UDRIVE
European Naturalistic Driving Study

# Driver distraction and inattention

Deliverable 43.1

DOI: 10.26323/UDRIVE_D43.1

UDRIVE

European Naturalistic Driving Study

# EUROPEAN COMMISSION

SEVENTH FRAMEWORK PROGRAMME
FP7-SST-2012.4.1-3
GA No. 314050

## eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment

| | | |
|---|---|---|
| **Deliverable No.** | UDRIVE D43.1 | |
| **Deliverable Title** | Driver Distraction and Inattention | |
| **Dissemination level** | Public | |
| **Written By** | Oliver Carsten (University of Leeds)<br>Daryl Hibberd (University of Leeds)<br>Jonas Bärgman (Chalmers/SAFER)<br>Jordanka Kovaceva (Chalmers/SAFER)<br>Marta Sofia Pereira Cocron (Technische Universität Chemnitz)<br>Mandy Dotzauer (DLR)<br>Fabian Utesch (DLR)<br>Meng Zhang (DLR)<br>Eric Stemmler (DLR)<br>Laurette Guyonvarch (LAB)<br>Fridulv Sagberg (TÖI/SAFER)<br>Fabio Forcolin (Volvo) | 27-06-2017 |
| **Checked by** | Jonas Bärgman | 29-06-2017 |
| **Approved by** | Marika Hoedemaeker (TNO) *QA manager*<br>Nicole van Nes (SWOV) *Project Coördinator* | 30-06-2017 |
| **Status** | Final | 30-06-2017 |

**Please refer to this document as:**

Carsten, O., Hibberd, D., Bärgman, J., Kovaceva, J., Pereira Cocron, M.S., Dotzauer, M.,... Forcolin, F. (2017). *Driver Distraction and Inattention.* UDRIVE Deliverable 43.1. EU FP7 Project UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D43.1

**Acknowledgement**:

The author(s) would like to thank Tsippy Lotan for her valuable comments on previous drafts and Marika Hoedemaeker for performing the quality assurance on the final draft.

**Disclaimer**:

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 314050.

# Executive Summary

## Introduction

In UDRIVE, the major focus of the work on driver inattention and distraction has been focused on obtaining a better understanding of whether and how drivers manage their secondary task activities — when they choose to engage, what tasks they select, whether they adjust their activity to different situations and whether they are willing to surrender secondary task activities as the primary task of driving becomes more demanding. In other words, the focus is on self-regulation, on how drivers manage their secondary task activity in the context of the dynamics of the traffic and road situation. That management includes the determination not to engage in such tasks in the first place or only to engage in some particular activities. NDS are particularly suited to such an investigation, since experimental studies in driving simulators and even on test tracks tend to suffer from an instruction effect, in that participants are typically instructed to carry out an activity at a given moment. Thus such experimental studies provide insight into how driver attention, driver information processing and driving performance are affected by secondary tasks, but are less useful when research is focused on driver management of task activity.

## Car driver engagement in secondary tasks

The major question here was in what activities are drivers engaging while driving and are there differences by driver category and driver country.

Overall the car drivers spent 10.2% of their driving time engaged in some kind of secondary tasks. Mostly, that was one single task at a time, but there were also rare instances of being engaged in two tasks simultaneously. The most prominent activity for car drivers was mobile phone use, followed by smoking and talking/singing (

Males were somewhat more likely than females to engage in secondary tasks (10.8% compared to 9.5% of driving time). There were also some differences between men and women in their activity patterns, with men doing a lot more smoking and more talking to passengers or singing, and women doing more mobile phone activity and personal grooming.

Within the mobile phone activities, the most time (38.4%) was spent on hand-free interaction, which would typically be touching the screen while the device was in a cradle, for example to interact with a navigation app. Next were hand-free conversation (22.6% )and handheld interaction, such as texting (22.0%). The last is most concerning, as it involves visual-manual interaction with the device, which is known to be highly risky because drivers tend to take their eyes off the road scene for a substantial time.

Again, there were differences by gender. Males were far more prone to spend time interacting with hands-free phones, while females spent more of their time in hands-free conversation and in handheld interaction, such as texting. Men spent a higher proportion of time in handheld conversation.

By country, participants from Poland spent a substantially higher proportion of their driving time in secondary task activities — almost 20%. Next were the participants from France and the UK, with the German drivers having the lowest engagement at 1.9%.

The rank order for mobile phone use by country was similar. Again Poland was the highest at nearly 10%, followed by France and the UK at less than half of that. The proportion of time on the mobile phone was remarkably low for the German drivers at 0.1%.

## Truck driver engagement in secondary tasks

The major question here is in what activities are drivers engaging while driving and to what extent were truck driver activities different from those of the car drivers.

The global number of hours of driving analysed was 83. The total time spent in all the secondary tasks sums up to about 20% of the total driving time that has been annotated. Therefore, as an example, it is possible to deduce that the time spent in phone-related activities is about 5% of the total annotated driving time.

In terms of total task time spent on a task, smoking was by far the most prominent, followed by food, phone and talking/singing. The last two were approximately equal in total task time. Inspection of the data revealed that there was one very long-lasting occurrence of a smoking task (30 minutes in total) which led to the large difference in duration of the smoking task between the two types of average that are shown in the figure. Also note that only one-third (eight) of the drivers smoked at all while driving. Thus the total amount of smoking was low, even if the average duration was long. Nevertheless, it can be seen that smoking is more prominent than for the car drivers. Also to be noted is electronic device interaction, which does not feature at all for the car drivers. Such devices are part of the working environment in trucks, especially here where the trucks were being operated for delivery.

When phone use was broken down into s sub-categories, the analysis revealed that the most frequent task was handheld interaction which accounted for about 2% of the total driving time annotated and for about 35% of the phone sub-tasks. This is potentially concerning, as it involves visual-manual interaction. Note that hands-free talking was hard to distinguish from the talking/singing category in the manual annotation. However, the annotators separated these categories to the best of their ability, i.e. coded the task as hands-free when they could observe reliably that the driver was making a phone call.

Thus, not surprisingly, there are quite large differences between the activities of the car drivers and the truck drivers. There are clearly instances of more risky activities involving visual-manual interactions.

### Car driver attitudes and engagement in secondary tasks

Knowing what are the motivators for inattention and distraction can help in developing targeted countermeasures aimed at particular segments of the population. The issue addressed here was to what extent do personality factors play a role.

A set of questionnaires probing driver attitudes and self-reported behaviours was collected from the participants on recruitment. From those questionnaires an overall score of driver propensity to negative attitudes on safety combined with more risky self-reported behaviours was calculated. Twelve factors went into this scale, with drivers being coded 0/1/2 on each factors, so that the maximum score (most risk-prone) for each individual was 24. The drivers were divided into those that did not and those that did engage in any particular task. There were some large differences in the average composite personality score between those that did not and those that did engage in particular tasks. Mobile phone use, smoking, eating and drinking, grooming and dual-task engagement all stand out.

Personality and attitude have been shown to have a clear influence on propensity to engage in secondary tasks. Future work could investigate more refined personality scoring and more detailed aspect of engagement, such as willingness to engage in specific tasks in specific situations.

### Did driving task complexity and secondary task complexity influence the decision to engage in secondary tasks?

The focus here was on whether, when drivers do decide to engage in secondary tasks, they consider their current driving activity, the external driving situation and the difficulty of the secondary task in deciding whether or not to carry out the activity at a particular moment. In other words, the issue was whether the drivers self-regulate. This analysis was only performed for car drivers.

There were signs of adaptation of engagement. A coding scheme was applied to classify task complexity and manoeuvre complexity. Duration of secondary task was affected by complexity of manoeuvre, such that task duration was longer in easy conditions and shorter in medium and hard conditions.

However, drivers did not adapt as well to overall conditions (not just manoeuvre) and to secondary task difficulty. Easier tasks were indeed dropped as the road environment shifted from easy to hard, but more demanding secondary tasks tended to be more frequent when conditions were hard.

There are thus some signs of adaptation by drivers, in particular when performing a manoeuvre. However, appropriate adaptation was not always pursued — drivers were seemingly more willing to perform medium and difficult tasks when the situation was more demanding. This suggests perhaps that they do not fully comprehend the demands imposed by harder tasks and are unwilling to suppress them as the road situation becomes more difficult. However, these results need to be confirmed by more extensive analysis of the UDRIVE database.

### Do drivers adapt their safety margins for performing secondary tasks?

#### Car drivers

By analysing driving in the period before, during and after the performance of a secondary task, it is possible to determine to what extent drivers increase their safety margins, in the knowledge that their reactions might be slower while performing a secondary task, so that they need to compensate for this, perhaps by slowing down on increasing their distance to a lead vehicle.

The analysis for the car drivers focused on the task that is most demanding in terms of attentional focus, visual-manual interaction such as texting. In a majority (56 percent) of cases drivers carried out this type of interaction while stationary, although a substantial minority of interactions took place while moving.

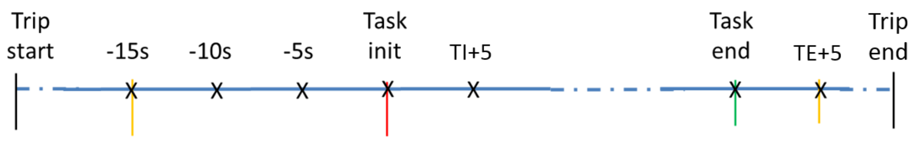
In line with this, it could be seen that speeds tended to be lower in the task period than in the preceding baseline period, but this was mainly due to the propensity to carry out visual-manual tasks while stopped. There are some interactions that take place at high speeds. Overall, mean speed in the task period was around 20 km/h.

Similarly, while most drivers (73 percent) slowed down to perform their visual-manual tasks, others (27 percent) were faster in the task period than the baseline period. For those that slowed down, the mean change was −13 km/h; for those that sped up, the mean change was +7 km/h.

#### Truck drivers

The speeds at phone task initiation were compared to the speed distribution across all annotated trips. This can be seen as a representation of when drivers feel more comfortable engaging in secondary tasks, compared to everyday driving (not performing a task). The majority of the phone tasks were initiated at low speeds (at or below 30 km/h) or when standing still. The speed range with the highest frequency of phone task initiations (but also most time spend driving) is 80 km/h and above. This may be when the environment is less complex and the driver even may have support systems, such as cruise control, engaged. Drivers in this study initiated phone-tasks less often at standstill and at 70 km/h, and more often at 5 and 20 km/h, compared to everyday (no-phone task) driving. Note the large difference in the proportion of phone task initiations at standstill for cars and trucks. Cars drivers were at standstill in 56% of visual-manual tasks, which were the majority of all truck tasks.

A more detailed analysis examined seven distinct points in time before during and after the start of a secondary task. Here after these points are called analysis points (APs). The APs are shown in the figure below.




**A visualization of the seven analysis points used in this analysis, as well as the start and end of the trip**

*Note: The dotted sections of the time-line represents a variable length, while the solid lines represents 5s equal separation between points.*

When all tasks were considered together, results showed a significant speed decrease between the first two time-points: −15 s and −10 s before task start (−.21 and −1.07 km/h, respectively) and task start. This may be due to drivers self-regulating, but further analysis showed that stopped vehicles affected the results. When the secondary tasks (of “all task”) where at least one analysis point had a speed of zero were removed — no decrease in speed between before the task and task start, but instead an increase in speed from the period between 15 s before the task start until task end, and 5 s after the task end. That is, drivers did not significantly decrease the speed before or during the task, but instead increased their speed after it was completed.

For phone tasks, the picture was similar: there was a significant increase in speed between a) before and up to five seconds into the task, and b) the end of or after the task. That is, drivers seem to have self-regulated by increasing their speed after the phone task ended, rather than reducing it before the task was initiated. When the tasks performed while stopped were removed, the trends were the same as with the standstills included. One way of interpreting the results of speed differences over the seven points between with and without standstill, and in relation to the low proportion of task initiations at stand-still, is that drivers who stop and interact with their phones slow down, but do not stop, until after they have actually initiated the phone interactions.

### Conclusions

There are indications of some self-regulation by drivers, but there are also instances of activity at high speeds. Further analysis is needed to get a better understanding of the circumstances of such high-speed activity, both for car and truck drivers. Equally, the situations in which car drivers were speeding up also need to be analysed in more detail.

## Methodology: potential of automated video analysis to support or replace manual annotation

Currently, manual annotation is needed to identify secondary task activity, and searching for secondary tasks in naturalistic data is a real challenge due to the low frequency of these event in normal driving. An automated procedure has been applied to provide candidate cases of secondary tasks to manual annotators.

The automatic annotation tool is based on deep learning algorithms. The first step consisted in a frame by frame detection of secondary task using a Convolution Neural Network (CNN) model. LAB used French UDRIVE pilot data to develop the tool. The development included manual annotations to create a learning data base used to train the CNN model. The tool output was the probability of the driver to be texting/phoning for each frame. A threshold value was then used to turn the probability into a prediction (“secondary task ongoing”/ “no secondary task ongoing”). This threshold value was derived from the Receiver Operating Characteristic (ROC) curve of the CNN model where an optimal point was defined giving True Positive Rate and False Positive Rate the same weight. Using these thresholds, the results for hand held phone call detection on the test (pilot) data was a True Positive Rate of 0.94 and a False Positive Rate of 0.04. For texting, True Positive Rate is 0.80 and a False positive Rate is 0.15. The tool was implemented on the UDRIVE database and a first set of trips was processed. The automatic annotation was included in SALSA and can be compared to manual annotations in future studies.

The results in the pilot database were promising. The first potential enhancement is to add hysteresis to the frame by frame detection. This evolution would benefit from the temporal nature of the data and potentially increase the performance compared to the frame by frame detection. The next step of our approach will then be to evaluate the models on UDRIVE data and tune it with data from other Operation Sites data to increase performance. Manual annotation from UDRIVE data set will be used in the next months as a ground truth to adapt the thresholds and tune the algorithms. Performance will be assessed using the manual annotations, and the automatically realised detections will be used for further studies on driver behaviour when engaged in secondary tasks.

## Table of contents





UDRIVE
European Naturalistic Driving Study

# 1 Introduction

This report constitutes the major output of the work on driver distraction and inattention within UDRIVE, which has been carried out within WP43, Distraction and Inattention. Since the introduction and widespread adoption of the mobile phone, the topic of driver distraction has received huge attention (see e.g. the review by Kircher et al., 2011). Past research consistently shows that driving performance decreases when drivers are engaged in secondary tasks (e.g. Brookhuis et al., 1991; Engström et al., 2005; Sayer et al., 2005; Horberry et al., 2006; Jamson and Merat, 2005; Consiglio et al., 2003; Hancock et al., 2003).

With the advent of Naturalistic Driving Studies (NDS, i.e. observations of driving in the real world with data collected from the vehicle and with no experimental intervention), the impact of distracted driving on safety, became a major, if not the major, topic addressed. NDS were applied to calculate the probability of a safety critical event or crash occurring while engaged in a secondary task, as compared to the probability of such an event in randomly selected or matched baseline episodes (i.e. time slots without the secondary task. The overall consensus has bene that visual-manual and visual tasks have a direct impact on driver attention to the roadway and traffic scene and have negative impacts on safety. There is much less consensus on the impact of cognitive tasks such as talking on a hands-free mobile phone, with laboratory and simulator studies tending to show a negative impact on safety and many real-world studies suggesting a positive (protective) effect, at least for forward collisions (Carsten and Merat, 2015).

