D7.2 Effectiveness evaluation of the interventions.

Interview with Laurie Brown

Deliverable 7.2 focuses on evaluating the effectiveness of the i-DREAMS interventions in improving drivers' safety outcomes. Field trials were carried out in five countries (Belgium, Germany, Greece, Portugal and the UK), across four transport modes (cars, trucks, buses and rail). Trial protocols were broadly the same for cars, trucks and buses, but different methods were used for rail. In this deliverable, comparisons are made between the different countries for which data are available. At the time of this interview, no final results were available yet regarding the bus trials.

Again, we talk with Laurie Brown who was the responsible author of this deliverable. In this interview the focus is on the effectiveness results per mode, taking into account differences that were noted between the different countries. But before focussing on those results, we will first remind all the readers about several aspects of the theoretical framework, such as the fact that:

- the i-DREAMS platform combines both real-time and post-trip interventions to respectively nudge and coach the drivers;
- these interventions aim to improve the outcomes proposed in the Logic Model of Change (LMoC);
- the intervention approach consists of four phases.

After clarifying these aspects, the findings with respect to the effectiveness of interventions, distinguishing the results of both the outcome as well as the process evaluation, are reported per mode for which data were available.

Hello Laurie, we meet again! Before going into the practical results, I first want to remind our readers about some theoretical aspects. To start, can you briefly explain to us again how we define real-time and post-trip interventions in i-DREAMS?

LAURIE: *"Of course! Real-time interventions are warnings that are provided to drivers during the trip via the in-vehicle i-DREAMS display to warn them for imminent danger. As you know, in the Safety Tolerance Zone (STZ) we distinguish three phases (see Figure 1): the 'normal phase' where there is minimal risk, the 'danger phase' where there is an increased collision risk and the 'avoidable accident phase' where immediate driver intervention is required to prevent a crash. A real-time warning is triggered if the driver enters the 'danger phase' and more severe (intrusive) warnings are triggered if they move into the 'avoidable accident phase'. Real-time warnings are triggered for the performance objectives (PO's) related to 'speeding', 'road sharing' and 'driver fitness'. No real-time warnings are provided for the PO 'vehicle control'. Post-trip interventions are provided in the form of feedback to the drivers via the i-DREAMS smartphone app after the trip. That feedback is provided in the form of 'scores' or the visualisation of the 'trips' and the events on a trip, or in the form of 'info & tips' about driving behaviours. Gamification is also used in the final phase, with a 'leader board' that provides a driver ranking of performances and 'goals and badges' which challenge and reward drivers to further improve driving behaviour."*

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

Figure 1: The three phases in the STZ

You referred to 'performance objectives (PO's)'. If I recall correctly, they are part of the Logic Model of Change (LMoC). Can you elaborate a bit on that too?

LAURIE: *"The LMoC (see Figure 2) presents four different levels of driver safety that are targeted by our i-DREAMS interventions. The highest level is the Safety Outcomes (SO's), 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 (SPG's). These are the behaviours that need to change for the safety outcomes to be realised. The second-lowest level is the Performance Objectives (PO's). 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 (CO's). These are the underlying behavioural determinants that need to change for the performance objectives to become realisable.*

Figure 2: Illustration of how the Logic Model of Change is applied in i-DREAMS

During a trip, events are detected at the level of the PO's[1]. When we discuss the results per mode later on, we will focus on the number of events that occurred per 100 km, so then we are talking about the PO-level in the LMoC."

[1] In i-DREAMS the following PO's are distinguished: acceleration, deceleration, steering (*under SPG 'vehicle control'*) - speeding (*under SPG 'speed management*') - tailgating, lane departure, forward collision avoidance, vulnerable road user collision avoidance, illegal overtaking (*under SPG 'sharing the road with others'*) – fatigue, distraction (*under SPG 'driver fitness'*).

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

And then lastly, before actually diving into the results. Can you again explain to us the four phases that were distinguished in the on-road field trials?

LAURIE: *"We distinguished four phases (see Figure 3) indeed to test the interventions at different points of the trial. Phase 1 was the baseline phase, Phase 2 introduced real-time warnings in the vehicle, which were present for the remainder of the trial, Phase 3 introduced post-trip feedback via the i-DREAMS app, and Phase 4 further added gamification functions to the app."*

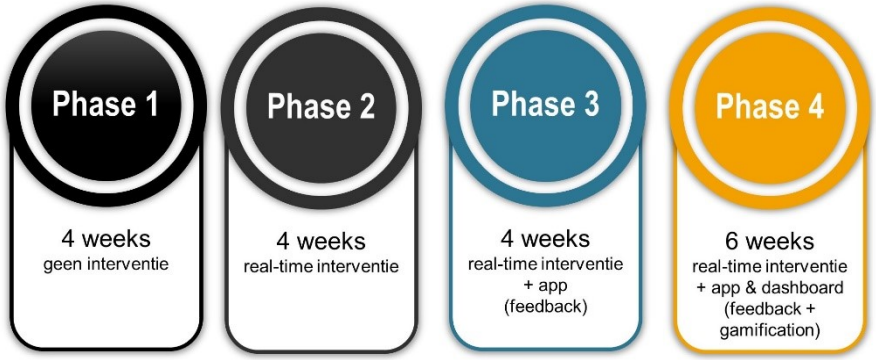

Figure 3: Four phases of the field trials

Ok, that brings us to the actual results then. Let us first focus on the car results. Regarding the outcome evaluation, what were the main conclusions there?

