.001
Attitude	Sign	0.135	Sign	0.678	Sign	<0.001	Sign	<0.001
Personal Norm	Sign	0.018	Sign	0.424	Sign	<0.001	Sign	0.037
Subjective Norm	Sign	0.023	Sign	0.049	Sign	<0.001	Sign	<0.001

Cars – Process – App Usage

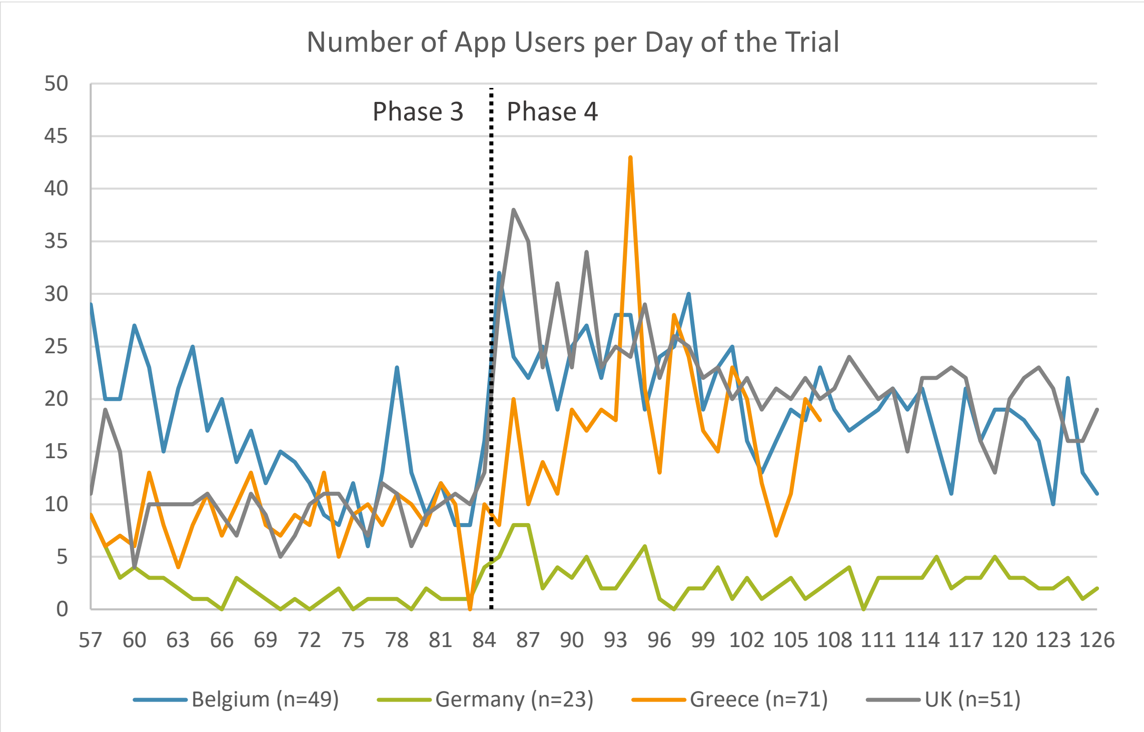


Figure 37: Number of app users per country, per day of trial (cars)

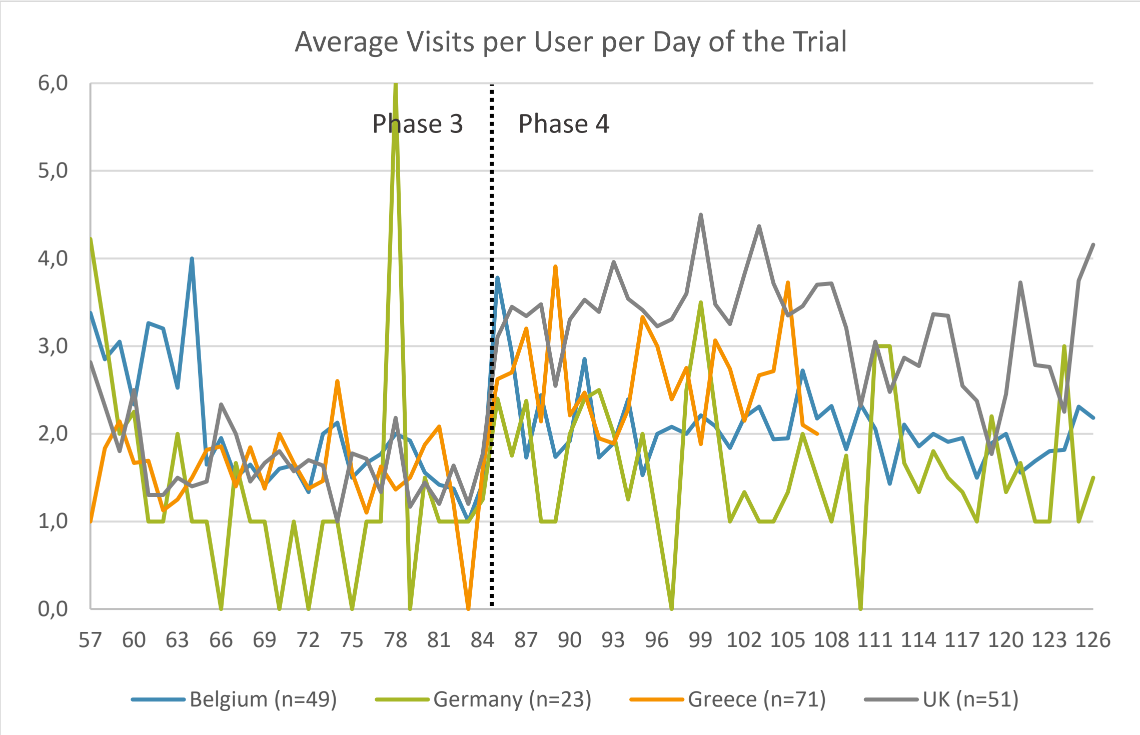


Figure 38: Average app visits per user per country, per day of trial (cars)

Cars – Process – Use of Technology and User Acceptance

Table 63: System fidelity questionnaire data (Belgium cars)

Belgium Cars (n=34)					
System Fidelity (UX1 – UX7)	**Totally Disagree**	**Disagree**	**No Opinion**	**Agree**	**Totally Agree**
Q1: Identifying myself in the vehicle is easy	3%	9%	3%	15%	71%
Q2: The display in the car works as it should	3%	24%	15%	38%	21%
Q3: The warnings are clear, I understand the meaning of all the symbols and sounds	3%	3%	6%	53%	35%
Q4: The alerts are correct. They correctly reflect the situation around my vehicle	9%	50%	18%	15%	9%
Q5: The alerts are always given in a timely manner allowing me to adjust my actions in a timely manner	3%	15%	24%	47%	12%
Q6: The warnings are sometimes distracting	29%	26%	18%	24%	3%
Q7: The alerts make me more aware of my actions while driving	6%	3%	6%	59%	26%

Table 64: System fidelity questionnaire data (UK cars)

UK Cars (n=54)					
System Fidelity (UX1 – UX7)	**Totally Disagree**	**Disagree**	**No Opinion**	**Agree**	**Totally Agree**
Q1: Identifying myself in the vehicle is easy	2%	4%	2%	11%	81%
Q2: The display in the car works as it should	4%	31%	2%	37%	26%
Q3: The warnings are clear, I understand the meaning of all the symbols and sounds	2%	7%	4%	41%	46%
Q4: The alerts are correct. They correctly reflect the situation around my vehicle	2%	43%	9%	33%	13%