Here, by contrast, the focus is not on the estimation of risk while performing secondary tasks, but rather on how drivers manage their secondary task activities — when they choose to engage, what tasks they select, whether they adjust their activity to different situations and whether they are willing to surrender secondary task activities as the primary task of driving becomes more demanding. In other words, the focus is on self-regulation, on how drivers manage their secondary task activity in the context of the dynamics of the traffic and road situation. That management includes the determination not to engage in such tasks in the first place or only to engage in some particular activities. NDS are particularly suited to such an investigation, since experimental studies in driving simulators and even on test tracks tend to suffer from an instruction effect, in that participants are typically instructed to carry out an activity at a given moment. Thus such experimental studies provide insight into how driver attention, driver information processing and driving performance are affected by secondary tasks, but are less useful when research is focused on driver management of task activity.

Previous NDS, principally carried out in the U.S. have reported on the prevalence of secondary task activity and on inattention more generally (e.g. Hickman et al., 2010; Dingus, 2014; Dingus et al., 2016), but they have not tended to examine in-depth which drivers engage, in what traffic situations they do it, what tasks and subtasks they choose to engage in, and whether they appear to be managing their safety margins. However, these have sometimes been conducted as on-road experiments that take the form of probing drivers on their willingness to do as task as opposed to observing them in a task (e.g. Lerner and Boyd, 2005), or have focused on drivers of high-end vehicles (e.g. Metz et al., 2014). The UDRIVE research and analysis affords an opportunity to examine drivers management of secondary tasks across a varied sample of drivers from a number of European countries, as well as across private and fleet driving.

The report covers driving of both cars and trucks. For car driving, a special feature is the ability to compare behaviour in different countries — France, Germany, Poland and the UK. Since these countries have substantial differences in overall road safety performance, the results may shed some light on whether driver distracting activities contribute to those differences. However, it should be noted that the participants do not constitute a representative sample of the driving population in each of these countries, so that cross-country comparisons can only be considered as indicative, rather than as conclusive. Car and truck analysis has been carried out separately, on account of the very different operational environments of the two vehicle types. However, the conclusions attempt to reflect on the overall findings and their implications.

# 2 Data handling and processing

## 2.1 Choices in sampling

The objective in the sampling of the overall UDRIVE data for this analysis was to represent as far as possible to full range of participants and their trips. Separate samples were drawn for car and truck driving. It was decided that the first stage in sampling should be participants, in order to maintain as much variety as possible and to give as many participants as possible a chance to be represented in the analysis. Within participants, trips were sampled to ensure that short trips as well as longer trips would be included in the analysis.

However, the data used is not a true statistical sample of the whole database. That is because data was processed as it became available. This was on account of the short time available between the end of data collection and the beginning of formal analysis. Data from some Operation Sites and some drivers was available sooner, and therefore those drivers are, in strict terms, over-represented in the database. There has been no compensation for this by, for example, any formal weighting procedure in the analysis. On the other hand, there is no reason to believe that the analysis has been distorted on account of using a sample that can to some extent be termed a “convenience sample.”

## 2.2 Car sample

### 2.2.1 Trips

A pool of 745 annotated trips were used to analyse the prevalence of secondary task activity in normal driving. These trips represented a total of 194 driving hours (shortest trip: 36 sec; longest trip: ≈3 hours) and 9148 km driven (shortest travelled distance: 7.32 m; longest travelled distance: 273.48 km). The videos were collected in different operation sites located in France, Germany, Poland and United Kingdom. A distribution of the annotated trips per country is showed in Table 2-1.

**Table 2-1: Distribution of the annotated trips by country**

| | Total number of trips | Total travelled time (hours) | Total travelled distance (km) |
|---|---|---|---|
| France | 217 | 58.57 | 2862.27 |
| Germany | 83 | 20.21 | 1249.75 |
| Poland | 150 | 41.40 | 1668.69 |
| United Kingdom | 295 | 73.83 | 3367.74 |

As can be seen in Figure 2-1, the majority of the trips (n=660) lasted between 1 and 30 minutes. Only two trips were shorter than one minute and 21 lasted longer than one hour. When considering the distribution of trips by distance travelled (Figure 2-2), it is possible to observe that 423 trips lasted from 1 to 10 Km and 213 tips from 10 to 50 Km. Few were shorter than 100 metres (n=2) or longer than 50 Km (n=27).

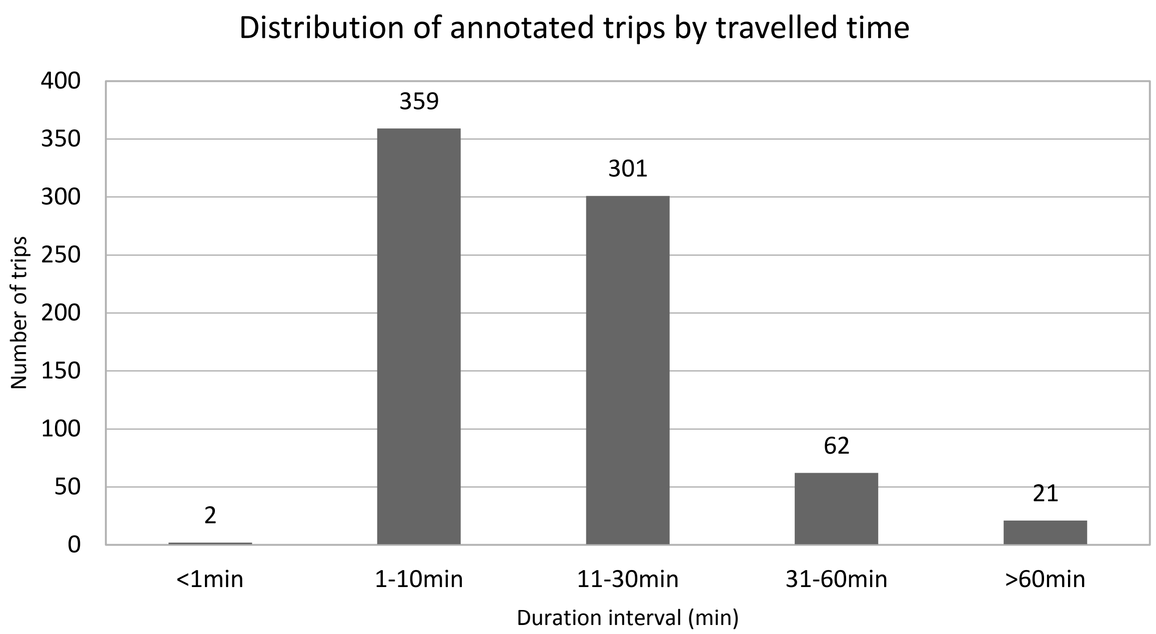



**Figure 2-1: Distribution of annotated trips by travelled time**

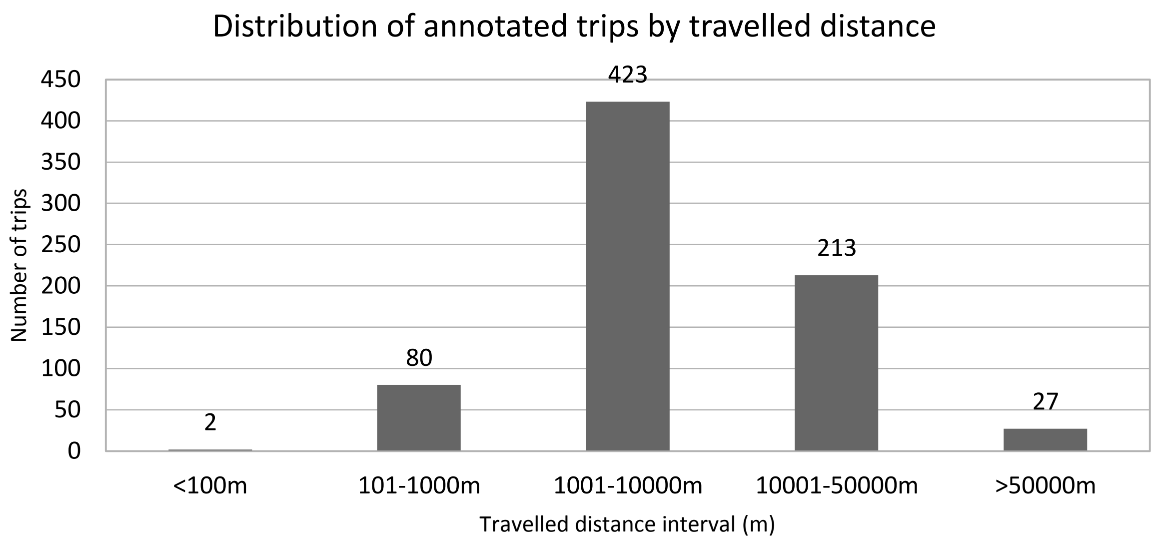



**Figure 2-2: Distribution of annotated trips by travelled distance**

### 2.2.2 Participants

The number of drivers in the analysis was 96 (49 male, 47 female). Their age ranged between 20 and 77 years (*M* = 44.48; SD = 13.05). Table 2-2 shows how participants were distributed among countries and gender. Table 2-3 looks at travel by gender group. It is possible to verify that women performed 31 more trips than men. However, men accounted for a slightly higher number of travelled hours (about 7 hours more), and more distance travelled (1482 more kilometres driven than females).

**Table 2-2: Distribution of participants by gender and country with respective trips statistics**

| | Male | Female | All |
|---|---|---|---|
| France | 14 | 12 | 26 |
| Germany | 8 | 4 | 12 |
| Poland | 12 | 8 | 20 |
| UK | 15 | 23 | 38 |

**Table 2-3: Trip statistics**

| | Male | Female | All |
|---|---|---|---|
| Total number of trips | 357 | 388 | 745 |
| Total travelled time (hours) | 100.47 | 93.53 | 194 |
| Total travelled distance (km) | 5315.33 | 3833.12 | 9148.45 |

## 2.3 Truck sample

The sampling of data for analysis of secondary tasks for trucks was based on a discussion about benefits and drawbacks with different sampling approaches. Part of this discussion was joint with UDRIVE WP4.2 where sampling for random and matched baselines for analysis of Safety Critical Events (SCEs) was discussed and decided (See D42.1). Below a description of the chosen sampling strategy is provided.

The criteria used for sampling segments of driving for the secondary task analysis for trucks were as follows:

- To be considered for inclusion, a driver should have at least 100 trips available for use in analysis in the database.
- For each driver a minimum of six trips were sampled randomly across that driver's entire dataset (when he had a minimum of 100 trips in the database).
- More trips were added for a driver if the total time driven did not exceed 2 hours. That is, the number of trips for each driver was at least six, but with additional trips added until a minimum of two hours were available for analysis for each driver.
- To avoid different sampling probabilities for different parts of the data within a driver – as additional data became available for a driver the sampling was made only on one dataset for each driver. That is, as soon as a driver reached above 100 trips, the entire set of trips (>100) was used as the basis for the sampling described in the bullets above.
- For a trip/record to be considered for inclusion in the sampling it had to be at least 30 s long and have a speed greater than zero for at least 50% of the samples.
- For annotated trips there were some initial and end conditions that had to be fulfilled. If these were not fulfilled the start and/or end of the trip was cropped until it was fulfilled. That is, a trip was annotated from 10 seconds priori the vehicle started moving for the first time in the video (or the start of the video itself if it was already moving) to 10 seconds after the

vehicle finally stopped moving (or the end of the video if the driver was still driving at the end of it).

Note that data were collected continuously and new trips for annotation and subsequent analysis were added as data came in to the central data centre (CDC). However, due to the limited time available for annotation, if a driver was included in the dataset for analysis of secondary tasks for trucks (i.e., fulfilling the criteria above), additional sampling was not done for that driver. That is, within drivers the probability of sampling a trip was equal across time/distance (as it was only done once per driver). Between drivers, however, the probability of a time trip being sampled was different, as some drivers had several hundred trips and others just barely had passed the 100 trip threshold. Figure 2-3 shows the relationship between number of annotated trips and the total annotated time for each driver.

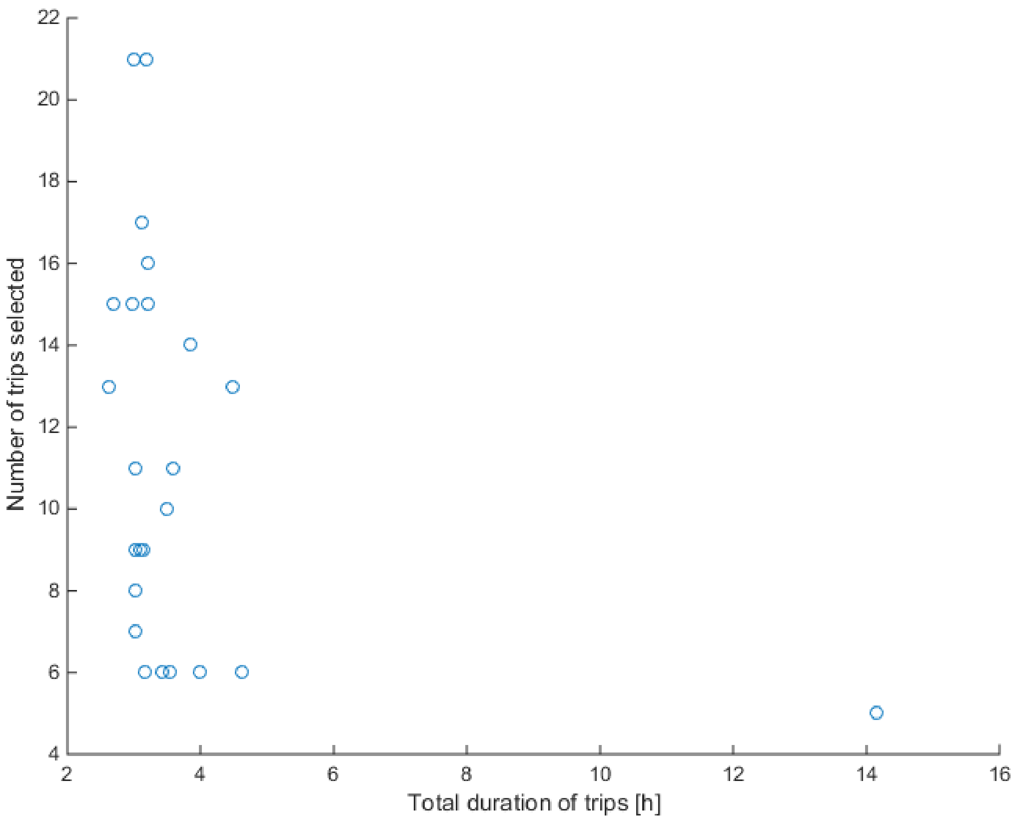



**Figure 2-3: The number of trips selected per driver versus the total duration of the annotated trips per driver**

## 2.4 Coding procedure

### 2.4.1 Overall process

Before video annotations started, the procedure on how to perform the manual annotations was defined in detail. For the secondary task variables, it was decided that annotations would be conducted in three distinct passes: Pass A, Pass B and Pass C. In each pass, annotations would focus on different aspects of the secondary task engagement or of the road environment during this specific moment

### 2.4.2 Pass A – global annotation of secondary tasks

This pass focussed on annotating the whole trips selected for engagement in secondary task activity, with tasks being defined in broad categories. Annotators visualised entire trips and, each time a driver performed a secondary task, a marker was placed inside the program (SALSA) to define the beginning, the end, and the type of task being performed. Annotators used the codebook developed for UDRIVE, which contained the detailed description of each secondary task type and the exact drivers' actions (e.g. hand / arm movement or gazes towards an object) that had to be considered to define the start and end of a task. The first videos available to be annotated allowed the refinement of the codebook, i.e. the clarification of ambiguities and the inclusion of tasks that were not defined in earlier versions.