LAURIE: *"Overall, car drivers showed a reduction in events per 100km after exposure to the i-DREAMS technology. So, there was an improved safety outcome. However, differences were found between the countries analysed, and between the different safety promoting goals (SPGs). It was not possible to form robust conclusions regarding 'fatigue' and 'distraction' events, due to a lack of data.*

Drivers in the UK were found to have the highest number of events per 100km for most SPGs ('total', 'vehicle control', 'road sharing'), but also showed the greatest reduction in events, and consistent reduction across the data collection phases. Statistically significant[2] decreases in events were found overall (i.e., Phase 1 to Phase 4) for all event types, with additional significant results most commonly found in Phase 1 to Phase 2, with the introduction of the real-time warnings.

Drivers in Germany had the highest number of 'speeding' events per 100km, but the lowest number of 'vehicle control' events. These drivers were also on average younger and less experienced than the drivers in other countries. Statistically significant decreases in events were found overall for 'speeding', and specifically from Phase 2 to Phase 3 for 'total' events, when the post-trip feedback was introduced.

[2] If a result is statistically significant, it means that it is unlikely to be explained by chance or random factors alone. In other words, there is only a very small chance of a statistically significant result occurring if there were no real effect in the study.

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

Drivers in Belgium had the lowest number of events for 'speeding' and 'road sharing', and substantially fewer 'road sharing' events compared with UK drivers. However, the overall reduction in events was small, and events initially increased. In the questionnaire data, Belgian drivers reported they engaged in far more risk-taking behaviours than drivers from Germany and the UK, however the data showed the opposite. Statistically significant decreases in events were found overall for 'road sharing' events, and specifically from Phase 3 to Phase 4 for 'total' and 'vehicle control' events, when the gamification features were added and the complete intervention scheme was present.

For the individual SPGs, the most improvement was seen in 'road sharing' events, with both Belgium and the UK showing a reduction in both 'medium' (STZ level 2) and 'high' (STZ level 3) severity events, which was consistent across all phases. Results were mixed for 'vehicle control' events, with Belgium and Germany showing a small increase, and the UK showing only a small decrease. 'Speeding' events also showed mixed results, increasing slightly in Belgium, but decreasing in Germany and the UK.

Analysis of questionnaire data showed that drivers subjective opinions were that their driving performance improved through exposure to the technology. In particular, the increase in scores for 'perceived knowledge' (knowing the benefits of safe driving, knowing what is needed to drive safely) were statistically significant for all countries."

Very interesting! Were you able to draw some general conclusions for cars about the process evaluation as well?

LAURIE: *"Yes, we were. The use of the i-DREAMS app varied between countries, with German drivers having particularly low engagement. For Belgium and the UK, drivers engaged more with the app in Phase 4 than Phase 3, and in both phases, usage was highest at the start of the phase, then gradually decreased throughout. Data suggests that the gamification functions (Phase 4) were more engaging and held attention more consistently. One suggestion is that the gamification functions prompted more regular visits, as for example their position on the leader board and their progress towards goals would change daily. Whereas records of their trips and events, and associated score (Phase 3 functions), were clearly of interest to participants, but may be something drivers look at less frequently as they could review multiple trips at once, or may only be reviewed if the driver felt a recent trip was particularly eventful. Push notifications in the app were an effective method of increasing engagement.*

Generic information (e.g. tips & facts) in the app was less appealing to users. Drivers found more interest in personalised feedback such as their trip information, goals progress, and position on the leader board.

With respect to user acceptance, the vast majority of participants felt the i-DREAMS system was easy to use and easy to understand. In Belgium and the UK, drivers generally felt that the system improved their performance and was a good idea, however a lower proportion said they would continue to use it. The reason for this could be that the data also showed lower trust, and a feeling that the alerts did not always accurately reflect the situation (i.e., false alerts). Furthermore, nearly half the UK drivers felt the system was 'annoying' (which could again be related to false alerts), though very few Belgian drivers said the same. Drivers in Germany were generally less accepting overall."

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

Where there any eye-catching differences between drivers within each country?

LAURIE: *"Yes, there were actually, but they are not always easy to explain. For Belgian drivers, the overall effect was a small improvement in performance, with events initially increasing before decreasing in the final phase. However, on closer examination two thirds of Belgian drivers showed a consistent improvement across all phases for 'total' events. The remaining third showed an increase in events in early phases, and almost no change in Phase 3 to Phase 4, but these are in the minority. It is noted however that fewer drivers showed improvement when 'speeding' events are looked at in isolation. There were little demographic differences between the two groups of drivers to explain these results, however some of the drivers which did not improve were more impacted by changing COVID-19 restrictions and technology issues, which could partly explain their outcomes.*

For German drivers, again approximately two thirds of drivers showed improvement, which was consistent across all phases, while the remaining third showed an increase in events. Of the drivers who did not improve, a higher proportion were male when compared to the drivers who did improve.