©i-DREAMS, 2019

Q5: The alerts are always given in a timely manner allowing me to adjust my actions in a timely manner	2%	11%	22%	43%	22%
Q6: The warnings are sometimes distracting	4%	24%	13%	41%	19%
Q7: The alerts make me more aware of my actions while driving	2%	6%	6%	54%	33%

Table 65: User acceptance questionnaire data (Belgium cars)

Belgium Cars (n=46)					
User Acceptance (EX1)	**Totally Disagree**	**Disagree**	**Neutral**	**Agree**	**Totally Agree**
a. Using the system increases my driving performance	0%	2%	11%	63%	24%
b. If I use the system, I will reach my destination safely	0%	11%	30%	52%	7%
c. I think the i-DREAMS system is easy to understand	0%	2%	9%	52%	37%
d. I think the i-DREAMS system is annoying	15%	57%	22%	7%	0%
e. Using the system is a good idea	0%	0%	15%	70%	15%
f. The system makes driving more interesting	0%	15%	48%	30%	7%
g. I would be proud to show the system to people who are close to me	0%	7%	35%	46%	13%
h. In general, people who I like would encourage me to use the system	4%	7%	50%	33%	7%
i. While using the system I can maintain safe driving behaviour	0%	15%	22%	50%	13%
j. I have the knowledge necessary to use the system	0%	0%	2%	70%	28%
k. I am afraid that I do not understand the system	39%	37%	15%	9%	0%
l. I am confident that the system does not affect my driving in a negative way	0%	4%	9%	57%	30%
m. Using the system information requires increased attention	2%	43%	17%	35%	2%
n. The system distracts me from driving	24%	50%	20%	7%	0%
o. I think using the i-DREAMS system makes me a safer driver	0%	4%	15%	72%	9%
p. I think using the i-DREAMS system makes me more aware of my surroundings (other vehicles, lane position, etc.)	2%	0%	13%	63%	22%
q. I think I can depend on the i-DREAMS system	4%	15%	13%	61%	7%
r. I will feel more comfortable doing other things (e.g., adjusting the radio) with the i-DREAMS system	13%	46%	28%	9%	4%
s. If I had a choice, I would continue to use the i-DREAMS system	0%	11%	41%	33%	15%
t. I would recommend the i-DREAMS system to other drivers	0%	4%	33%	48%	15%

Table 66: User acceptance questionnaire data (Germany cars)

Germany Cars (n=29)					
User Acceptance (EX1)	**Totally Disagree**	**Disagree**	**Neutral**	**Agree**	**Totally Agree**
a. Using the system increases my driving performance	17%	28%	24%	31%	0%
b. If I use the system, I will reach my destination safely	10%	14%	34%	34%	7%
c. I think the i-DREAMS system is easy to understand	3%	7%	10%	55%	24%
d. I think the i-DREAMS system is annoying	14%	31%	31%	14%	10%
e. Using the system is a good idea	3%	7%	28%	52%	10%
f. The system makes driving more interesting	21%	24%	21%	17%	17%
g. I would be proud to show the system to people who are close to me	10%	10%	38%	28%	14%
h. In general, people who I like would encourage me to use the system	10%	17%	41%	28%	3%
i. While using the system I can maintain safe driving behaviour	3%	21%	28%	31%	17%
j. I have the knowledge necessary to use the system	3%	0%	7%	52%	38%
k. I am afraid that I do not understand the system	52%	38%	0%	10%	0%
l. I am confident that the system does not affect my driving in a negative way	3%	17%	14%	41%	24%
m. Using the system information requires increased attention	0%	45%	14%	38%	3%
n. The system distracts me from driving	14%	52%	14%	17%	3%
o. I think using the i-DREAMS system makes me a safer driver	14%	21%	24%	38%	3%
p. I think using the i-DREAMS system makes me more aware of my surroundings (other vehicles, lane position, etc.)	24%	24%	21%	24%	7%
q. I think I can depend on the i-DREAMS system	34%	34%	14%	17%	0%
r. I will feel more comfortable doing other things (e.g., adjusting the radio) with the i-DREAMS system	45%	31%	17%	7%	0%
s. If I had a choice, I would continue to use the i-DREAMS system	10%	31%	17%	34%	7%
t. I would recommend the i-DREAMS system to other drivers	10%	24%	34%	21%	10%

©i-DREAMS, 2019

Table 67: User acceptance questionnaire data (UK cars)

UK Cars (n=54)					
User Acceptance (EX1)	**Totally Disagree**	**Disagree**	**Neutral**	**Agree**	**Totally Agree**
a. Using the system increases my driving performance	4%	7%	26%	50%	13%
b. If I use the system, I will reach my destination safely	0%	6%	44%	39%	11%
c. I think the i-DREAMS system is easy to understand	0%	0%	7%	54%	39%
d. I think the i-DREAMS system is annoying	4%	28%	26%	28%	15%
e. Using the system is a good idea	2%	0%	19%	61%	19%
f. The system makes driving more interesting	4%	17%	28%	37%	15%
g. I would be proud to show the system to people who are close to me	0%	9%	24%	43%	24%
h. In general, people who I like would encourage me to use the system	2%	7%	43%	35%	13%
i. While using the system I can maintain safe driving behaviour	0%	9%	20%	50%	20%
j. I have the knowledge necessary to use the system	0%	0%	6%	31%	63%
k. I am afraid that I do not understand the system	22%	56%	17%	4%	2%
l. I am confident that the system does not affect my driving in a negative way	4%	11%	11%	43%	31%
m. Using the system information requires increased attention	6%	22%	24%	35%	13%
n. The system distracts me from driving	11%	33%	30%	19%	7%
o. I think using the i-DREAMS system makes me a safer driver	4%	7%	24%	44%	20%
p. I think using the i-DREAMS system makes me more aware of my surroundings (other vehicles, lane position, etc.)	6%	7%	19%	39%	30%
q. I think I can depend on the i-DREAMS system	11%	19%	30%	33%	7%
r. I will feel more comfortable doing other things (e.g., adjusting the radio) with the i-DREAMS system	7%	28%	41%	19%	6%
s. If I had a choice, I would continue to use the i-DREAMS system	11%	19%	24%	35%	11%
t. I would recommend the i-DREAMS system to other drivers	7%	15%	20%	37%	20%

Table 68: Questionnaire data – ADAS acceptance, stats results, per country (cars)

Evaluation Measure	**BE (n=45)**		**DE (n=29)**		**UK (n=54)**	
	Test	**sig**	**Test**	**sig**	**Test**	**sig**
b. Using ADAS increases my driving performance	Wilcox	.468	Wilcox	<.001	Wilcox	.303
c. My interaction with ADAS is clear and understandable	Sign	.108	Sign	.019	Wilcox	<.001
e. Using ADAS is a good idea	Sign	.004	Sign	.013	Sign	.100
f. I can maintain safe driving behaviour while using ADAS	Sign	<.001	Sign	.267	Wilcox	.166
g. I will feel more comfortable doing other things (e.g., adjusting the radio) with ADAS	Wilcox	<.001	Wilcox	.002	Sign	0.59

h. Using ADAS information requires increased attention	Sign	.265	Wilcox	.858	Wilcox	.986
j. I trust the information I receive from ADAS	Sign	.002	Wilcox	<.001	Wilcox	.042
k. ADAS distract me while driving	Wilcox	<.001	Wilcox	.335	Sign	.377

Cars – Differences Between Drivers

Table 69: Driver ADAS, accident and offence history, and confidence, per driver change type (Belgium)