The manual annotations performed focused on the interaction with non-embedded systems or objects. The interactions with controls such as air conditioning or radio were not part of the manual annotations. Table 2-4 shows the types of secondary tasks that were annotated.

**Table 2-4: Description of the secondary task categories for manual annotations**

| | |
|---|---|
| Phone | The driver is interacting with a mobile phone in the following manner: locating/searching; holding it in hand, lap, or some other way; pressing buttons; talking; navigating (the phone is placed somewhere and the driver looks to it frequently; the mobile phone display is always on and does not go to standby mode); putting it away.<br>If the driver is moving the lips and it is not clear if he/she is singing, talking with a passenger or talking on a hands-free phone, the annotation category is "talking singing audience unknown" |
| Electronic Device | The driver is interacting with an electronic device.<br>An electronic device can be: iPad, calculator, camera or a nomadic GPS (devices that are not integral to the vehicle).<br>Types of interactions are: locating/searching; reaching or starting to glance around; manual interaction (with finger or stylus); visual interaction; putting away. |
| Food and Drink | The driver is performing an action with a/towards a food-related or drink-related item. E.g.: reaching for cup, utensils, plate or food; eating/drinking with and without utensils; putting utensils or food away.<br>Chewing only, does not count as a secondary task (e.g. chewing gum). However, if driver reaches for chewing gum, this has to be counted as a "Food and Drink", that ends when the driver puts away the chewing gum package. |
| Smoking | The driver is glancing around for, reaching, lighting, smoking or extinguishing a cigar/cigarette. The interaction with smoking related items should also be annotated in this category.<br>If the driver is chewing tobacco and tobacco is simply in mouth during the analysis window, do not annotate as a secondary task. If the driver is reaching, spitting, putting it away, then annotate "Smoking" (same principle as for chewing gum). |
| Reading and Writing | The driver is writing in a paper or reading material that is in the vehicle, but not a part of the vehicle (i.e., not reading external signs, or centre stack display).<br>This could be reading directions, paper material, and packaging.<br>If a driver is writing/typing or reading a on a mobile phone or tablet, these actions |



UDRIVE European Naturalistic Driving Study

| | |
|---|---|
| | belong to the respective categories "Phone" and "Electronic Device". |
| Personal Grooming | The driver interacts with any item related to personal hygiene, health, or adornment.<br><br>This includes: reaching for comb, brush, make-up, razor, dental floss, contact lenses, glasses (not currently being worn), hat (not currently being worn). Removing, adjusting, or putting on clothing or jewellery are also included in this category.<br><br>If the driver quickly swipes hair out of eyes or idles twirling of hair this is not included in the annotations.<br><br>Personal grooming activities are annotated if the driver has an object in the hand. An exception is if driver is looking in the rear-view mirror to remove something/contact lens from the eye. |
| Talking/ Singing audience unknown | The driver is moving the lips but it is not clear if he/she is singing or talking on a phone using a hands-free device (headset, in-vehicle integrated system, or hands-free speakerphone). When it's clear that the driver sings or talks with a passenger this is not annotated |
| Other/ Unknown | The driver is interacting with some other object not included in the other categories or doing something that cannot be clearly identified. |

Pass A annotations had the aim of identifying the secondary task engagement without going into detail. In general, secondary tasks began when the driver started the interaction with a specific object and ended when the visual or physical contact stopped. Consecutive secondary task interactions belonging to one category were annotated as being one secondary task as long as the interval between them was not greater than 10 seconds. An example is eating crisps while driving. If the interval between the end of an action (putting the crisp in the mouth) and the beginning of another (picking another crisp from the bag) was greater than 10 seconds, then this would be annotated as two distinct "Food and Drink" secondary tasks.

The procedure accounted for the possibility that drivers could do more than one secondary task at the same time. For this reason, it was possible to manually annotate up to three simultaneous tasks. They were recorded in different layers of the annotation panel, and were all annotated following the same principles.

### 2.4.3 Pass B – detailed annotation of secondary tasks

For Pass B, annotations were not performed on full trips but only on segments of the trip that contained a mobile phone or an electronic device interaction. The aim was to identify which specific actions drivers perform while engaging in secondary tasks with both types of devices. The following tables (Table 2-5 and Table 2-6) show the categories defined for each case.

**Table 2-5: Description of the mobile phone categories for Pass B manual annotations**

| | |
|---|---|
| Search | The following actions are included: glancing to find the mobile phone, reaching towards his/her mobile phone, and/or flipping phone open or pressing a button to answer a call. This category should finish when the driver has answered the mobile phone. |
| Handheld Conversation | The driver is talking on a handheld phone or has phone up to ear as if listening to a phone conversation or waiting for person they are calling to pick up the phone. |

| | |
|---|---|
| Hands-free Conversation | The driver is talking or listening on a phone using a hands-free device such as a headset, in-vehicle integrated system, or hands-free speaker phone. If driver has an earpiece or headset, the driver must be observed talking repeatedly. |
| Hold | The driver is holding a mobile phone but not manipulating it. Could be holding it in hand, lap, or some other way. |
| Handheld Interaction | The driver is pressing buttons or a touch screen on the mobile phone. The driver can be writing a text message, browsing the internet, or interacting with other phone applications. These are mainly physical interactions (e.g. with finger) that will alternate with small pauses (just looking, looking back to the road). |
| Handheld Reading | The driver is looking at the screen and clearly reading it, without a physical interaction. If there is a physical interaction with the phone, this should be annotated as interaction. |
| Hands-free Interaction | The driver looks to the mobile phone regularly without holding it. It can happen if drivers receive navigation instructions using the mobile phone. |
| Related | The driver is interacting with a mobile phone in some manner. Includes plugging phone into charger, cleaning screen, putting on headset, etc. |

**Table 2-6: Description of the electronic device categories for Pass B manual annotations**

| | |
|---|---|
| Search | Driver reaches or starts to glance around for an electronic tablet device (e.g. iPad). |
| Interaction | The driver is pressing buttons or using the touch screen on the electronic tablet device. This interaction can be done with the hand or with a stylus. |
| Hold | The driver is holding and looking at an electronic tablet device, but not pressing any buttons. This category includes if the driver holds the device in the hands, lap, or places it in another place, but continues to read/look at it (even if intermittently). |
| Related | The driver is interacting with an electronic tablet device in some manner not described in other categories. Includes plugging tablet into charger, cleaning screen, headset, holding without manipulating, etc. |
| Other Device | The driver is interacting in some way with an electronic device that is not included in other categories and is not integral to the vehicle (e.g. calculator, camera, nomadic GPS). |

### 2.4.4 Pass C – environment annotations

In Pass C, the surrounding environment while engaging in a secondary task was annotated. Annotations of these variables were not only performed while the driver was conducting a secondary task, but started 30 seconds before and ended 10 seconds after. The following variables were annotated: weather, road surface, light conditions, presence of an intersection and right of way, presence of a construction zone, crest and site obstruction. Table 2-7 presents the categories utilised for each of the mentioned variables. All these variables were time series variables, which means that each frequency unit (1 second) had a value. Thus during a video sequence one variable could assume several distinct values.

**Table 2-7: Description of the environment variables for Pass C manual annotations**

| | |
|---|---|
| Weather | Categories: no adverse conditions; wind gusts; fog; mist or light rain; raining; snowing; sleeting; rain and fog; snow and fog; other; unknown |
| Road Surface | Categories: dry; wet; snowy; icy; muddy; sand, oil or dirt; gravel; uneven or potholes; other; unknown |
| Light conditions | Categories: day light; dawn; dusk; darkness lighted road; darkness not lighted road; unknown |
| Intersections | Categories: no intersection; intersection with traffic light and subject vehicle (SV) does not have priority; intersection with traffic light and SV has priority; intersection with traffic light and unclear priority; intersection without traffic light and subject vehicle (SV) does not have priority; intersection without traffic light and SV has priority; intersection without traffic light and unclear priority; roundabout |
| Construction zone | Whenever a construction zone was visible in the video, the beginning and end points were marked. If the vehicle was approaching or was otherwise affected by a construction zone (construction equipment, barrel, etc. are visible) this was also noted as "related"<br><br>Categories: no; yes; related; unknown |
| Crest presence | Whenever a crest was present, the beginning and end points were marked.<br><br>Categories: no; yes; unknown |
| Sight obstruction | A sight obstruction was considered whenever the driver could not see at least two seconds ahead of the road where he was traveling.<br><br>Categories: no; yes ; unknown |

## 2.5 Final sample

The numbers of tasks analysed for the car drivers and truck drivers is shown in Table 2-8. Car drivers from The Netherlands were not included in the analysis because, at the time that it was performed, data for only a handful of drivers was available.

**Table 2-8: Descriptive statistics about countries and vehicle types for secondary task analysis**

| | Number of drivers used in secondary task analysis | The average number of trips annotated per driver (SD) | The total driving duration annotated | Number of secondary tasks found |
|---|---|---|---|---|
| **Cars** | | | | |
| France | 26 | 8.3 (3.5) | 58.57 hours | Phone=178<br>Electronic Device=22<br>Smoking=15<br>Reading / writing=2<br>Food/drink=18<br>Personal Grooming=86<br>Talking/Singing=147<br>Unknown=178 |

| | | | | |
|---|---|---|---|---|
| | | | | Total=646 |
| Great Britain | 38 | 7.8 (4.0) | 73.83 hours | Phone=152<br>Electronic Device=11<br>Smoking=53<br>Reading / writing=3<br>Food/drink=61<br>Personal Grooming=51<br>Talking/Singing=144<br>Unknown=96<br>Total=571 |
| Germany | 12 | 6.9 (4.5) | 20.21 hours | Phone=4<br>Electronic Device=1<br>Smoking=0<br>Reading / writing=0<br>Food/drink=10<br>Personal Grooming=7<br>Talking/Singing=117<br>Unknown=10<br>Total=149 |
| Poland | 20 | 7.5 (3.1) | 41.40 hours | Phone=138<br>Electronic Device=153<br>Smoking=19<br>Reading / writing=12<br>Food/drink=28<br>Personal Grooming=63<br>Talking/Singing=62<br>Unknown=116<br>Total=591 |
| **Trucks** | | | | |
| Netherlands | 24 | 11.4 (4.8) | 104 hours | Control = 622,<br>Electronic device = 194,<br>Food/drink=182,<br>Personal grooming = 121,<br>Phone = 266,<br>Reading/writing = 217,<br>Smoking = 25,<br>Talking/singing = 117<br>Total = 1744 |

## 2.6 An alternative process for car driving: automatic analysis of video

The objective was to reduce the time needed for manual annotation by automatically detecting candidate episodes of secondary tasks activity in the UDRIVE data. Those candidates would then be confirmed in a second step by manual annotators. Thus the objectives are two-fold: retrieving enough true positive cases to run statistical analysis, while limiting the false positive rate on account of the limited resources available for manual confirmation. We focused on the automatic detection of handheld phone conversations and texting with a phone. Since automatic detection had to be processed over a large database, the processing time had to be tuned to provide candidates in a reasonable amount of time. This tool for automatic video annotation was developed to expand the available data on secondary tasks in naturalistic driving studies.

**Method**

A machine learning algorithm known as Convolutional Neural Networks (CNN) was used to detect phone usage from video recordings of the drivers. CNN is a well-established technique for image classification which, as with any supervised technique, requires training of the model with a set of labelled images. In our case, the algorithms were trained on a subset of UDRIVE pilot data from France (four different camera views of the driver recorded synchronously). The model performance was then assessed by comparing its predictions with actual labels on new images not used during the training phase. More details of the algorithm can be found in Dellinger et al. (2017).

As with many classification algorithms, CNN does not directly predict a label but instead, provides a measure of the likelihood that each label matches a given image. In our case, CNN estimates the probability that the driver is engaged on handheld phoning hand or texting on each video frame. We then had to decide how conservative the threshold value should be. This threshold led to classification of the current frame as positive (driver is phoning/texting) or negative (driver is not phoning/texting). Using a high threshold value decreases the number of false positives (classification as phoning/texting when the driver is not performing this secondary task) at the cost of a higher number of misses. If no or little information is available regarding the cost of misclassification, the model performance can still be assessed by computing its Receiver Operating Characteristic (ROC) curve, which relates the True Positive Rate (TPR) to the False Positive Rate (FPR) for each threshold value.

This curve, in turn, can be characterised by its Area Under Curve (AUC), allowing an assessment of model performance on a scalar value. For a given sample, AUC is actually proportional to the Mann-Witney U (Hanley, 1982), and can therefore be interpreted as the probability for a randomly chosen positive instance to be rated higher than a randomly chosen negative instance. As such, AUC ranges between 0 and 1, with a higher score indicating a better performance in the discrimination task. AUC is therefore a useful metric to compare model performance at various threshold values at the same time. If no classification cost is available, assumptions about the relative importance of True Positive Rate and False Positive Rate can be made in order to derive a corresponding threshold. In this study, equal importance was given to both rates, which defines the optimal threshold as the closest point of the ROC curve to the point located at (0,1)

**Results**

Two CNN models were trained on a subset of the UDRIVE pilot data to detect the presence of phone call or texting activity in video frames. Then, the models were used on 256 trips resulting in 140 hours of video with 10 drivers to provide automatic estimation of the likelihood of handheld phone use and texting. Videos were recorded at 10Hz leading to a test dataset of more than 5,000,000 frames. Respectively, 3% and 8% of those video frames were labelled as “Phone call” or “Texting” during the manual annotation process. The results for the handheld phone call activity are presented in Figure 2-4. The left plot displays the CNN detection score for all video frames and shows that video frames displaying a phone call activity are associated with high CNN probabilities, while video frames without call activity are spread out over lower values. Three candidate thresholds are displayed at 0.10, 0.50 and 0.90. The ROC curve is plotted on the right and nearly reaches the optimal point located at (0,1). The overall good performance of the CNN classification is translated into the high AUC score of 0.97. The point of the ROC curve which minimises the distance to the (0,1) point corresponds to a threshold value of 0.68 with a True Positive Rate of 0.94 and a False Positive Rate of 0.04. This point is displayed as a blue diamond on both plots.

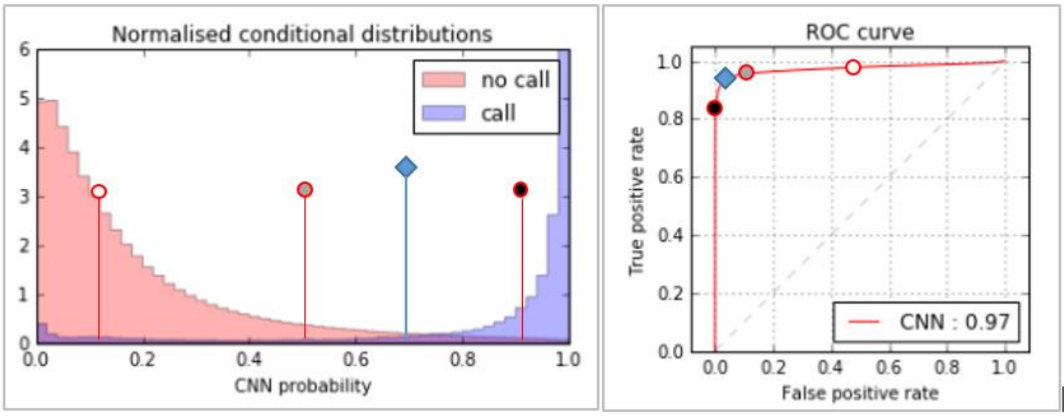



**Figure 2-4: CNN prediction values for Phone Call detection**

Similar plots are presented in Figure 2-5 for the detection of Texting activity. Video frames without Texting activity are here again associated with low CNN probabilities values, but in this case, a certain number of frames with Texting are also scored with low values. Those frames are then classified as non-positive for most of the threshold values, which decreases the true positive rate, also known as recall, while still achieving a correct AUC of 0.88 in detecting Texting activity.