For UK drivers, three quarters of drivers improved, showing a consistent decrease in events between phases, and only a quarter did not improve. Furthermore, the average decrease in events for drivers who improved was substantially greater than the average increase in events for drivers who did not. The drivers who did not improve had a higher proportion of male drivers and were noticeably younger and less experienced than the drivers who did improve.

In all countries the drivers who did not improve more often described themselves as 'confident' or 'very confident' drivers compared with those who did improve.

For Belgium and the UK, the data suggests a link between app usage and performance outcome; nearly all of the drivers who had high app usage showed improved outcome, and drivers who did not improve typically had lower app usage. It would be interesting to investigate this further to determine whether there is a direct correlation between these results, or if the low app usage was symptomatic of an overall lower engagement with the trial and less willingness to change behaviour."

And what were the main findings related to the truck trials?

LAURIE: *"Truck data were analysed for five Belgian truck companies that participated. The results were less clear than for car drivers, however we were still able to draw some conclusions.*

With respect to the outcome evaluation, the overall event numbers were substantially lower for truck drivers than car drivers. However, interventions appear to be less effective, with generally little change between phases for most event types. Most improvement was seen during Phase 4 of the trial (where all interventions are present), though changes were not statistically significant. The only statistically significant decrease in events was seen for 'total high' events between Phase 1 to Phase 2. This could suggest that the real-time warnings had significant impact on the most severe events.

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

With respect to the process evaluation, phase 3 interventions appeared to be the least effective, though it is also noted that only two thirds of drivers engaged with the app. Regarding the app, more use was seen in Phase 4 compared to Phase 3. This suggests that the gamification features held drivers' attention more, and indeed the leader board was the most commonly area visited.

Further analysis showed differences between drivers within the sample. When only drivers whose outcome improved (i.e., events decreased) are considered, they showed a consistent decrease across the data collection phases. However, a slightly larger proportion of drivers had an overall increase in event numbers. As with car drivers, the data suggests a link between app usage and performance outcome; nearly all of the drivers who had high app usage showed improved outcome."

That brings us to the rail mode. There, a totally different approach was used, if I'm not mistaken?

LAURIE: *"Yes, and for good reason. Trains (heavy rail) and trams (light rail) were included to broaden the application of the i-DREAMS platform which was originally designed for use in road vehicles. It 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[3].*

In contrast trams operate within a mixed-traffic environment, both driving on 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."

How did you approach the simulator study?

LAURIE: *"The drivers completed a series of four drives: a familiarisation, a baseline, an intervention and a manipulation drive. Furthermore, they completed a series of questionnaires and discussions about their drives, the equipment, and their experiences. We recruited 30 participants who all had at least six months of tram driving experience. At the end, two of them dropped out due to simulator sickness. During the simulator study we tested parts of the i-DREAMS platform options, namely alerts for speeding, VRU[4] detection and fatigue."*

Were you able to draw any conclusions from the simulator study?

LAURIE: *"Yes, we were. Due to the mixed traffic, multi-user environment that tram drivers operate in, a system to help improve*

[3] Information can be found in the WP8 deliverables.

[4] VRU = Vulnerable Road User

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

safety and mitigate risk has potential to be useful. From ours tram simulator study we learned 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 stations, 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 existing system[5] 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 alarms are being triggered.

And what about the focus group study. How did you approach that?

LAURIE: *"The 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. Six focus groups were conducted online with drivers and trainers. They were contributing to the focus groups as part of their scheduled working day. Each focus group involved between three and five participants and lasted 50 to 60 minutes. The focus groups found that drivers had mixed opinions about the potential introduction of the app, with both positive and negative views being discussed."*

Could you be a bit more precise on what those opinions where?

LAURIE: *"Sure! 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 ok 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."*

[5] The primary purpose of the fatigue system, currently already in use in the UK, is to detect fatigue events by providing a warning when the system thinks eyes are closed (or not visible). The system appears to be less well accepted. 'False positives' might play a role in this.

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

OK, may I conclude that the i-DREAMS system cannot be directly translated into trams in its current format and that some adaptation is required?

Laurie: *"That is absolutely a correct conclusion. In the deliverable we actually made some suggestions on methods and issues which can guide this adaptation. We made some general suggestions, but also a couple of more specific ones related to fatigue warnings, speed warnings, VRU event warnings and post-trip feedback from the app."*

Thank you, Laurie. With all of these insights, I think we can conclude this interview.

Edith Donders

i-DREAMS DisCom manager

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

Deliverable 7.2 is part of WP7:

Evaluation of safety interventions

Download the report here

i-DREAMER in the spotlight

LAURIE BROWN

Graduated as *mathematician* in *2012*

Employed at *Loughborough University* since *2012*

Passionate about *reading, everything Disney, and my cats.*

Tasks in i-DREAMS: *Co-ordination of the car field trial at Loughborough. Lead on Deliverable 7.2 analysing the effectiveness of the real-time and post-trip interventions, and responsible for analysis of UK data for this WP.*