Question / Response Option		**BE (All)**	**BE (A)**	**BE (B)**
Number of participants (drivers)		53	29	17
Participant gender	Male	33 (62%)	18 (62%)	11 (65%)
	Female	18 (34%)	10 (34%)	6 (35%)
	Unknown	2 (4%)	1 (3%)	0 (0%)
Participants mean age (years)[14]		46.7	48.0	49.3
Standard deviation of age (years)		18.2	17.8	18.9
Years driving experience (range, average)		2 - 55, 27.1	3 - 55, 28.1	2 - 52, 29.9
Which ADAS are present in your car? (Percentage replied equipped)	Automatic emergency braking	25%	28%	24%
	Blind spot warning	10%	14%	6%
	Drowsiness alert	14%	14%	12%
	Forward collision warning	27%	28%	35%
	High speed alert	27%	34%	18%
	Lane keeping assistance	24%	24%	24%
	Night vision & pedestrian detection	2%	0%	6%
In the last three years, have you been involved in an accident with your car, which was self-inflicted?	No	90%	83%	100%
	Yes, once	10%	17%	0%
	Yes, twice	0%	0%	0%
Within the last three years, have you been fined for a traffic offence while driving your car? (Excluding parking offences)	No	39%	41%	29%
	Yes - not specified	12%	21%	0%
	Yes - speeding	45%	31%	71%
	Yes - running a red light	0%	0%	0%
	Yes - multiple offences (speeding + running a red light, speeding + phone offence)	4%	7%	0%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	0%	0%	0%
	Discreet, average	61%	59%	59%
	Sportive, ambitioned	39%	41%	41%
	Risk-taking, offensive	0%	0%	0%
How confident you are concerning your own driving skills?	Insecure	0%	0%	0%
	Neutral	27%	31%	18%
	Confident	53%	45%	71%
	Very confident	20%	24%	12%

[14] Note that 5 drivers are not classified as Type A or Type B as they were excluded from outcome analysis. The average age of these participants was 28.8, hence the average for Type A and Type B drivers can both be higher than the average for all drivers.

Table 70: Driver ADAS, accident and offence history, and confidence, per driver change type (Germany)

Question / Response Option		DE (All)	DE (A)	DE (B)
Number of participants (drivers)		29	16	9
Participant gender	Male	19 (66%)	10 (63%)	7 (78%)
	Female	10 (34%)	6 (38%)	2 (22%)
Participants mean age (years)[15]		32.2	32.0	34.4
Standard deviation of age (years)		9.6	10.3	10.3
Years driving experience (range, average)		1 - 35, 11.4	2 - 35, 12.4	1 - 25, 11.1
Which ADAS are present in your car? (Percentage replied equipped)	Automatic emergency braking	31%	38%	22%
	Blind spot warning	21%	25%	22%
	Drowsiness alert	21%	19%	22%
	Forward collision warning	34%	38%	22%
	High speed alert	28%	38%	22%
	Lane keeping assistance	28%	31%	22%
	Night vision & pedestrian detection	7%	0%	22%
In the last three years, have you been involved in an accident with your car, which was self-inflicted?	No	86%	88%	78%
	Yes, once	7%	13%	0%
	Yes, twice	7%	0%	22%
Within the last three years, have you been fined for a traffic offence while driving your car? (Excluding parking offences)	No	72%	69%	78%
	Yes - not specified	0%	0%	0%
	Yes - speeding	24%	31%	11%
	Yes - running a red light	0%	0%	0%
	Yes - multiple offences (speeding + running a red light, speeding + phone offence)	3%	0%	11%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	7%	0%	22%
	Discreet, average	62%	69%	56%
	Sportive, ambitioned	28%	31%	11%
	Risk-taking, offensive	3%	0%	11%
How confident you are concerning your own driving skills?	Insecure	0%	0%	0%
	Neutral	21%	25%	0%
	Confident	59%	63%	67%
	Very confident	21%	13%	33%

[15] Note that 4 drivers are not classified as Type A or Type B as they were excluded from outcome analysis. The average age of these participants was 28.0, hence the average for Type A and Type B drivers does not then equal the average for all drivers.

Table 71: Driver ADAS, accident and offence history, and confidence, per driver change type (Greece)

Question / Response Option		EL (All)	EL (A)	EL (B)
Number of participants (drivers)		80[16]	30	18
Participant gender	Male	48 (60%)	17 (57%)	10 (56%)
	Female	32 (40%)	13 (43%)	8 (44%)
Participants mean age (years)		31.5	29.7	34.0
Standard deviation of age (years)		10.1	8.1	13.4
Years driving experience (range, average)		1 - 41, 10.7	1 - 27, 10.0	1 - 41, 12.6
Which ADAS are present in your car? (Percentage replied equipped)	Automatic emergency braking	16%	17%	17%
	Blind spot warning	5%	3%	6%
	Drowsiness alert	9%	10%	11%
	Forward collision warning	20%	13%	33%
	High speed alert	14%	10%	17%
	Lane keeping assistance	15%	13%	17%
	Night vision & pedestrian detection	3%	3%	6%
In the last three years, have you been involved in an accident with your car, which was self-inflicted?	No	83%	73%	89%
	Yes, once	16%	27%	11%
	Yes, twice	1%	0%	0%
Within the last three years, have you been fined for a traffic offence while driving your car? (Excluding parking offences)	No	93%	93%	100%
	Yes - not specified	0%	0%	0%
	Yes - speeding	6%	7%	0%
	Yes – DUI (intoxicated)	1%	0%	0%
	Yes - running a red light	0%	0%	0%
	Yes - multiple offences (speeding + running a red light, speeding + phone offence)	0%	0%	0%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	9%	7%	6%
	Discreet, average	70%	67%	72%
	Sportive, ambitioned	19%	23%	17%
	Risk-taking, offensive	3%	3%	6%
How confident you are concerning your own driving skills?	Insecure	0%	0%	0%
	Neutral	26%	20%	44%
	Confident	53%	53%	28%
	Very confident	21%	27%	28%

[16] Note that a lot of drivers are not classified as Type A or Type B as they were excluded from outcome analysis. Also, 8 drivers included in the outcome analysis do not have questionnaire data.

Table 72: Driver ADAS, accident and offence history, and confidence, per driver change type (UK)

Question / Response Option		**UK (All)**	**UK (A)**	**UK (B)**
Number of participants (drivers)		54	37	12
Participant gender	Male	33 (61%)	22 (59%)	9 (75%)
	Female	21 (39%)	15 (41%)	3 (25%)
Participants mean age (years)[17]		45.4	47.9	37.8
Standard deviation of age (years)		13.6	14.1	10.9
Years driving experience (range, average)		2 - 60, 25.0	2 - 60, 27.1	2 - 34, 18.3
Which ADAS are present in your car? (Percentage replied equipped)	Automatic emergency braking	2%	3%	0%
	Blind spot warning	0%	0%	0%
	Drowsiness alert	0%	0%	0%
	Forward collision warning	9%	5%	17%
	High speed alert	0%	0%	0%
	Lane keeping assistance	0%	0%	0%
	Night vision & pedestrian detection	0%	0%	0%
In the last three years, have you been involved in an accident with your car, which was self-inflicted?	No	85%	84%	92%
	Yes, once	13%	16%	0%
	Yes, twice	2%	0%	8%
Within the last three years, have you been fined for a traffic offence while driving your car? (Excluding parking offences)	No	80%	78%	83%
	Yes - not specified	0%	0%	0%
	Yes - speeding	19%	19%	17%
	Yes - running a red light	2%	3%	0%
	Yes - multiple offences (speeding + running a red light, speeding + phone offence)	0%	0%	0%
Please select with which of the following driving styles you identify the most.	Less experienced, hesitant	2%	3%	0%
	Discreet, average	74%	81%	50%
	Sportive, ambitioned	15%	8%	33%
	Risk-taking, offensive	9%	8%	17%
How confident you are concerning your own driving skills?	Insecure	2%	3%	0%
	Neutral	11%	14%	8%
	Confident	61%	65%	67%
	Very confident	26%	19%	25%

[17] Note that 5 drivers are not classified as Type A or Type B as they were excluded from outcome analysis. The average age of these participants was 45.6, hence the average for Type A and Type B drivers does not then equal the average for all drivers.