In the case of the Texting detection task, the point of the ROC curve which minimises the distance to the (0,1) point is associated with a threshold value of 0.06 with a True Positive Rate of 0.80 and a False Positive Rate of 0.15.

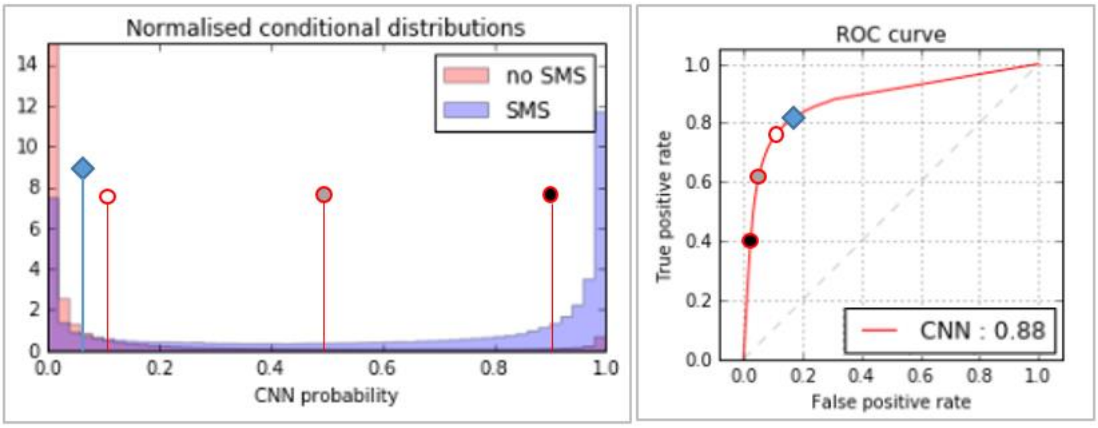



**Figure 2-5: CNN prediction values for Texting detection**

**Discussion/ Impact**

This study represents a new approach in data processing for NDS. It included the application of a deep learning algorithm using a dataset from pilot data as a learning database. As a first step, the model performances were estimated on this French Pilot dataset with promising results. This dataset was recorded with the same video devices and settings as used in the field during the UDRIVE project. Moreover, models were evaluated under various conditions. Nevertheless, real-world conditions of recording may have an impact on model efficiency. For example, discrepancy in camera positioning/tuning over the 120 UDRIVE cars or the variety of phones used database may cause divergences from the learning data base. This could lead to substantially different ROC curves. The next step of our approach will therefore be to evaluate the models on UDRIVE data and tune the models from other Operation Sites to give the better performance. For that purpose, manually annotated data is available from UDRIVE. It will be used in the next month as the ground truth to choose the threshold and tune the algorithms for these new data.

Concerning the CNN models, in the current implementation, they are trained on static video frames and no temporal structure is included in te training phase. The model prediction for a given frame is therefore fully independent from the previous and following frames, where in fact a significant amount of information could be found. Including this temporal aspect into a post-processing of the model output would represent an efficient enhancement. Sharp unrealistic drops in signal could be smoothed out

The UDRIVE data reuse group will gather UDRIVE partners willing to continue their work on UDRIVE data after the end of the project. It will be a great opportunity to extend this work and provide good quality secondary task candidates for future studies.

Adapted metrics based on precision and recall will be assessed to better estimate detection performances and expected gain compared to fully manual annotations. The first phone task candidates will be available in the next few months. Hands on steering wheel detection can also be used to study safety margin in normal driving. Automatic detection of phone use in the UDRIVE data base will reduce the time needed for manual annotation and increase the number of secondary task samples available for analysis

# 3 Car driver engagement in secondary tasks

## 3.1 What was the prevalence of secondary task activity?

The descriptive analysis performed for the car sample revealed that a secondary task occurred in 387 trips of the 745 annotated for the car sample. This means that for 51.95% of the annotated trips, drivers engaged in at least one secondary task. When considering these trips, a total of 1957 secondary tasks were performed, which represents 10.19% (19 h 42 min) of the total time travelled. The distribution of the number of secondary tasks per trip (Figure 3-1) shows that most of the trips (approx. 42%, n= 164) contained the performance of two to five secondary tasks, while in about 36% of the trips (n= 138) only one secondary task was performed. For the other 85 trips (22%), six or more secondary tasks were identified in each trip.

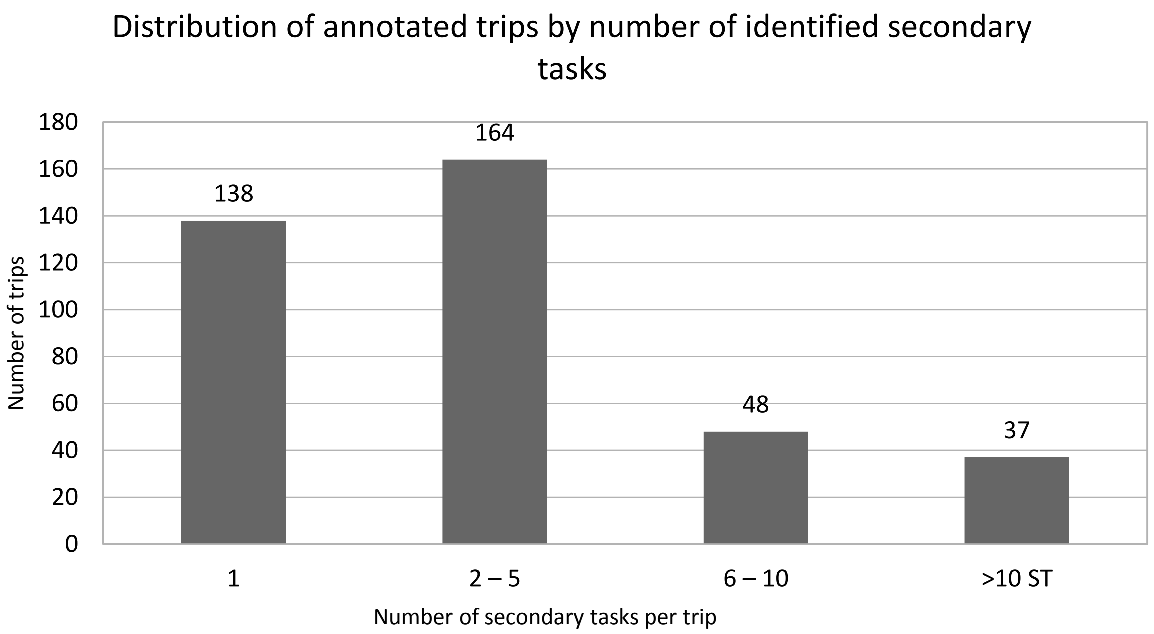



**Figure 3-1: Distribution of annotated trips by number of identified secondary tasks**

The manual annotation of secondary tasks accounted for the possibility of registering tasks performed simultaneously, i.e. while the driver was already engaged in one secondary task. Involvement in up to three secondary tasks was identified. Results revealed that for the 1957 tasks performed, 125 were initiated while the driver was already conducting another secondary task and one was performed while the driver was engaging in two other secondary tasks. When considering the total time of secondary task performance (10.19%), the analysis of single or simultaneous secondary tasks showed that single secondary task engagement accounted for approximately 19 hours (9.71% of the driven time and almost all the time regarding the overall secondary task engagement), while double secondary tasks engagement (two simultaneous tasks) accounted for a total of only 56 minutes (0.48% of the driven time). The amount of time drivers spent engaged in three simultaneous secondary tasks was 2.5 seconds.

A detailed overview on the type of secondary tasks performed allowed investigation of how often and how long drivers engaged in each of the tasks. First, the number of tasks in each category was counted (Table 3-1). Data revealed that the mobile phone tasks (n=472) and the tasks belonging to the talking/singing to audience unknown (n=470) were the most frequent ones, followed by the group of tasks “other” that integrated miscellaneous activities that did not belong to the categories defined in the codebook (n=400). Reading and writing was the task type that accounted for a lower number of task initiations (n=17).

**Table 3-1: Number of secondary tasks by task type**

| Secondary task type | Total number |
|---|---|
| Mobile Phone | 472 |
| Electronic device | 187 |
| Food and Drink | 117 |
| Smoking | 87 |
| Reading and writing | 17 |
| Personal Grooming | 207 |
| Talking/Singing | 470 |
| Other | 400 |

As a second step to analyse the type of task engagement, the total amount of time drivers dedicated to a task type was compared to the total time driven (194 hours). As can be observed in Figure 3-2, mobile phone interaction was the most represented task with a bit more than eight hours of total engagement (which accounted for 4.24% of the total time driven). Smoking followed with a total engagement time of approximately 5 hours and 39 minutes (2.91% of the time travelled). Talking/Singing to audience unknown (2h 39 min; 1.37% of the time travelled) scored third, a category related to the mobile phone usage as it might indicate a hands-free phone interaction. The least represented secondary task was the read and writing category with a total of six minutes engagement.

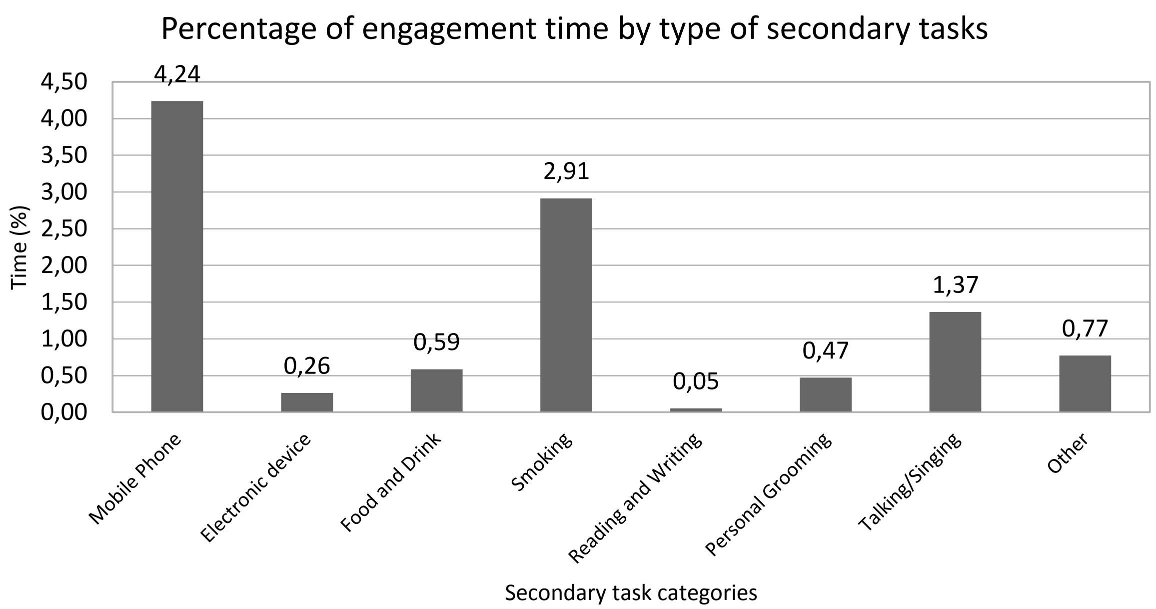



**Figure 3-2: Percentage of engagement time by type of secondary task**

The majority of the secondary tasks were between 20 and 125 seconds long with a mean at 122.76 s (Figure 3-3). There were a few outliers with extremely long durations. This is due to the fact that intermittent activities, such as smoking (with a maximum of 1800 s) are coded as one long episode. As long as the cigarette is burning, it is coded as active secondary task, independently from the position of the cigarette (mouth, hand). However, since these episodes were relatively rare, they did not have much influence on the mean of the distribution.

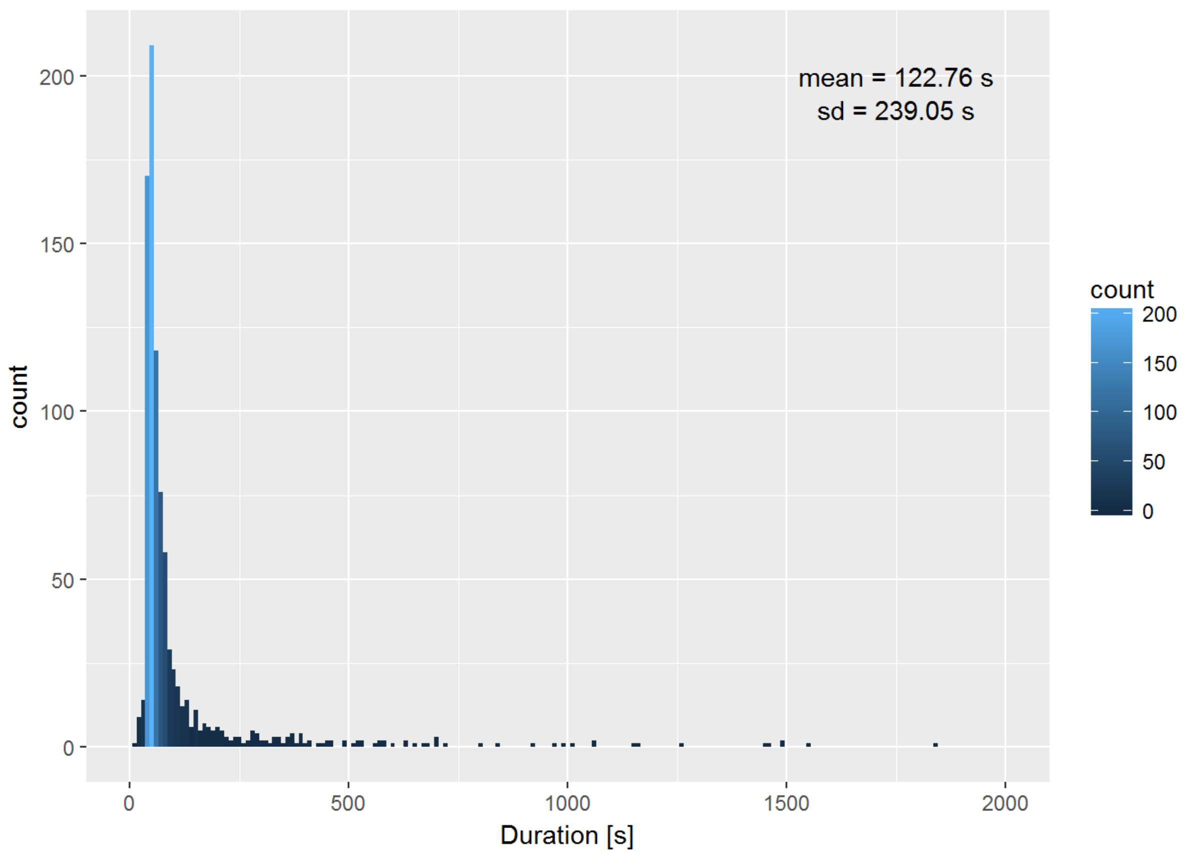



**Figure 3-3: Overview of duration of secondary tasks**

## 3.2 To what extent is the willingness to engage in secondary tasks while driving dependent on gender and cultural factors?

In order to verify the extent to which gender influences the performance of secondary tasks, the overall time each group was engaged in single and simultaneous secondary tasks was compared to the respective total time driven (M= 100.47 h; F= 93.53 h) . Table 3-2 shows that men spent a slightly higher percentage of their travel time engaged in secondary tasks (10.80%) compared to women (9.54%). The performance of two simultaneous secondary tasks was relatively similar among groups (M=0.49 %, F= 0.47%) and the only moment three tasks were performed simultaneously happened while a male participant was behind the wheel.

**Table 3-2: Overall secondary task engagement by gender group**

| Secondary task | Male | Female |
|---|---|---|
| Overall engagement (%) | 10.80 | 9.54 |
| 1 secondary task (%) | 10.31 | 9.07 |
| 2 simultaneous secondary tasks (%) | 0.49 | 0.47 |
| 3 simultaneous secondary tasks (%) | 0.000006 | 0.00 |

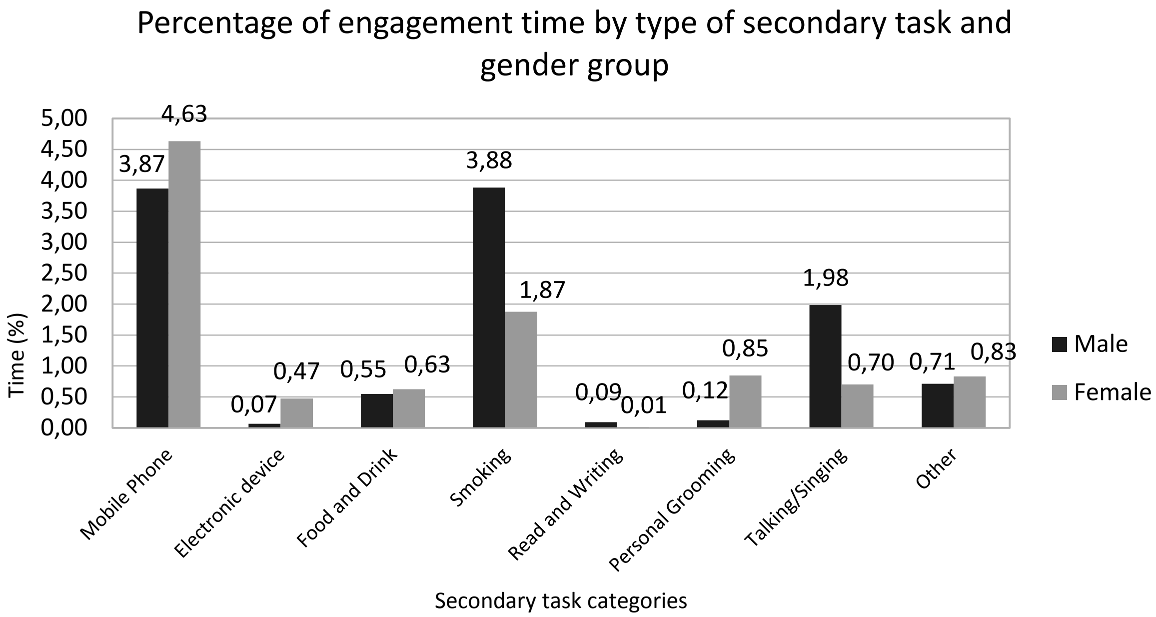



**Figure 3-4: Percentage of engagement time by type of secondary task and gender group**

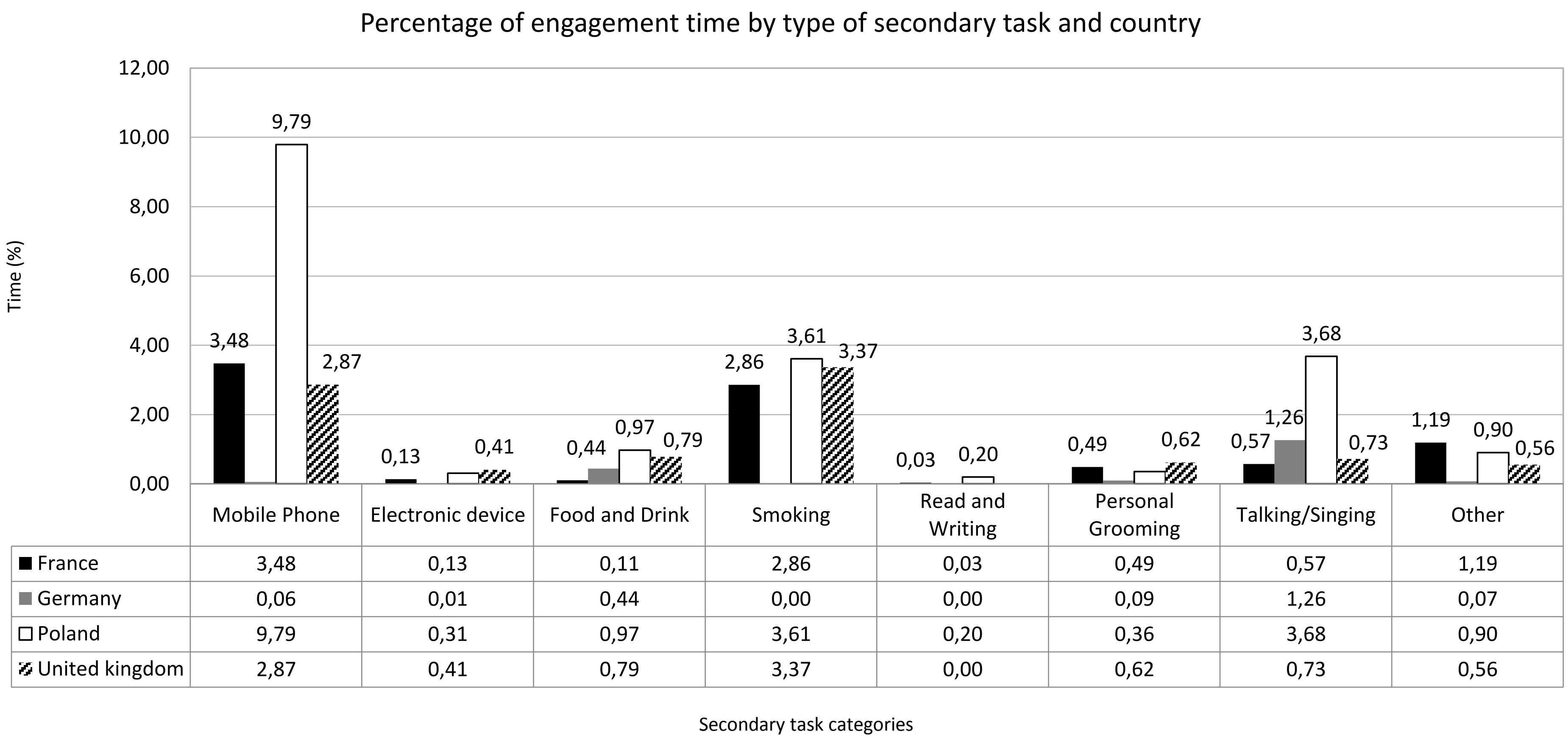
As can be observed in Figure 3-4, slight differences for the task engagement between men and women can be seen for the food and drink, read and write, and other categories. Five other tasks presented a more pronounced difference between gender groups. Women engaged longer in phone (M= 3.87%; F= 4.63%) and electronic device tasks (M= 0.07%; F= 0.47%) and also performed more personal grooming activities (M= 0.12%; F= 0.85%). The biggest differences between gender groups was verified for the smoking and talking/singing tasks. Men smoked 2.01% longer than women (M= 3.88%; F=1.87%) and also spent about 1.28% more of their time engaging in activities that might be related to interaction with hands-free phone devices (M= 1.98%; F=0.70%).

The influence of cultural factors on the performance of secondary tasks was also verified, by comparing the task engagement among countries. In term of overall task engagement, the Polish sample was the one that registered higher overall secondary task engagement (19.42%), followed by the United Kingdom (8.80%) and France (8.25%). As it can be seen in Table 3-3, the French sample was the one with higher percentage of simultaneous task performance (two tasks). The German sample was the one with the lowest overall secondary task engagement and performed no simultaneous secondary tasks while driving.

**Table 3-3: Percentage of overall secondary task engagement by country**

| Secondary task | France | Germany | Poland | United Kingdom |
|---|---|---|---|---|
| Overall engagement (%) | 8.25 | 1.93 | 19.42 | 8.80 |
| 1 secondary task (%) | 7.62 | 1.93 | 19.03 | 8.27 |
| 2 simultaneous secondary tasks (%) | 0.63 | 0.00 | 0.39 | 0.54 |

The time while performing each type of task was compared to the total driven time by country (France= 58.57 h; Germany= 20.21 h; United Kingdom= 73.83 h; Poland= 41.40 h). Figure 3-5 shows the prevalence of each secondary task type by country. It is possible to observe that for the electronic device, the food and drink, read and writing, and the personal grooming categories there were no major differences among the five countries. This homogeneity changes when the other categories are considered. A contrast among countries is present for smoking, as the German sample did not perform this task compared to the other country samples (French sample= 2.86%, Polish sample= 3.61%, United Kingdom sample= 3.37%). For the talking/singing category, the Polish sample constituted the highest activity (3.68%) followed by the German sample (1.26%). Finally, the mobile phone task is the one containing the largest difference among countries, and the Polish sample obtains the highest task representation. For the Polish drivers, 9.79% of the driven time was spent interacting with a mobile phone, while for the French drivers this accounted for 3.48% and for the UK drivers this represented 2.87%. There was negligible mobile phone interaction for the German drivers (0.06%).




| | Mobile Phone | Electronic device | Food and Drink | Smoking | Read and Writing | Personal Grooming | Talking/Singing | Other |
|---|---|---|---|---|---|---|---|---|
| France | 3,48 | 0,13 | 0,11 | 2,86 | 0,03 | 0,49 | 0,57 | 1,19 |
| Germany | 0,06 | 0,01 | 0,44 | 0,00 | 0,00 | 0,09 | 1,26 | 0,07 |
| Poland | 9,79 | 0,31 | 0,97 | 3,61 | 0,20 | 0,36 | 3,68 | 0,90 |
| United kingdom | 2,87 | 0,41 | 0,79 | 3,37 | 0,00 | 0,62 | 0,73 | 0,56 |

**Figure 3-5: Percentage of engagement time by type of secondary task and country**

 Page 20

### 3.3 What specific mobile phone sub-tasks do drivers typically engage in and to what extent is the willingness to engage in these tasks dependent on gender and cultural factors?

In order to better understand what drivers do while interacting with mobile phones (the task with the highest engagement percentage time) this task was further annotated and broken down into different actions. As it can be seen from Table 2-5, actions while interacting with the mobile phone were coded into the following categories: search, handheld conversation (HH conv.), hands-free conversation (HF conv.), hold, handheld interaction (HH int.), handheld reading (HH read), hands-free interaction (HF int.), and related.

At the time of data extraction for the production of this deliverable, not all 472 identified mobile phone tasks had been annotated for Pass B, but only 304. These 304 were used to perform a detailed analysis of the mobile phone involvement while driving, which represented about 7 hours and 40 minutes. Figure 3-6 shows the mobile phone actions in relation to the overall time spent involved in these 304 mobile phone tasks. When the search and related categories are not considered, the general handheld utilisation of such device accounted for about 36% of the time (HH conv.= 5.82%; Hold= 6.53%; HH int.= 21.97%; and HH read= 1.68%). If one considers that both search and related categories also contain manipulations of the mobile phone, then the percentage of handheld utilisation increases to 39%. From the four handheld categories, the handheld interaction (HH int), i.e. pressing buttons or touching the screen on the mobile phone, was the action with highest relative engagement time. Hands-free interaction with the mobile phone (HF conv. and HF int.) accounted for about 61% of the overall time interacting with such device for the 304 tasks analysed.

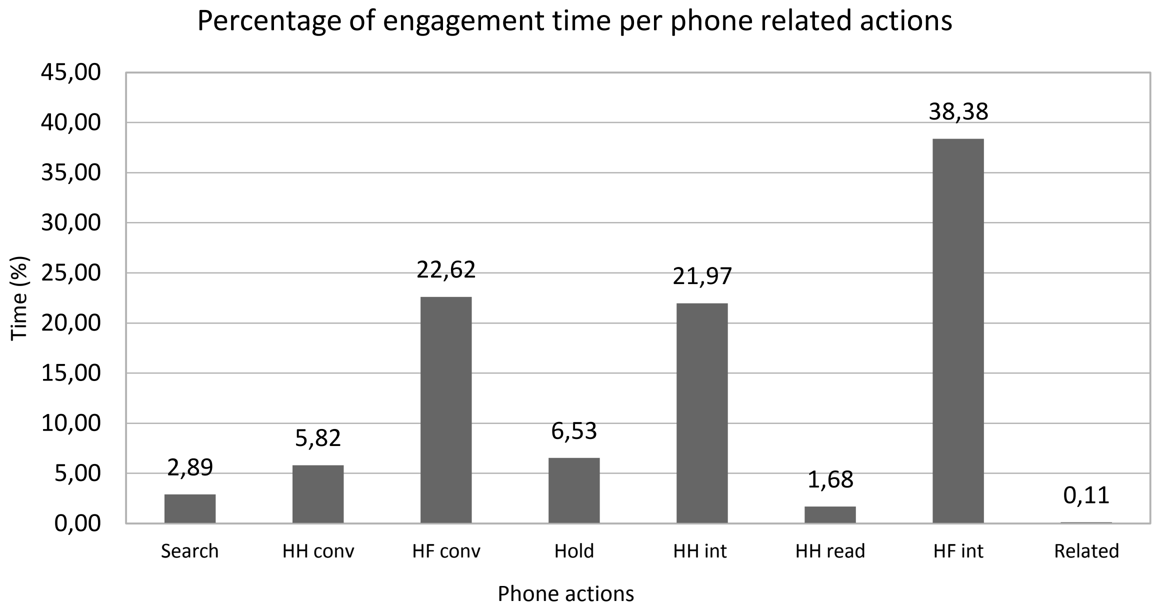



**Figure 3-6: Overall percentage of engagement in mobile phone actions**

A gender analysis was also performed to verify the willingness to engage in each of the mobile phone sub-tasks while driving. Regarding the 304 mobile phone tasks taken into account for this analysis, it is possible to state that men and women were similarly engaged in mobile phone tasks (approximately 3 hours and 50 minutes for both gender groups). Figure 3-7 reveals that, when the search and related categories are not considered, men were involved in handheld action for about 34% of the total time interacting with a mobile phone, while women did it for about 38% of their respective time. When these two categories are included, women still obtain the highest handheld time percentage (M=37.11%; F= 41.69%). Conversely, male participants showed a longer hands-free mobile phone engagement (M= 63.67%; F=58.31%, when HF conv. and HF int. are considered together). Men spent more than half of their engagement time looking at the mobile phone regularly without holding it (HF int., an action frequently performed when the driver receives navigation instructions from the device). Women spent the majority of their engagement time talking or listening on a phone using a hands-free device (HF conv.) and pressing buttons or a touch screen on the cell phone (HH int.).