Annex 2: Questionnaires

For information, the set of questions asked to car drivers is included here.

Screening Questionnaire

Question Ref	Question Text
SQ_Participant_ID	*This is the participant ID generated using the on-boarding system*
SQ_Gender	Participant gender
SQ_Nationality	Participant nationality
SQ_Year_of_birth	Participant year of birth (year only, not full DOB for data protection)
SQ_Age	Participant age at time entered study.
SQ_ Age_got_driving_license	At what age did you receive your driver's license (i.e., when did you drive legally by yourself)?
SQ_Years_driving	Number of years license held for.
SQ_Vehicle_brand	Vehicle brand (e.g., Renault)
SQ_Vehicle_model	Vehicle model (e.g., Clio)
SQ_Vehicle_age	How old is your vehicle in years?

Screening Telephone Call

Question Ref	Question Text
STC_Second_Nat	Participant second nationality (if applicable)
STC_Highest_lev_education	Participant highest level of education
STC_Current_occupation	Participant current (main) occupation
STC_Employment_stat	Participant current employment status
STC_Net_income	What is the monthly net income for your household?
STC_Med_condition_decleration	Can you declare that you are not suffering from a medical condition that would be considered a legal exclusion to drive?
STC_First_registered	Vehicle date of first registration (include year only)
STC_Fuel_type	Vehicle fuel type
STC_ Engine_CC	Vehicle engine size/cylinder capacity (cc) (e.g., 500cc):
STC_Engine_HP	Vehicle engine power (horsepower) (e.g., 105 HP):
STC_Vin_No	Vehicle vin/Chassis/Frame no.:
STC_Gearbox	Vehicle gearbox type
STC_Disability_mod	Has this vehicle been modified to cope with physical limitations of the driver?
STC_Veh_modification_description	If yes, briefly describe which modifications took place?
STC_Number_other_drivers	How many other drivers use the above vehicle in an average week?
STC_Drvr_1_split	How is the use of this car split between all of the drivers who use it?
STC_Drvr_2_split	
STC_Drvr_3_split	
STC_Drvr_4_split	
STC_Weekly_km	How many kilometres (make an estimate) do you travel on average per week with this car (during COVID-19)?
STC_Urban	How much do you drive on urban roads (e.g. roads with a maximum speed limit of 30 km/h or 50 km/h)?...............%
STC_Rural	How much do you drive on rural roads (e.g. roads with a maximum speed limit of 70 km/h or 90 km/h)? ....%
STC_Motorway	How much do you drive on motorways e.g. roads with a maximum speed limit 120 km/h)? ....%

Entry Questionnaire

Question Ref	Question Text
EQ1a_Adaptive_cruise_control EQ1b_Forward_collision_warning EQ1c_NV_PD EQ1d_Traffic_sign_recognition EQ1e_Lane_keeping_Assistance EQ1f_Blind_spot_warning EQ1g_Drowsiness_alert EQ1h_Parking_assist EQ1i_High_speed_alert EQ1j_Automatic_emergency_braking EQ1k_Other EQ1l_other_please_specify	Which Advanced Driving Assistance Systems are present in your car? EQ1c_NV_PD = night vision and pedestrian detection
EQ2a_Adaptive_cruise_control EQ2b_Forward_collision_warning EQ2c_NV_PD EQ2d_Traffic_sign_recognition EQ2e_Lane_keeping_Assistance EQ2f_Blind_spot_warning EQ2g_Drowsiness_alert EQ2h_Parking_assist EQ2i_High_speed_alert EQ2j_Automatic_emergency_braking EQ2k_Other	How often do you use the following Advanced Driving Assistance Systems that are present in your car? EQ2c_NV_PD = night vision and pedestrian detection
EQ3a_Useful EQ3b_Increase_perform EQ3c_Understandable EQ3d_Easy EQ3e_Good_idea EQ3f_Maintain_safe EQ3g_Comfortable EQ3h_Attention EQ3i_Accident_risk EQ3j_Trust EQ3k_Distract	Indicate to what extent you agree with the following statements about ADAS in general. ADAS are useful while driving Using ADAS increases my driving performance My interaction with ADAS is clear and understandable I find ADAS easy to use Using ADAS is a good idea I can maintain safe driving behaviour while using ADAS I will feel more comfortable doing other things (e.g., adjusting the radio) with ADAS Using ADAS information requires increased attention Using ADAS information decreases the accident risk I trust the information I receive from ADAS ADAS distract me while driving
EQ4a_Speed_limit_built_up EQ4b_Speed_limit EQ4c_Sleepy EQ4d_Tired EQ4e_Mobile_phone EQ4f_VRU_close EQ4g_Illegal_overtake EQ4h_Safe_distance EQ4i_Driving_lane	Please estimate: over the last year, how often did you as a car driver… drive faster than the speed limit inside built-up areas? drive faster than the speed limit? drive when you were so sleepy that you had trouble keeping your eyes open? realize that you were actually too tired to drive? used a hand-held mobile phone while driving? drive to close to a vulnerable road user (pedestrian, moped, cyclist, etc.)? illegally overtake another vehicle? drive without respecting a safe distance to the vehicle in front? cross the outer edges of the driving lane?
EQ5_Driving_style	Please select with which of the following driving styles you identify the most.
EQ6_Driving_confidence	How confident you are concerning your own driving skills?
EQ7_Driving_is	Driving is….
EQ8a_Skill EQ8b_Hazards EQ8c_Crash_risk	How do you think you compare to the average driver? Regarding general driving skills, I am: Regarding the ability to cope with hazards in traffic, I am:

	Regarding your risk of being involved in a crash, I am:
EQ9a_Police_close_following EQ9b_Overtake EQ9c_Fast EQ9d_Small_gap EQ9e_Faster_speed_limit EQ9f_Risky_overtake EQ9g_Speed_drive_careful EQ9h_Know_risks EQ9i_Closer_recommended EQ9j_Closer_flow	Please indicate to which extent you agree with the following statements. People stopped by the police for close-following are unlucky because lots of people do it It is quite acceptable to take a slight risk when overtaking I know exactly how fast I can drive and still drive safely Some people can drive safely even though they only leave a small gap behind the vehicle in front Even driving slightly faster than the speed limit makes you less safe as a driver I think it is okay to overtake in risky circumstances as long as you drive within your own capabilities It's okay to drive faster than the speed limit as long as you drive carefully I know exactly what risks I can take when I overtake. It is quite acceptable to drive closer to the vehicle in front than is recommended Sometimes you have to drive in excess of the speed limit in order to keep up with the traffic flow
EQ10a_Attention EQ10b_Keeping_distance EQ10c_Adjusting_speed EQ10d_Conforming_speed_limit	Please rate your own driving skills in regard to the following situations or manoeuvres. Paying attention to other road-users Keeping sufficient following distance Adjusting the speed to the conditions Conforming to the speed limits
EQ11a_Benefits EQ11b_Needed_safe EQ11c_Skills EQ11d_Competent EQ11e_Important EQ11f_Comfortable EQ11g_Personally_important EQ11h_Obligation EQ11i_Friends_safe EQ11j_Colleagues_safe EQ11k_I_control EQ11l_Safe_easy	Please indicate to which extent you agree with the following statements. I know the benefits of safe driving I know what is needed to drive safely I have the skills to drive safely I feel competent enough to drive safely Safe driving is important to avoid crashes Safe driving makes me feel comfortable For me personally, safe driving is important Safe driving should be a personal obligation My friends think safe driving is important My colleagues find it important to drive safely I control whether I drive safely or not For me, safe driving is easy to do
EQ12_Accident_three_years	Within the last three years, have you been involved in an accident with your car, which was <u>self-inflicted</u>?
EQ13a_accident_1 EQ13b_accident_2 EQ13c_accident_3 EQ13d_accident_4	If yes, how severe was this accident / were these accidents? Accident 1 Accident 2 Accident 3 Accident 4
EQ14_Traffic_offence	Within the last three years, have you been fined for a traffic offence while driving with your car?
EQ15a_Speeding_offence EQ15b_DUI_offence EQ15c_Tailgating_offence EQ15d_Phone_offence EQ15e_Parking_offence EQ15f__Illegal_overtaking EQ15g_Running_light EQ15h_Running_stop EQ15i_Running_yield EQ15j_Not_stop_ped EQ15k_Other	If yes, for which offence have you been fined within the last three years? Speeding Driving under the influence Tailgating (unsafe following distance) Using handheld phone while driving Parking offence Illegal overtaking Running a traffic light Running a stop sign Running a yielding sign

	Not stopping at a pedestrian crossing Other (please specify)
EQ16a_Sit_read EQ16b_Watching_TV EQ16c_Sitting_inactive EQ16d_Car_passenger EQ16e_Lying_down EQ16f_Sitting_talking EQ16g_Sitting_lunch_alcohol EQ16h_Car_stopped	How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? Sitting and reading Watching TV Sitting, inactive in a public place (e.g., a theatre or a meeting) As a passenger in a car for an hour without a break Lying down to rest in the afternoon when circumstances permit Sitting and talking to someone Sitting quietly after a lunch without alcohol In a car, while stopped for a few minutes in the traffic
EQ17_General_sleep_rating	In general, how would you rate your sleep in the last 3 months?
EQ18_Diagnosed_sleep_disorder EQ18a_Yes_what	Have you ever been diagnosed with a disorder or condition which affects your sleep, e.g. obstructive sleep apnoea? If yes, which condition or disorder?
EQ19_Fight_sleep_in_car	How often do you have to fight sleepiness in order to stay awake while driving the car?
EQ20_Stop_because_sleepiness	In the past 12 months, have you had to stop the car due to sleepiness?
EQ21_Sleepiness_Wanted_to_stop	In the past 12 months, have you wanted to stop the car due to sleepiness, but been unable to?
EQ22_Asleep_while_driving	In the past 12 months, have you fallen asleep whilst driving the car?
EQ23_Crash_blame_sleep	In the last 10 years have you experienced an incident or crash with your car where sleepiness was partly or solely to blame?
EQ24_n1_Epilepsy EQ24_n2_Parkinsons_Disease EQ24_n3_Multiple_Sclerosis EQ24_n4_Stroke EQ24_n5_Migraines EQ24_n6_Dizziness EQ24_n7_Other EQ24_m1_Limited_flexibility EQ24_m2_Arthritis EQ24_m3_Artificial_limbs EQ24_m4_Paralysis EQ24_m5_MMD EQ24_m6_Other EQ24_c1_High_blood_pressure EQ24_c2_Low_blood_pressure EQ24_c3_Heart_attack EQ24_c4_Pacemaker EQ24_c5_Bypass_surgery EQ24_c6_Other EQ24_h1_Difficulty_hearing EQ24_h2_Deafness EQ24_h3_Hearing_aid EQ24_v1_Near_sighted EQ24_v2_Farsighted EQ24_v3_Reading_glasses EQ24_v4_Colour_blindness EQ24_v5_Blind_one_eye EQ24_v6_Poor_night_vision EQ24_v7_Other	Do you have any diseases of the following categories that you are aware of? If yes, which ones? N = Neurological 1 – Epilepsy, 2 – Parkinson's disease, 3 – Multiple sclerosis, 4 – Stroke, 5 – Migraines, 6 – Dizziness, 7 – Other M = Muscles, skeletal 1 – Limited flexibility, 2 – Arthritis, 3 – Artificial limbs, 4 – Paralysis, 5 – Muscle and movement disorders, 6 – Other C = Cardio-vascular 1 – High blood pressure, 2 – Low blood pressure, 3 – Heart attack, 4 – Pacemaker, 5 – Bypass surgery, 6 – Other H = Hearing 1 – Difficulty hearing, 2 – Deafness, 3 – Hearing aid V = Vision 1 – Objects far aware are blurry, 2 – Objects close up are blurry, 3 – Reading glasses needed, 4 – Colour blindness, 5 – Blind in one eye, 6 – Poor night vision, 7 – Other

User Experience Questionnaire

Question Ref	Question Text
UX1_Identify_easy	Identifying myself in the vehicle is easy
UX2_Display_works	The display in the car works as it should (e.g. turns on when the car is turned on, closes when the car is turned off, does not freeze while driving ...)
UX3_Warnings_clear	The warnings are clear, I understand the meaning of all the symbols and sounds
UX4_Alerts_correct	The alerts are correct. They correctly reflect the situation around my vehicle (e.g. speed indications, road markings, other road users in the vicinity ...)
UX5_Alerts_timely	The alerts are always given in a timely manner allowing me to adjust my actions in a timely manner
UX6_Warnings_distracting	The warnings are sometimes distracting
UX7_alerts_aware	The alerts make me more aware of my actions while driving
UX8_Additional_comments	If you would like to add something regarding the statements above or if you would like to share additional experiences, comments or suggestions with us, you can do so here
UX9_Alerts_missing	Are there any particular alerts that you think might be useful but are currently missing?
UX10_App_install_easy	Installing the app on my mobile via the link I received by email was easy for me
UX11_App_update_easy	Updating the app is easy for me
UX12_App_overview_easy	I can easily find the overview of my trips in the app
UX13_App_filter_easy	I find it easy to filter trips according to a certain period
UX14_Trip_info_clear	The trip information (start and end point, duration, distance and scores) shown per trip is clear to me
UX15_Visualise_trip_easy	Visualizing my completed trips on a map in the app is easy
UX16_Specific_events_easy	I find viewing specific events identified during a ride easy.
UX17_Detailed_info_clear	The detailed information (e.g., dashcam videos) shown per event on a trip is clear to me
UX18_Scores_clear	The information shown via the 'scores' button is clear to me
UX19_Period_scores_easy	Selecting a period of which I want to see 'scores' is easy for me
UX20_Message_useful	I find the 'messages' button a useful addition in the app
UX21_Respond_message_easy	I find responding to posted messages in the app easy
UX22_Settings_useful	I find the information behind the 'settings' button useful
UX23_Info_tile_useful	The 'Info' tile in the middle of the screen shows new pros, cons and facts about safe driving every day. This information is useful to me.
UX24_Thumbs_up_useful	I can rate the information I get from the 'Info' tile with a [thumbs up] or a [thumbs down]. I find this option useful.
UX25_Tips_useful	The 'Tips' tile in the middle of the screen offers new practical tips on safe driving every day. This information is useful to me.
UX26_Tile_thumbs_up	I can rate the information I receive via the 'Tips' tile with a [thumbs up] or a [thumbs down]. I find this option useful.
UX27_Goals_Useful	The 'Goals and badges' tile in the middle of the screen offers me challenges for different parameters to improve my own driving performance. I find this interesting.
UX28_New_goals_motivating	Achieved goals regarding a specific parameter are made more challenging through new goals. I enjoy these new goals because they motivate me to do even better.
UX29_Badges_interesting	By completing enough goals on a specific parameter, I can collect badges. I find this interesting.
UX30_Scoreboard_useful	The 'Scoreboard' tile in the centre of the screen gives me the chance to compare my own driving performance with that of my colleagues. I find this interesting.