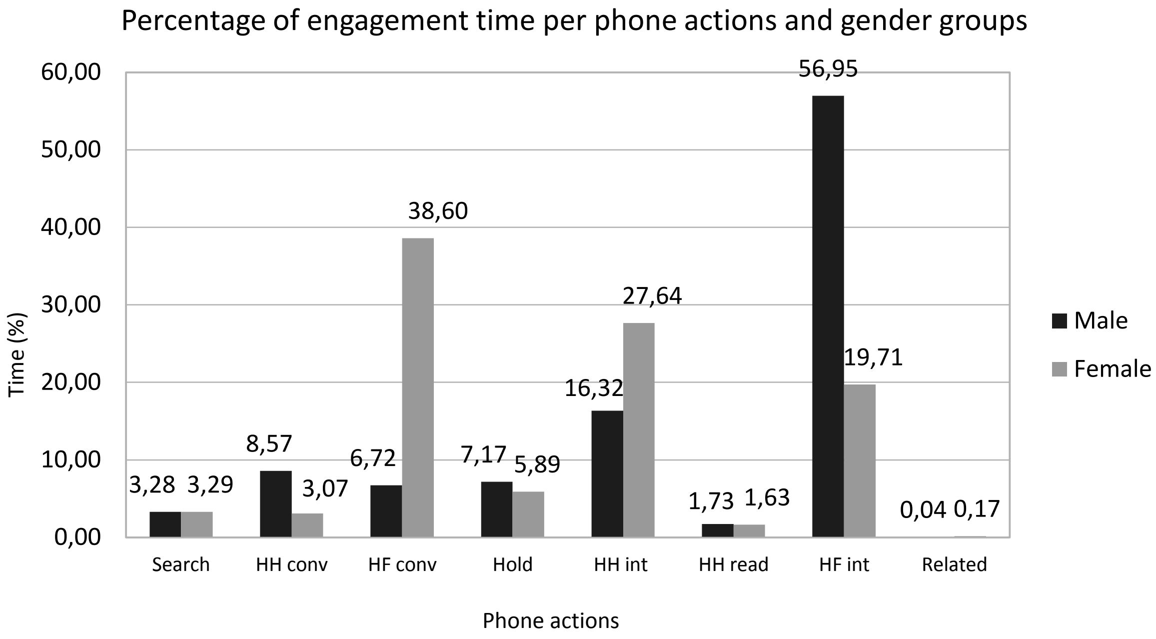
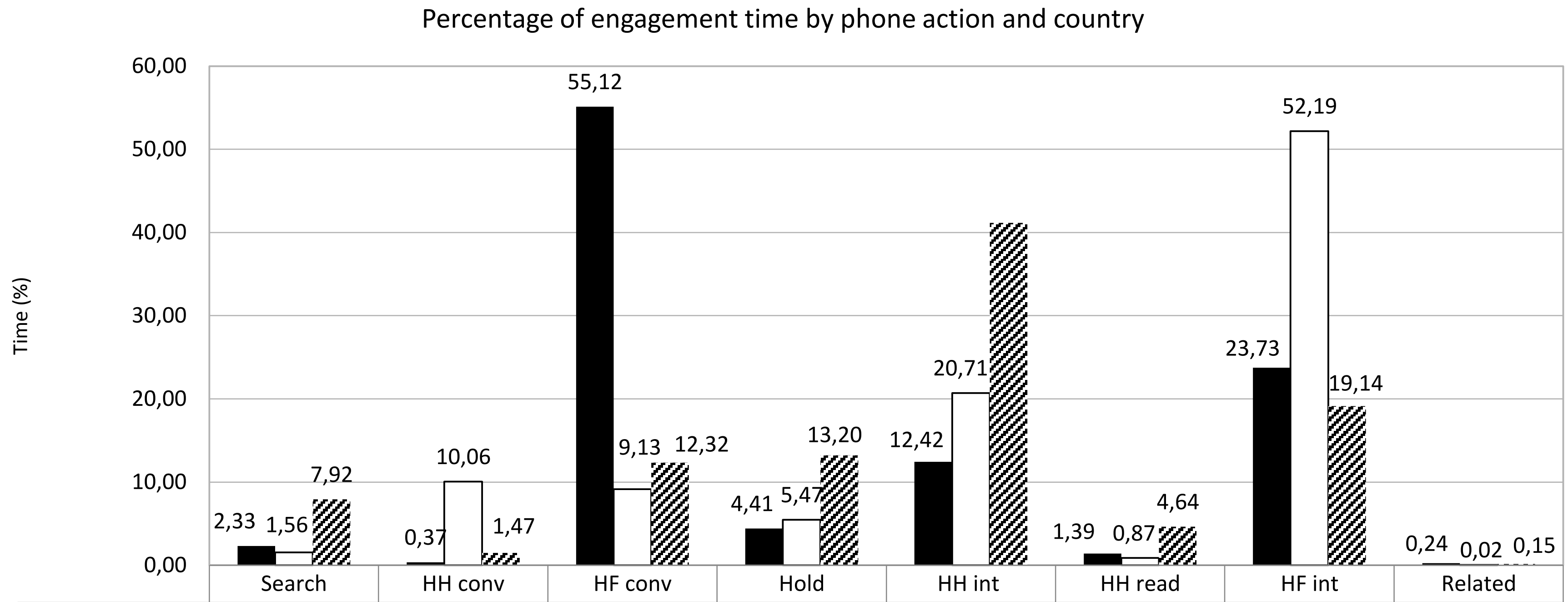



**Figure 3-7: Percentage of engagement time per phone action and gender group**

The comparison between the samples from the various countries is presented in Figure 3-8. Note here that inside the 304 annotated mobile phone tasks, none belonged to the German sample. The reason for this to happen was not a methodological one; it was just due to the time when these data arrived in the database. Results from France, Poland and United Kingdom are presented (respective approximate overall mobile phone engagement time for this sample: France= 2 hours and 10 min; Poland=4 hours and 20 min; United Kingdom= 1 hour and 10 min). The analysis of handheld mobile phone actions (HH conv., Hold, HH int., HH read.) reveals that the United Kingdom sample was the one with higher “related” engagement times (60.46%). This is also true when both search and related categories are considered inside this group (68.54%). The Polish sample followed in terms of handheld engagement (37.10%; 38.68% with search and related cat.), and the French sample had the lowest values (18.58%; 21.15% with search and related cat.). When taking into account the hands-free categories, the scenario is the inverse. The French sample spent about 79% of their time engaged in hands-free mobile phone sub-tasks, the Polish sample about 61% and the sample from United Kingdom approximately 31%.




| | Search | HH conv | HF conv | Hold | HH int | HH read | HF int | Related |
|---|---|---|---|---|---|---|---|---|
| France | 2,33 | 0,37 | 55,12 | 4,41 | 12,42 | 1,39 | 23,73 | 0,24 |
| Poland | 1,56 | 10,06 | 9,13 | 5,47 | 20,71 | 0,87 | 52,19 | 0,02 |
| United kingdom | 7,92 | 1,47 | 12,32 | 13,20 | 41,16 | 4,64 | 19,14 | 0,15 |

**Figure 3-8: Percentage of engagement time by phone action and country**

 Page 23

## 3.4 What was the influence of personality factors?

This section looks at the impact of personality on secondary task engagement. First, there is a consideration of individual personality components as measured by a range of questionnaires completed by drivers at the outset of the study. The second part of the section uses this suite of personality metrics to produce a summary measure of driver personality – a negative driver personality traits score. The potential utility of this score as a means of predicting driver engagement in secondary tasks whilst driving is considered.

### 3.4.1 Method

Here, the personality questionnaire data that was collected when participants were recruited was analysed to investigate whether personality affected propensity to use a mobile phone while driving. The sample of drivers and trips used was the same as that used in the prevalence analysis above (Section 3.1). There were 745 trips by 96 drivers that were annotated for mobile phone use. Of those trips, 387 by 88 drivers involved secondary task engagement at some time during the trip. There was complete questionnaire date for 92 drivers out of the 96.

Drivers were presented with a suite of questionnaires at the beginning of the project, including:

- Driver Attitude Questionnaire (DAQ) (20 items assessing attitudes towards speeding and close following behaviours);
- Driver Behaviour Questionnaire (DBQ) (19 items assessing the prevalence of errors and violations in the driver's everyday behaviours);
- Driver Skills Questionnaire (DSQ) (15 items assessing how drivers behaviour in a series of described driving scenarios including speeding behaviour, travelling with passengers, engaging with distractions, journey planning etc.)
- Traffic Locus of Control (TLOC) Questionnaire(17 items assessing views towards the factors that cause road accidents);
- Arnett Inventory of Sensation Seeking (AISS) (20 items assessing the risk-taking and sensation-seeking nature of a driver's personality).

The five questionnaires above have been subjected to factor analysis in prior work, which has been assumed valid for the purposes of this subsequent analysis (Warner et al., 2010; Department for Transport, 2005; Özkan and Lajunen, 2005; Lajunen et al., 2004; Parker et al., 1996; Arnett, 1994; French et al., 1993; West et al., 1993). The subscales derived from factor analysis have been used as independent variables to represent attributes of driver personality.

For each subscale of a questionnaire (e.g. Driver Attitude Questionnaire, speeding subscale) and for the overall summary scale (e.g. DAQ, overall), drivers were ranked by their score before being split into two groups. The break point between the two groups was selected such that all drivers fell into one group or the other, with no overlap in scores between the two groups. In some cases, this results in slight disparities in the number of members of these two groups. This process was done so as to dichotomise a continuous variable for subsequent exploratory data analysis (Table 3-4). This was selected in preference to a median split as it resulted in more balanced groups in terms of the number of members.

**Table 3-4: Splitting driver sample along personality dimensions**

| Questionnaire Sub-scale | Group membership | | Group description |
|---|---|---|---|
| | Group 1: Mean score /# members | Group 2: Mean score /# members | |



UDRIVE European Naturalistic Driving Study

| | | | | | |
|---|---|---|---|---|---|
| DAQ speeding | 27.8 | 46 | 38.2 | 46 | High score = more negative attitude towards speeding |
| DAQ close following | 34.4 | 49 | 43.1 | 43 | High score = more negative attitude towards close following |
| DAQ overall | 62.6 | 45 | 79.9 | 47 | High score = more negative attitude towards speeding and close following behaviours |
| DBQ errors | 1.4 | 49 | 2.0 | 39 | High score = more reported driving errors |
| DBQ aggressive violations | 1.5 | 42 | 2.4 | 36 | High score = more reported aggressive driving violations |
| DBQ ordinary violations | 1.2 | 38 | 2.2 | 53 | High score = more reported ordinary driving violations |
| DBQ all violations | 1.5 | 38 | 2.3 | 40 | High score = more reported driving violations |
| DSQ speed | 6.1 | 43 | 11.2 | 49 | High score = more reported speeding behaviour |
| DSQ calmness | 11.3 | 39 | 15.5 | 52 | High score = more reported calm driving behaviour |
| DSQ social resistance | 5.7 | 52 | 9.6 | 38 | High score = more reported resistance to others' advice |
| DSQ focus | 10.2 | 52 | 14.8 | 38 | High score = more reported cautious driving and resistance to distraction |
| DSQ planning | 6.1 | 48 | 10.4 | 45 | High score = more reported planning ahead before and during driving |
| DSQ deviance | 2.0 | 35 | 4.0 | 58 | High score = more reported rule-breaking and deviant behaviours |
| TLOC self | 2.2 | 42 | 3.5 | 50 | High score = rate their own driving as contributing to the cause of road accidents |
| TLOC other | 3.6 | 47 | 4.5 | 45 | High score = rate the driving of others as contributing to the cause of road accidents |
| TLOC vehicle and environment | 2.9 | 47 | 4.0 | 44 | High score = rate vehicle and environmental factors as contributing to the cause of road accidents |
| TLOC fate | 1.9 | 46 | 3.3 | 45 | High score = rate coincidence or fate as contributing to the cause of road accidents |
| AISS novelty | 22.4 | 48 | 29.3 | 42 | High score = drivers seek out novel experiences |
| AISS intensity | 18.4 | 50 | 25.9 | 42 | High score = drivers seek out high intensity experiences |
| AISS overall | 41.8 | 42 | 52.6 | 47 | High score = drivers seek out highly novel or high intensity experiences |

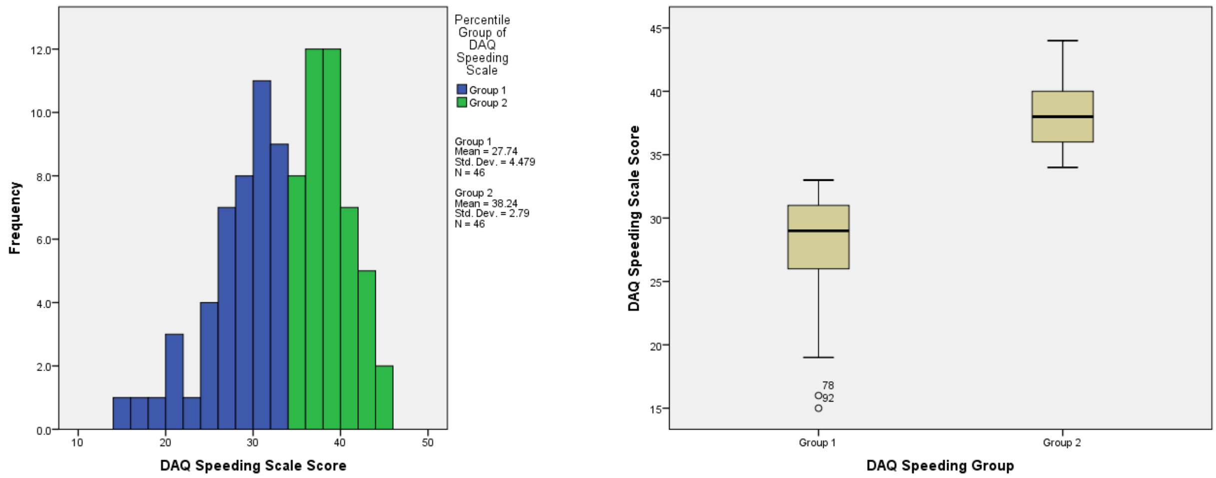



**Figure 3-9: Example of categorization along a driver personality dimension [Left: histogram showing category membership; Right: boxplot showing non-overlapping categories]**

### 3.4.2 Proportion of total driving time with phone use

745 trips were annotated for the presence of secondary tasks, with 96 car drivers represented in the sample. Individual drivers contributed between 3 and 22 trips to the dataset, and thus, to allow for differences in exposure, this analysis focuses on the proportion of total driving time that drivers engaged in mobile phone secondary tasks, rather than the number of task engagements.

127 instances of mobile phone interaction were seen in the annotated videos, involving 50 drivers. For the first step of this analysis, all drivers were retained in the sample, and independent-samples t-tests were used to compare the proportion of time spent interacting with a phone across driver personality categories.

The two DAQ Speeding categories differed significantly in the proportion of driving time for which they interacted with a mobile phone [$t(47.94)=2.276$, $p=.027$, $d=.047$]. Drivers who held a more negative view of speeding behaviour were less willing to interact with a mobile phone during driving (0.7% of time) than those who held more positive attitudes towards speeding (5.7% of time) (Figure 3-10). This suggests that drivers who hold negative attitudes towards aberrant speeding behaviours may also hold negative attitudes towards secondary task engagement. All remaining comparisons between questionnaire subscale categories showed no difference in the proportion of time that drivers spent interacting with their phone.