UX31_additional_statements	If you would like to add additions regarding statements above or if you would like to share additional experiences, comments or suggestions with us, you can do so here:
UX32_Useful_info_missing	Is there any information that might be useful to you but is currently missing from the app?

Exit Questionnaire

Question Ref	Question Text
EX1a_Increase_performance EX1b_Destination_safely EX1c_Easy_understand EX1d_Annoying EX1e_Good_idea EX1f_More_interesting EX1g_Proud EX1h_People_Like_use EX1i_Maintain_safe EX1j_Knowledge EX1k_Do_not_understand EX1l_Affect_negative_way EX1m_Increased_attention EX1n_Distracts EX1o_Safer EX1p_More_aware EX1q_Depend EX1r_Comfortable_other_things EX1s_Continue_use EX1t_Recommend	Indicate to what extent you agree with the following statements about the i-DREAMS system. Performance expectancy Using the system increases my driving performance If I use the system, I will reach my destination safely Ease of use / effort expectancy I think the i-DREAMS system is easy to understand I think the i-DREAMS system is annoying Attitude towards using technology Using the system is a good idea The system makes driving more interesting Social influence I would be proud to show the system to people who are close to me In general, people who I like would encourage me to use the system Facilitating conditions While using the system I can maintain safe driving behaviour I have the knowledge necessary to use the system Anxiety I am afraid that I do not understand the system I am confident that the system does not affect my driving in a negative way Perceived Safety Using the system information requires increased attention The system distracts me from driving Perceived Usefulness I think using the i-DREAMS system makes me a safer driver I think using the i-DREAMS system makes me more aware of my surroundings (other vehicles, lane position, etc.) Trust I think I can depend on the i-DREAMS system I will feel more comfortable doing other things (e.g., adjusting the radio) with the i-DREAMS system Behavioural Intention to Use If I had a choice, I would continue to use the i-DREAMS system I would recommend the i-DREAMS system to other drivers
EX2a_Clear_in_general EX2ai_Why EX2aii_Suggestions_improve EX2b_Visual_symbols EX2bi_Why EX2bii_Suggestions_improve EX2c_Sounds EX2ci_Why EX2cii_Suggestions_improve	Indicate to what extent you find the i-DREAMS system clear in general How clear do you find the i-DREAMS system in general? Why? Suggestions to improve? How clear do you find the visual symbols of the system in general? Why? Suggestions to improve? How clear do you find the sounds of the system in general? Why? Suggestions to improve?

EX3a_Know_benefits EX3b_Know_safely EX3c_Skills_safely EX3d_Competent EX3e_Safe_important EX3f_Safe_comfortable EX3g_Personally_important EX3h_Personal_obligation EX3i_Friends_safe_important EX3j_Colleagues_safe_important EX3k_Control_safety EX3l_Safe_easy	Please indicate to which extent you agree with the following statements I know the benefits of safe driving I know what is needed to drive safely I have the skills to drive safely I feel competent enough to drive safely Safe driving is important to avoid crashes Safe driving makes me feel comfortable For me personally, safe driving is important Safe driving should be a personal obligation My friends think safe driving is important My colleagues find it important to drive safely I control whether I drive safely or not For me, safe driving is easy to do
EX4a_Power EX4b_Achievement EX4c_Hedonism EX4d_Stimulation EX4e_Self_direction EX4f_Universalism EX4g_Benevolence EX4h_Tradition EX4i_Conformity EX4j_Security	Please rate the importance of the following values as a life-guiding principle for you. Use the 8-point scale in which: 0 indicates that the value is opposed to your principles and 1 indicates that the value is not important for you, 4 indicates that the value is important 8 indicates that the value is of supreme importance for you. Power (social power, authority, wealth) Achievement (success, capability, ambition, influence on people and events) Hedonism (gratification of desires, enjoyment in life, self-indulgence) Stimulation (daring, a varied and challenging life, an exciting life) Self-direction (creativity, freedom, curiosity, independence, choosing one's own goals) Universalism (broad-mindedness, beauty of nature and arts, social justice, a world at peace, equality, wisdom, unity with nature, environmental protection) Benevolence (helpfulness, honesty, forgiveness, loyalty, responsibility) Tradition (respect for tradition, humbleness, accepting one's portion in life, devotion, modesty) Conformity (obedience, honouring parents and elders, self-discipline, politeness) Security (national security, family security, social order, cleanliness, reciprocation of favours)
EX5a_Queue_turn_left EX5b_Fail_notice EX5c_Fail_check EX5d_Brake_too_quckly EX5e_On_turning_left_cyclist EX5f_Miss_give_way EX5g_Attempt_overtake EX5h_Underestimate_speed EX5i_Hit_something EX5j_Intend_drive_A EX5k_Get_wrong_lane EX5l_Switch_one_thing EX5m_Attempt_drive_third_gear EX5n_Forget_carpark EX5o_Misread_sign EX5p_Realise_no_recollection	Please estimate, how often do you… queue to turn left onto a main road, you pay such close attention to the main stream of traffic that you nearly hit the vehicle in front of you fail to notice that pedestrians are crossing when turning into a side street from a main road fail to check your rear-view mirror before pulling out, changing lanes, etc. brake too quickly on a slippery road or steer the wrong way in a skid on turning left nearly hit a cyclist who has come up on your inside miss “Give Way” signs and narrowly avoid colliding with traffic having right of way attempt to overtake someone that you had not noticed to be signalling a right turn underestimate the speed of an oncoming vehicle when overtaking hit something when reversing that you had not previously seen

	intending to drive to destination A, you "wake up" to find yourself on the road to destination B get into the wrong lane approaching a roundabout or a junction switch one thing, such as the headlights, when you meant to switch on something else, such as the wipers attempt to drive away from the traffic lights in third gear forget where you left your vehicle in a car park misread the signs and exit from a roundabout on the wrong road. realize that you have no clear recollection of the road along which you have just been traveling
EX6a_Avoid_dark EX6b_Avoid_Urban EX6c_Avoid_Motorway EX6d_Avoid_bad_weather	Please indicate to which extent you agree with the following statements. I try to avoid driving in the dark I try to avoid driving in urban areas I try to avoid using highways / motorways I try to avoid driving in bad weather
EX7a_Explore_strange EX7b_Restless_at_home EX7c_Frightening_things EX7d_Wild_parties EX7e_Pre_planned_trips EX7f_Unpredictable_friends EX7g_Bungee_jump EX7h_Exciting_experiences	Please indicate to which extent you agree with the following statements. I would like to explore strange places I get restless when I spend too much time at home I like to do frightening things I like wild parties I would like to take off on a trip with no pre-planned routes or timetables I prefer friends who are excitingly unpredictable I would like to try bungee jumping I would love to have new and exciting experiences, even if they are illegal
EX8a_Driving_too_slow EX8b_Weaving_traffic EX8c_Slower_than_reasonable EX8d_Slow_vehicle EX8e_Run_red_light EX8f_Towards_at_night EX8g_Right_behind_lights_on EX8h_Speeds_up_pass EX8i_Slow_in_parking EX8j_Pulls_in_front EX8k_Obscene_gesture EX8l_Someone_way_over_limit EX8m_Someone_yells EX8n_Truck_kicks_up	Please indicate how angry you would feel if you came across the following situations while driving. Someone is driving too slowly in the passing lane holding up traffic Someone is weaving in and out of traffic Someone is driving slower than reasonable for the traffic flow A slow vehicle on a mountain road will not pull over and let people by Someone runs a red light or stop sign Someone coming toward you at night does not dim their headlights At night someone is driving right behind you with bright lights on Someone speeds up when you try to pass them Someone is slow in parking and holding up traffic Someone pulls right in front of you when there is no one behind you Someone makes an obscene gesture toward you about your driving Someone is driving way over the speed limit Someone yells at you about your driving A truck kicks up sand or gravel on the vehicle you are driving
EX9_Rate_experience_participating	How would you rate your experiences participating in this study? Please choose only one of the following
EX10_Bring_to_our_attention	Is there anything in particular that you would like to bring to our attention or any suggestion that can help improving the i-DREAMS project or future projects of the same nature? Please write your answer here:

Annex 3: Tram Simulator Trial Detailed Methodology

A3.1 Procedure

Prior to the simulator session, all drivers were provided with an information pack which gave an overview of the study and included the participant information sheet. On the day of the simulator study, drivers arrived at the designated simulator room in the tram depot, were welcomed, and asked to provide verbal confirmation that they did not have any COVID symptoms. The background of the study was then explained as well as the procedure for the session. A timeline of the simulator session is provided below (Table 73).

Table 73: Simulator session timeline

Duration	Scenario	Procedure
-	Before the session	• Drivers provided with an information pack
10 mins	Welcome and briefing	• Welcome drivers and explain the study • Participant information sheet and signed consent • Complete entry questionnaire • Familiarise drivers with measures
5 mins	Familiarisation drive	• Driving without the *i*-DREAMS system active • Practice tramstops and stopping • Practice additional measures (e.g., KSS, counting)
30 mins	Baseline drive	• Driving without the *i*-DREAMS system active • Urban and suburban driving inc. transitions • Differences in workload • Events include signals, signal changes, speed changes, VRU's, manipulated events
5 mins	Comfort break	*No smoking or caffeine intake*
5 mins	Questionnaire	• After baseline drive questionnaire and discussion
30 mins	Intervention drive	• Driving with the *i*-DREAMS system active • Urban and suburban driving inc. transitions • Differences in workload • Fixed timing interventions (speed and VRU detection) • Events include signals, signal changes, speed changes , VRU's, manipulated events (different to baseline)
10 mins	Manipulation drive	• Manipulation of driving behaviour to trigger *i*-DREAMS system • Speed warning, VRU detection and fatigue warning discussion
15 mins	Closing	• Debrief • Exit questionnaire / feedback from drivers

Note. KSS, Karolinska Sleepiness Scale

It was reported by the tram company that during typical tram driving, there was a noticeable difference in workload between the urban/town centre and suburban sections. Therefore, to increase workload in the simulated environment, drivers were asked to complete an additional counting task during a section of the main drives (baseline and intervention). This task was a modification of the Serial Sevens Subtraction Task (SSST, Hayman, 1942), whereby participants are asked to count backwards in sevens. For the simulator study, tram drivers were asked to count backwards in groups of 10, starting at 1000, for three tramstops inclusive during the urban town centre section of the main drives (baseline and intervention). Drivers were instructed to count backwards out loud without rushing and reminded that their primary task was to drive the tram. If drivers reached zero before the last tramstop, they were instructed to start again. Both the KSS and counting were practiced during the familiarisation drive.

A3.2 Simulator drives

The study was conducted using a simulated environment of the tram route frequently driven by participants, with pre-programmed speed limits which were set for tramstops, crossings, and all track routes. The weather was programmed as daytime driving with fine weather conditions. Traffic and pedestrians were programmed to be randomly generated (apart from those included in the manipulated events, see section A3.3 below).

In total, participants were asked to complete four drives during the simulator session: a five-minute familiarisation drive, two 30-minute main drives (Drive A and Drive B), and a five-to-ten-minute manipulation drive.

Familiarisation drive

The familiarisation drive was designed to be a drive of approximately five minutes for the participants to familiarise themselves with the simulator controls and measures used during the study (e.g., KSS and the counting task). The route consisted of three suburban tramstops heading towards the urban town centre.

Baseline and intervention drives

Of the two 30-minute main drives (Drive A and Drive B), one of the drives was aimed at collecting baseline data with no interventions from the *i*-DREAMS technology, and the second drive was designed to be an intervention drive, with the *i*-DREAMS technology switched on. The same route was used for both Drive A and Drive B, however the manipulated events that were included were different in Drive A compared to Drive B (see section A3.3 for more detail). The main drives were designed to be counterbalanced to avoid learning, so that either Drive A or Drive B could be used for baseline or intervention and were alternated between the participants.

The route tram drivers regularly operate contains segregated suburban track and multi-user urban track. Coinciding with this are differences in workload, with potentially lower workload in the segregated suburban sections and increased workload on the urban sections. To be reflective of typical tram driving, the chosen route featured both sections of suburban and urban track, with transitions between the two (as shown in Figure 39). In total (including start and end tramstops) 15 tramstops were included in the route, 10 in the suburban sections, and five in the urban section.

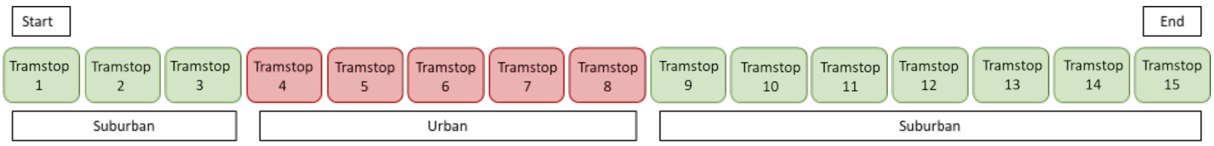


Figure 39: The route and tramstops used for Drive A and Drive B (both intervention and baseline drive)

<u>Manipulation drive</u>

The final simulator drive was the manipulation drive, a short drive starting at the end of the urban section used in the previous drives, lasting for five tramstops. During this drive, drivers were asked to manipulate their driving behaviour (e.g., speed) to trigger the *i*-DREAMS system. This ensured that all drivers experienced the *i*-DREAMS system and warnings.

A3.3 Manipulated Events

During the baseline and intervention drives (Drive A and Drive B), a series of manipulated events were included. These manipulated events were pre-programmed events used predominantly for training purposes (e.g., to practice certain events that may happen during real-world tram driving), that occur at certain locations within the simulated tram route. As such, the type and design of event that could be included was limited to what was available within the tram simulator settings. To avoid learning effects, the events were different for each of the main drives (Drive A and Drive B). However, an attempt was made to match events, resulting in matched pairs. In total, six events were included in each main drive, four from the matched pairs, and two events that were the same in both drives (these were included in the town centre of the urban section to increase workload). Table 74 below details the events in each of the main drives and the location, including a brief description of the event.