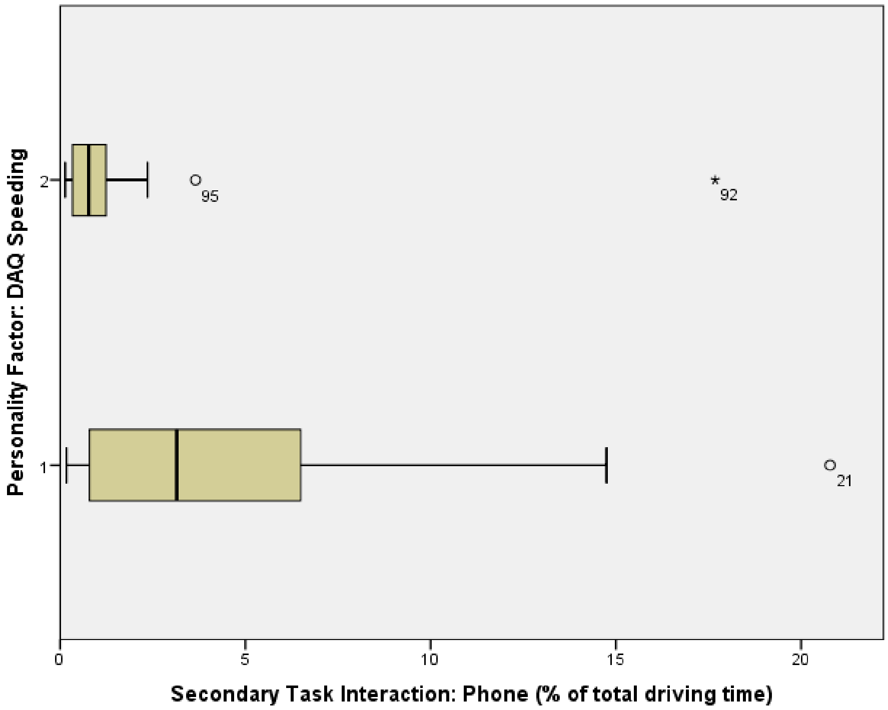



**Figure 3-10: Phone interaction per DAQ Speeding category; Category 1 has a less negative view towards speeding behaviour. For display purposes, non-users and two instances of phone interaction > 20% of driving time have been omitted.**

Drivers in each personality category were subsequently categorised simply as mobile phone users or non-users, depending on whether they were observed interacting with their mobile phone during any annotated trip (Table 3-5).

**Table 3-5: Phone use by personality sub-group**

| Questionnaire subscale | Group 1 Users (%) | Group 2 Users (%) | Difference (Group 2 % Use – Group 1 % Use) |
|---|---|---|---|
| DAQ speeding | 65 | 35 | -30 |
| DAQ close following | 55 | 44 | -11 |
| DBQ errors | 51 | 46 | -5 |
| DBQ aggressive violations | 47 | 51 | 4 |
| DBQ ordinary violations | 50 | 61 | 11 |
| DSQ speed | 44 | 57 | 13 |
| DSQ calmness | 46 | 54 | 8 |
| DSQ social resistance | 48 | 55 | 7 |
| DSQ focus | 56 | 42 | -14 |
| DSQ planning | 54 | 47 | -7 |
| DSQ deviance | 51 | 50 | -1 |
| TLOC self | 62 | 40 | -22 |
| TLOC other | 49 | 51 | 2 |

| | | | |
|---|---|---|---|
| TLOC vehicle and environment | 53 | 48 | 5 |
| TLOC fate | 43 | 58 | 15 |
| AISS novelty | 42 | 60 | 18 |
| AISS intensity | 44 | 57 | 13 |

Table 3-5 suggests that driver personality can affect whether or not they engage with a mobile phone secondary task:

- High negative attitudes toward speeding is linked to a lower likelihood of interacting with a mobile phone — 35% of drivers in the more negative DAQ Speeding group interacted with a phone, compared to 65% in the less negative group.
- Drivers who report committing more ordinary violations (e.g. close following, inappropriate speed, poor driving etiquette) are more likely to engage in a mobile phone task (61%) compared to those reporting lower ordinary violations (50%).
- Drivers who report more speeding behaviour on the DSQ show higher mobile phone use (57% of drivers) than those reporting lower levels of speeding behaviour.
- Drivers who report greater focus and a strong ability to ignore distractions show a lower level of phone use (42% of drivers) than those who report being less effective at these skills (56% of drivers interact with a phone). This subscale appears to capture to an extent, a driver's willingness to engage with a distracting task in the vehicle.
- Drivers who score highly on the Self subscale of the TLOC — who perceive the driver as having a significant contribution to the cause of road accidents — have a lower level of mobile phone engagement (40% of drivers) than those who score lower on this locus of control factor (62%). This suggests that higher awareness of how their own behaviour contributes to road accidents can potentially modulate the extent to which drivers engage with a distracting mobile phone task.
- Those drivers who perceive Fate to be an important contributor to road accidents show higher engagement with mobile phones (58% of drivers) compared to those who attribute less road accident causality to fate or coincidence (43% of drivers). This aligns with the above finding, in that drivers who perceive greater control over the cause of accidents are less willing to perform a mobile phone secondary task.
- High sensation-seeking drivers are more likely to interact with a mobile phone task. Those who score highly on the AISS Novelty subscale where more likely to perform a mobile phone task (60% of drivers) than those with a lower score (42%). The same pattern was observed for individuals scoring high on the Intensity subscale (57% performed a mobile phone task) compared to those scoring low (44%). Figure 3-11 shows that of those who engage in phone interaction, there are more instances of prolonged use in the high sensation-seeking group (Category 2) than the lower sensation-seeking group.

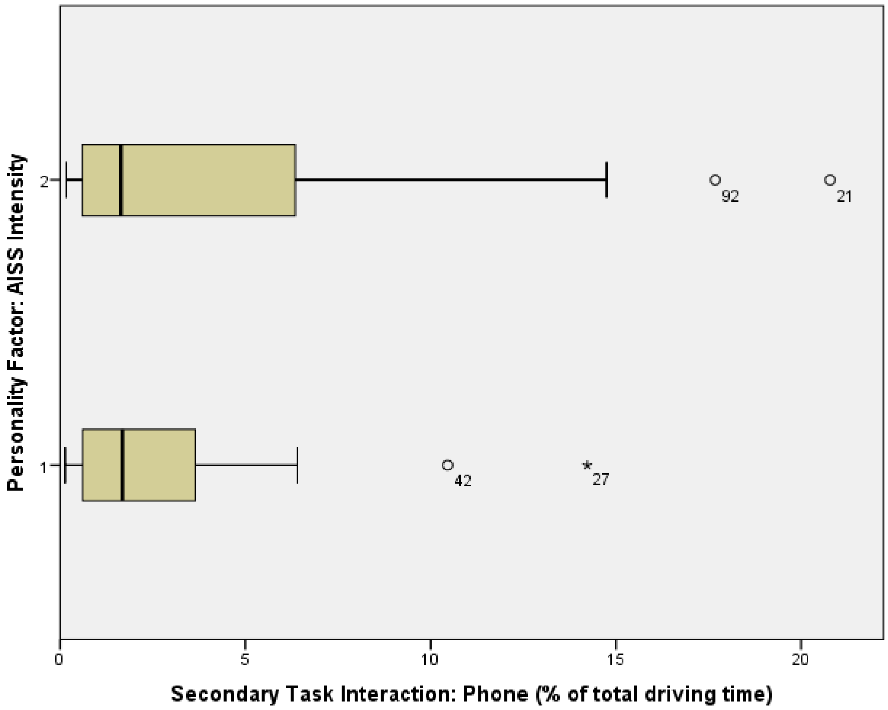



**Figure 3-11: Phone interaction per AISS Intensity category (1 = high sensation seeking behaviour). For display purposes, non-users and two instances of phone interaction > 20% of driving time have been omitted.**

Overall, then the suite of driver personality questionnaires appear to have some value for predicting driver engagement with a mobile phone whilst driving. Another conclusion is that general attitudes towards safety in driving influence the choice of whether or not to use a mobile phone.

### 3.4.3 Proportion of total driving time with electronic device use

This analysis focuses on the proportion of total driving time that drivers engaged in electronic device tasks. Independent samples t-tests were used to compare the two subgroups of participants created for each personality subscale (see Table 3-4).

27 instances of electronic device interaction were seen in the annotated videos, involving 17 drivers.

There were no significant between-group differences in the proportion of driving time for which drivers in interacted with an electronic device, regardless of personality subscale.

Drivers in each personality category were subsequently categorised simply as electronic device users or non-users, depending on whether they were observed interacting with their device during any annotated trip (Table 3-6).

**Table 3-6: Electronic device use by personality sub-group**

| Questionnaire subscale | Group 1 Users (%) | Group 2 Users (%) | Difference (Group 2 % Use – Group 1 % Use) |
|---|---|---|---|
| DAQ speeding | 13 | 22 | 9 |
| DAQ close following | 22 | 12 | -11 |
| DBQ errors | 22 | 10 | -12 |
| DBQ aggressive violations | 18 | 17 | -1 |
| DBQ ordinary violations | 21 | 17 | -5 |

| | | | |
|---|---|---|---|
| DSQ speed | 19 | 16 | -2 |
| DSQ calmness | 21 | 13 | -7 |
| DSQ social resistance | 15 | 18 | 3 |
| DSQ focus | 17 | 18 | 1 |
| DSQ planning | 19 | 16 | -3 |
| DSQ deviance | 11 | 21 | 9 |
| TLOC self | 21 | 14 | -7 |
| TLOC other | 23 | 11 | -12 |
| TLOC vehicle and environment | 21 | 14 | -8 |
| TLOC fate | 24 | 11 | -13 |
| AISS novelty | 15 | 19 | 4 |
| AISS intensity | 22 | 10 | -12 |

Table 3-6 suggests that driver personality can affect whether or not they engage with a mobile phone secondary task:

- High negative attitudes toward close following seems to be linked to a lower likelihood of interacting with an electronic device— 12% of drivers in the more negative DAQ Speeding group interacted with a phone, compared to 22% in the less negative group.
- Drivers who report making less driving errors are more likely to engage in an electronic device task (22%) compared to those reporting making more errors (12%). This suggests that drivers who identify themselves as more error prone, may reduce their interaction with in-vehicle devices.

Overall, the infrequent use of electronic devices in this driver sample is likely to confound the analysis, and thus these findings should be viewed as initial trends.

### 3.4.4 Composite personality measure

The composite personality score includes six questionnaires given to drivers at the outset of the study, measuring driver attitudes, driver behaviour, driving skill, traffic locus of control, sensation-seeking and driving history. The raw data for each of these questionnaires involved an ordinal scale with 4–6 intervals. These were converted to a three-point scale in each case as detailed in Table 3-7. The cut-off points for the revised scale were determined based on inspection of the raw data to assess which response categories were logically grouped together or infrequently used. The objective was to produce a new scale from 0–2, which could be considered to have three distinct categories (e.g. negative, neutral, positive). All scales were recoded such that a high score reflected the most negative driver attitude or behaviour, such as high acceptance of close-following behaviour, high reported speeding, high reported mobile phone use (see Table 3-7).

**Table 3-7: Construction of composite personality measure**

| Questionnaire | Original Scale | Revised Scale |
|---|---|---|
| Driver Attitudes Questionnaire | 1 - I strongly disagree<br>2 - I disagree<br>3 - I neither agree or disagree<br>4 - I agree<br>5 - I strongly agree | 0 – Disagree<br><br>1 – Neutral<br>2 – Agree |
| Driver Behaviour Questionnaire | 1 - Never<br>2 - Hardly ever<br>3 - Occasionally<br>4 - Quite often<br>5 - Frequently<br>6 - Nearly all the time | 0 – Never<br>1 – Rarely<br><br>2 – Regularly |
| Driver Skills Questionnaire | 1 – Never or very infrequently<br>2 - Infrequently<br>3 – Quite infrequently<br>4 – Quite frequently<br>5 - Frequently<br>6 – Very frequently or always | 0 – Never<br>1 – Rarely<br><br>2 – Regularly |
| Traffic Locus of Control Scale | 1 – Not at all possible<br>2 – Not fairly possible<br>3 - Possible<br>4 – Fairly possible<br>5 – Highly possible | 0 – Not a factor in accident causation<br>1 – Weak factor in accident causation<br>2 – Strong factor in accident causation |
| Arnett Inventory of Sensation Seeking | 1 – Does not describe me at all<br>2 – Does not describe me very well<br>3 – Describes me somewhat<br>4 – Describes me very well | 0 – Does not describe me<br><br>1 – Describes me to some extent<br>2 – Describes me well |
| Driving History Question – Driver at Fault Accidents | 0<br>1<br>2, 3, 4, or 5 | 0 – No accidents<br>1 – 1 accident<br>2 – More than 1 accident |
| Driving History Question – Reported bad behaviours | 1 - Never<br>2 - Hardly ever<br>3 - Occasionally<br>4 - Quite often<br>5 - Frequently<br>6 - Very often | 0 – Never<br>1 – Rarely<br><br>2 – Regularly |

This recoding procedure resulted in 17 personality subscales scored on a 0-–2 scale. Traffic locus of control and sensation-seeking are not inherently positive or negative characteristics so these 5 subscales were removed from further consideration. The remaining twelve subscale scores relating to driver behaviour, attitudes and skills were summed to give a personality score out of 24. This score was used as a continuous variable in subsequent analyses. Ninety-three drivers with trips annotated for secondary task engagement submitted sufficient personality data to calculate a summary score. Figure 3-12 shows the distribution of these scores amongst the sample. The distribution is positively skewed with a mean personality score of 7.6 (range: 2.7–14.4; IQR: 6.2–8.8). High scores reflect high reporting of negative behaviours, attitudes and skills. The dataset shows few instances of drivers giving themselves high ratings, suggesting a reluctance to criticise themselves during self-report.

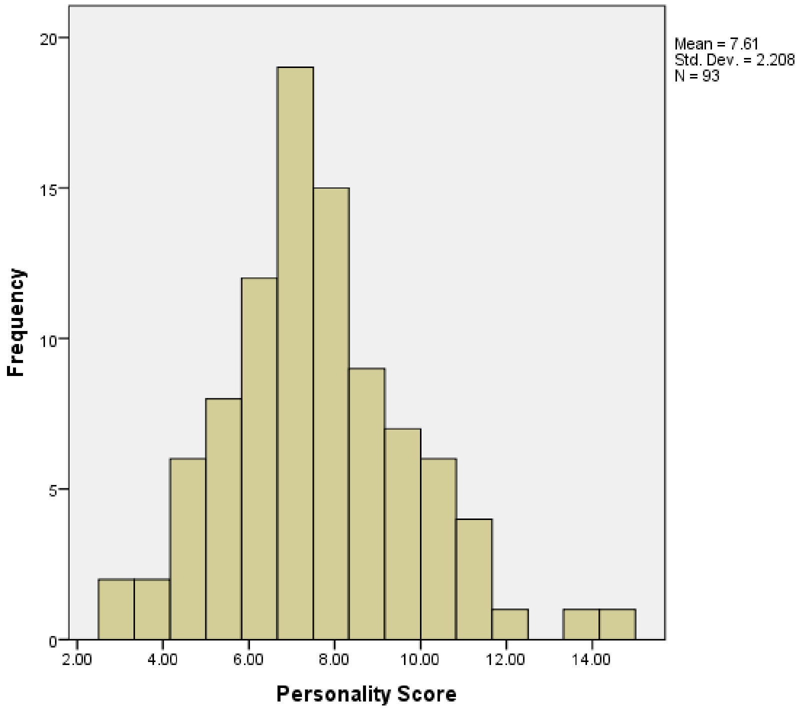



**Figure 3-12: Histogram of driver personality scores**

### 3.4.5 Secondary task engagement and personality score

The composite personality measure was used to determine whether there was any difference in personality score between those that did and did not engage with each type of secondary task during their driving. Each driver was categorised as a user/non-user based on whether the secondary task type was performed during any annotated drive; for example a driver who used a mobile phone in one out of three annotated trips was categorised as a 'user'. This simple categorisation provides a dichotomy of whether an individual driver is willing to perform a specific task during driving.

Independent-sample t-tests were used to compare the dependent variable (personality score) for two between-subjects groups (users vs. non-users). Prior exploratory data analysis showed the personality variable to be normally distributed (Kolmogorov-Smirnov test) hence parametric analysis methods were applied. Figure 3-13 shows the mean personality score per secondary task engagement group. In all cases, the tendency is for those drivers who engage in a particular secondary task to have higher scores (more negative personality traits) than those who do not engage in the task. This effect reaches significance in certain cases.