Table 74: Overview of the events included in the scenario drives to be used for baseline or intervention

Main drive	Event number	Location	Description	Matched pair
(Drive A)	Event 4	Before tramstop 4	Pedestrian runs between platforms as tram approaches	Matched with event 5
	Event 6	Approaching tramstop 5	Group of pedestrians in swept path	Matched with event 19
	Event 18	Between tramstop 6 and tramstop 7	Signal change	Included in both drives
	Event 21		Cyclist riding towards tram	Included in both drives
	Event 39	Between tramstop 9 and 10	Pedestrian runs across foot crossing	Matched with event 40
	Event 41	Tramstop 11	Pedestrian emerges behind tram	Matched with event 49
(Drive B)	Event 5	Approaching tramstop 5	Cyclist crosses trams path	Matched with event 4
	Event 18	Between tramstop 6 and tramstop 7	Signal change	Included in both drives
	Event 21		Cyclist riding towards tram	Included in both drives
	Event 19		Lorry in swept path	Matched with event 6

	Event 40	Between tramstop 9 and 10	Pedestrian runs across foot crossing	Matched with event 39
	Event 49	After tramstop 14	Pedestrian crosses behind stopped tram	Matched with event 41

A3.4 Recorded variables

The following variables were recorded during the simulator study:

- Questionnaire data
    - Participant demographics
    - Experience (tram driving, sleepiness, safety systems)
    - Use and opinion of safety systems
- Simulator data
    - Time
    - Distance travelled
    - Speed
    - Speed limit
    - Signal status
- Subjective sleepiness (KSS scores)
- Heart rate data and inferred sleepiness level (wearable)
- Triggering of *i*-DREAMS system – speed and VRU detection (Mobileye and gateway)

A3.5 Data analyses

Analysis of *i*-DREAMS alerts – speeding and VRU detection

The information regarding the *i*-DREAMS speeding and VRU alerts was downloaded from the system into separate database. Data was sorted per participant, and per drive (e.g., baseline or intervention). Speeding alerts were contained within their own database detailing the level of the Safety Tolerance Zone (STZ) the driver was at during the drive. Results were collated in terms of two risk levels of the STZ – STZ level 2, dangerous driving and STZ level 3, avoidable accident (STZ 1 would be deemed normal driving). These alerts were extracted and counted for both baseline and intervention per participant. For the VRU detection, the relevant information was extracted into a separate database. The warnings were then identified within the data, separating the detection of a VRU from a VRU collision warning (which is when the alert is given). This data was compiled per participant for the intervention and baseline drives. The number of times the alerts triggered, the mean and the range per data set is presented below.

Qualitative data from participant interviews

Participants were interviewed after the baseline drive and questioned about their recollection, anticipation and prediction related to one of the events they had encountered during the drive. Drivers were also asked about any contrast in their driving between segregated suburban track and urban track and about the effect on their driving of the presence of pedestrians and cyclists. After the intervention drive and during and after the manipulation drive participants were interviewed again on the subject of the warnings offered by the *i*-DREAMS system: fatigue, VRU event and speed. As the fatigue warnings were based on both time on task, and also the physiological measurement of fatigue using the wearables to detect changes in heart rate. As the testing time was restricted to fit with drivers shifts, was conducted during the day, and there was no experimentally induced sleep restriction, the drivers were unlikely to experience the *i*-DREAMS fatigue warning during the main simulator drives. Therefore, the drivers were shown the symbol visually with the concept explained and asked questions about their opinion of this within a tram driving operation.

Since the nature of the study was an in-depth investigation of the interaction of tram drivers with technology in their cabs a thematic (bottom up) approach was employed (Braun & Clarke, 2006). This decision was made in order to allow themes to develop both from the research questions and the narratives of the research participants. To this end, the recordings of the interviews (413 minutes total) were transcribed verbatim, de-identified and analysed by a team of two researchers using thematic analysis (Braun & Clarke, 2006) via NVivo (ver. 12.0) software. The two researchers had been present at the majority of the interviews (Researcher 1 attended 18 of the 28 interviews and Researcher 2 attended ten interviews; five of these were in common).

Researcher 1 analysed data from the first three participants and produced codes (or nodes as they are known in NVivo) which coincided with the questions which had been asked and therefore also with the research questions. The dialogue from each interview was coded using open coding (Saldana, 2009) which identified units of information that addressed the research aims. The previous knowledge of the research team and the research questions were also taken into consideration. The two researchers then met to discuss this analysis and subsequently to refine and agree a coding dictionary which was then used to analyse the remainder of the interviews.

Given the size and richness of the dataset, this approach created a robust process, ensuring all data were captured and considered during the analysis.

Sleepiness analysis from KSS and wearable data

Sleepiness data was provided by verbal KSS scores and Inter-beat Interval (IBI) wearable data, which was attributed to KSS bands (KSS1-5, 6-7, 8-9). To assess the sleepiness state of the participants, the verbal KSS scores were first analysed, focusing on the mean KSS score provided by the drivers, comparing between both Drive A and Drive B, and baseline and intervention drives. The number of drivers reporting scores in each of the KSS bands from the wearable data was also reported.

The second part of the analysis was a comparison between the KSS bands attributed to the wearable data and the verbal KSS scores. Analysis consisted of the number of complete sets and the number of times the drivers entered the KSS bands within the baseline and intervention drives, as well as a comparison between the first and last KSS scores and the first and last data points from the wearables.

Questionnaire data

At the beginning of the tram simulator trial sessions all of the participating drivers were asked to complete an entry questionnaire. This was completed before any of the simulator drives. The questionnaire provided some demographic data and was focused on participants' relationships with safety systems, both those which are currently situated within the tram cab and any additional systems or warnings which participating tram drivers considered would be beneficial for the driving context. The demographic data was analysed to provide an understanding of the participating drivers' previous experience with the tram simulator and their enjoyment of driving. The remainder of the analysis concentrated on participants' responses to direct questions about safety systems; these were a mixture of ratings and open-ended questions. In order to relate these responses to the *i*-DREAMS system, a consideration was made of issues and views which are currently addressed by *i*-DREAMS, those which show potential for integration into the system and those which are outside the scope of the *i*-DREAMS system.

A3.6 Limitations of the study design

As with every study, there are some limitations associated with the one being described here. Firstly, the sampling method used was self-selection, with participants volunteering to take part, and the timings of sessions being organised by the Operations Director and control room to ensure there was adequate cover for the drivers shifts. Therefore, this method may have influenced which drivers participated in the study. Although the gender balance in the study does represent the gender split in tram driving generally (predominantly male), the sample does not represent each gender equally. However, this was a result of the self-selection aspect of the sampling method in that fewer female drivers volunteered as participants. It is also likely to be due to there being fewer female drivers available in the first place. As the drivers were participating during their shift time, there were certain restrictions. For example, the sessions had to fit within the drivers' shift schedules. There were also time restrictions with the participants which meant that the study (including briefing, questionnaires, interview questions, the simulator drives and debrief) had to be designed to fit within a two-hour timeframe, and questions and drives had to be adapted to fit that. Due to these restrictions, it also meant that sleepiness and fatigue were unable to be experimentally manipulated.

The simulator drives were planned as a test of the *i*-DREAMS system, but there were several limitations associated with this part of the study as well. It has been argued that simulator studies cannot accurately replicate real life, which therefore is likely to affect the driving of the participants. For example, simulated environments contain none of the risks the drivers would face in real tram driving. This was also true for workload. Drivers reported that the workload differed between the suburban and urban sections of the route, which was difficult to replicate within the simulator. The planned addition of the counting task was an attempt to increase workload during the urban sections, however, this had to be adapted to ensure the drivers could complete it alongside driving the tram. The events included within the simulator drives were also pre-programmed and could not be altered by the experimenters; this necessarily influenced the design of the study. Several participants also struggled with negotiating the pre-programmed events, which resulted in their simulator drives ending early.

Finally, as the *i*-DREAMS system was designed to be used in road vehicles, adjustments had to be made to test it within the tram environment, which did not always work, for example, the false VRU alerts.

A3.7 References

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology, Qualitative Research in Psychology, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Hayman M (1942) Two minute clinical test for measurement of intellectual impairment in psychiatric disorders. Arch Neurol Psychiatry 47:454–464

Saldana, J. (2009) An introduction to codes and coding. The Coding Manual for Qualitative Researchers. London, UK: Sage.