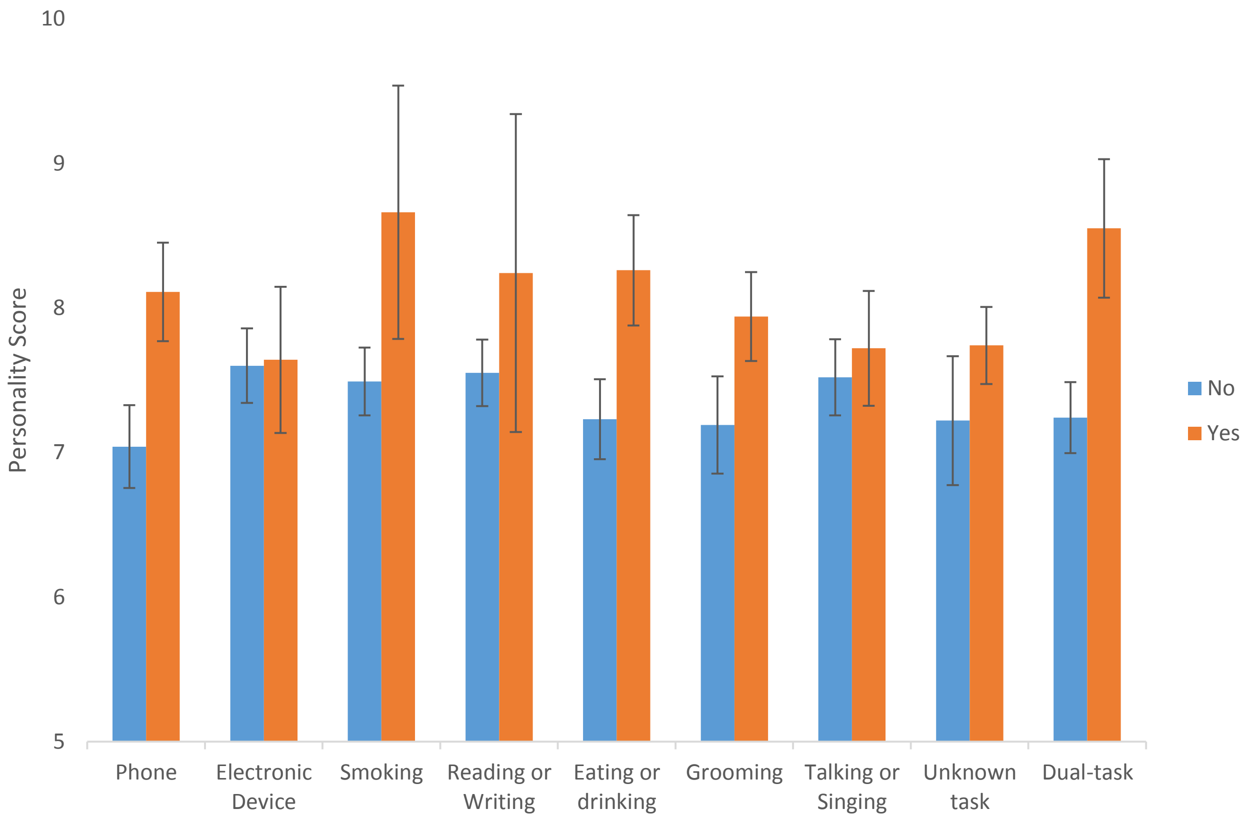


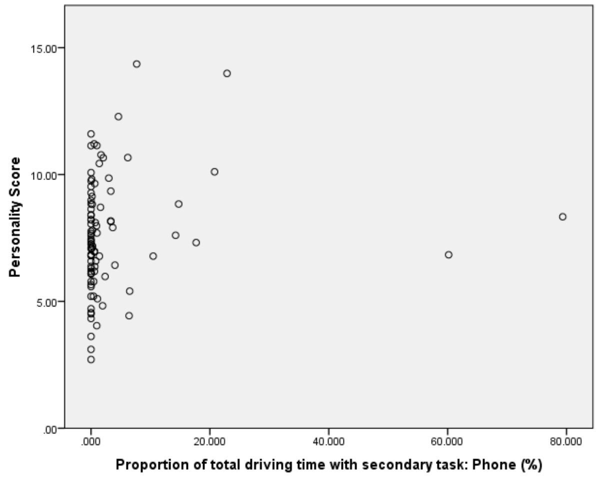

**Figure 3-13: Personality score x Secondary task engagement; dual-task category refers to any combination of two tasks**

For mobile phone use, the 47 drivers who used a phone has a significantly higher personality score (M=8.12, SD=2.33) than the 46 drivers who did not use a phone (M=7.04, SD=1.94), t(91) = -2.530, p=.013. This suggests that those drivers who hold more negative driving attitudes or who exhibit more risky behaviours are those most likely to interact with a phone whilst driving. Food and drink consumption also showed the same effect, with the 34 drivers who consumed during driving having a higher personality score (M=8.26, SD=2.23) than those 59 drivers who did not consume (M=7.23, SD=2.12), t(91)=-2.199, p=.030. There were differences in personality score between the smokers and non-smokers (difference = 1.17) and the reader/writers and non-reader/writers (difference = 0.69). These differences did not reach significance, partly due to the low number of drivers who engaged in each of these secondary tasks (9 drivers and 8 drivers respectively). The identical pattern of results were observed when the personality score was adjusted to include a TLOC and AISS component. In this case, the reading/writing group difference also reached statistical significance.

91% of drivers engaged in a single secondary task during at least one drive, hence no comparison of personality score was performed between users and non-users here. In the case of dual-task engagement, there was a significant difference in personality score between the two groups, with the 67 drivers who did not engage in two tasks simultaneously having a lower mean score (M=7.24, SD=2.01) than the 26 drivers who did multi-task (M=8.55, SD=2.44) ), t(91) = -2.650, p=.009. This means that those drivers who performed multiple task scored higher for negative driving personality traits. The largest mean difference between the user and non-user groups was observed for dual-task performance. This suggests that drivers with more negative driving personality traits are more likely to allow themselves to engage in the most demanding secondary task interactions.

### 3.4.6 Secondary task proportion and personality score

A simple linear regression was calculated to predict proportion of driving time spent using a phone based on driver personality. Personality score was found not to be a significant predictor of phone use time, however Figure 3-14 demonstrates that this analysis is likely to be skewed by the high proportion of drivers who do not use their mobile phone whilst driving. When mobile phone use for between 2-50% of driving time is considered, there appears to be a trend for negative personality score to increase with increasing phone use.



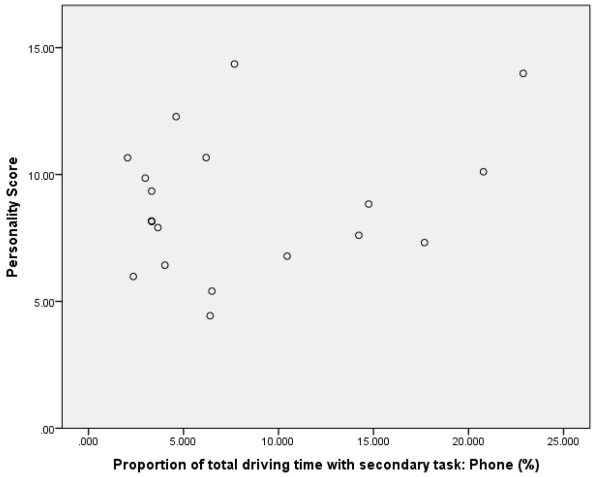




**Figure 3-14: Proportion of driving time using phone by personality score; left - all data; right - non-users and excessive user outliers excluded**

## 3.5 Did driving task complexity and secondary task complexity influence the decision to engage in secondary tasks?

### 3.5.1 Introduction

Secondary tasks are tasks that are not related to drivers' primary task of safely controlling and manoeuvring the vehicle. Research shows that drivers frequently engage in such tasks while driving. Secondary tasks range from conversing with passengers and drinking and eating to phoning, even reading and writing messages. Engaging in secondary tasks means that drivers' attention is diverted leading to degrading driving performance and resulting in serious concerns for road traffic safety. According to NHTSA, 25% of all crashes are due to driver inattention. Half of those were caused by driver distraction (Ranney et al., 2000). Driving performance is affected differently depending on the type of secondary task involvement. According to NHTSA, driver distraction can be divided into four distinct categories (Ranney et al., 2000):

1. Visual distraction (e.g. looking away from the roadway)
2. Auditory distraction (e.g. listening to music)
3. Biomechanical distraction (e.g. manually adjusting in-vehicle controls such as the radio or the A/C). Such tasks are sometimes termed "visual-manual".
4. Cognitive distraction (e.g. daydreaming or being lost in thoughts)

However, many other researchers classify auditory distraction, including hands-free mobile phone conversations, within the cognitive category.

Most activities drivers engage in involve more than one type of distraction. A driver may type a text message while driving. In addition to being visually distracted, the driver is also biomechanically (from this point forward referred to motor) and visually distracted. Some tasks related to safe vehicle operation also have characteristics of secondary tasks, for example, activating the windscreen wipers or adjusting climate controls. The willingness to engage in secondary tasks depends in part on the imposed workload of the task. Secondary tasks may be scheduled in moments of low task demands resulting from the driving task (Schömig et al., 2011; Lerner & Boyd, 2005).

The complexity of a driving situation determines the experienced workload of drivers. Workload is comprised of stress and strain. Stress represents the task demand, while strain constitutes the resulting impact upon the individual. Therefore, the stressor (i.e. task demand) is objectively the same for everyone, but how the stressor is perceived by the individual depends on the individual's coping skills, current state, and other constitutions and predispositions.

According to Schießl (2009), in the context of driving, three sources contributing to workload are identified:

1. Driving manoeuvre
2. Environmental factors
3. Secondary task

Driving manoeuvres differ in their degree of difficulty. While following another vehicle mainly requires controlling the vehicle longitudinally (i.e. distance to the vehicle and speed), overtaking another vehicle on a single carriageway requires additional resources. In addition to controlling the vehicle laterally, relevant information needs to be perceived and processed before deciding on and initiating the takeover manoeuvre. Environmental factors also moderate the degree of perceived workload. Environmental factors include, for example, infrastructural aspects such as the road type or the presence of junctions/intersections or density of traffic. Weather conditions, such as extreme heat, rain, fog also affect workload. Secondary tasks, such as phoning, eating or talking to a passenger, are also sources of workload.

If driving complexity affects perceived workload, then workload from driving can be determined based on the type of manoeuvre driven as well as environmental factors. In naturalistic driving studies, directly measuring workload (i.e. physiological indicators) is not possible. Collecting subjective ratings before, during, and after each drive is also not suitable for naturalistic driving studies. Therefore, workload has to be indirectly measured based on the driving manoeuvre and environmental/situational factors. Based on Nagel (1994), Schießl (2009) identified 15 driving manoeuvres and defined them based on driving performance parameters. Of these 15 manoeuvres, ten were identified in the present set of data (see Table 3-8).

**Table 3-8: Overview of defined driving manoeuvres**

| | |
|---|---|
| 1. Stopping | 6. Approaching intersection |
| 2. Following a road | 7. Crossing intersection |
| 3. Approaching an object | 8. Turning left or right |
| 4. Stopping behind an obstacle | 9. Changing lane |
| 5. Following a vehicle | 10. Parking |

These driving manoeuvres contribute to the driving task difficulty. According to Fuller (2005), drivers actively try to keep the driving task difficulty within manageable boundaries. Drivers may try to reduce task difficulty when driving manoeuvre demands are high. This may be achieved by stopping the performance of secondary tasks or by only choosing less complex secondary tasks.

The definition and categorization of the driving manoeuvres are not independent. The new driving manoeuvre depends on the previous manoeuvre; therefore, manoeuvre detection is done sequentially. This approach has two advantages (Schießl, 2009):

1. Only a limited number of manoeuvres can follow a previous manoeuvre limiting the search area
2. Describing the changes in parameters based on the previous manoeuvre should lead to a more accurate detection as manoeuvres can be better differentiated

In order to detect manoeuvres correctly and efficiently, parameters describing a driving manoeuvre were defined followed by defining changes in parameters used to differentiate between possible follow-up alternatives. For example (as illustrated in ), a driver's manoeuvre is *starting/continuing driving*. Based on the starting manoeuvre, a handful of follow-up manoeuvres are possible. If the car travels at a constant speed and velocity is greater than 20 km/h, the following detected manoeuvre is *following a road* (Schießl, 2009).

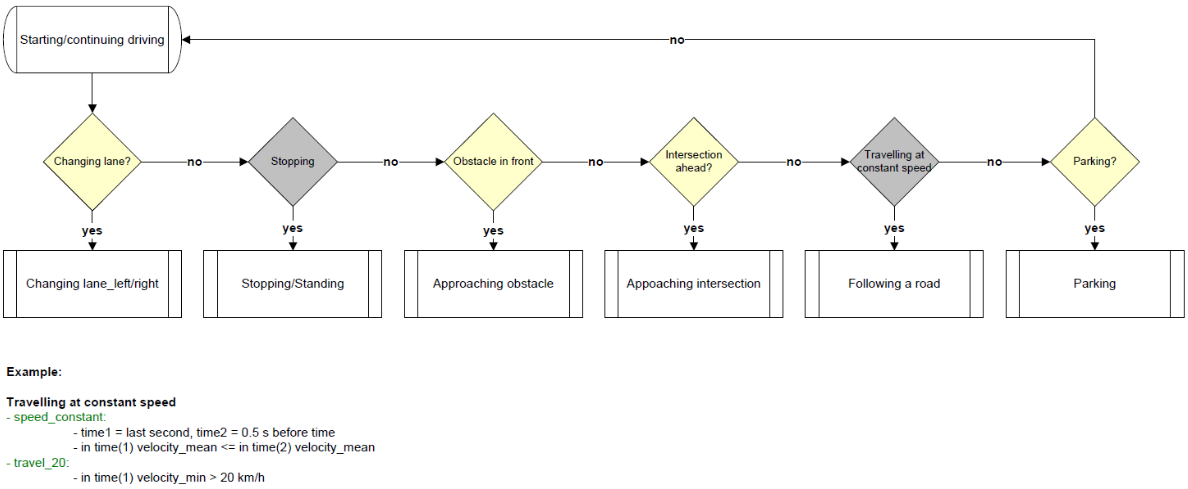



**Figure 3-15: Example of driving manoeuvre detection**

Actions in grey diamond shapes are easily quantified based on CAN-Bus data, while actions in yellow diamond shapes require more information. Map data is needed to identify intersections; GPS data is needed to know the exact position of the vehicle. Distance or time to intersection can be derived from the information. In order to identify vehicles in front of the host vehicle, radar or laser scanners are needed. Within the project UDRIVE, variable such as TTC and THW are derived from MobilEye data.

Environmental factors affect workload (i.e. task complexity). Traffic density, road type and associated legal speed limits, weather conditions, and road surface conditions are considered moderating factors.

Here the focus is on whether and how driving task complexity and secondary task complexity influence the decision to engage in secondary tasks. For this purpose, annotated trips including secondary task engagement are selected for the analysis. For each of the secondary task involvements, the driving manoeuvre or the sequence of driving manoeuvres is determined, and depending on the number of actions required for the manoeuvre, the complexity is determined. Secondary task complexity depends on the type of secondary task as well as the combination of type of distraction; therefore, a visual-motor task is more complex than an auditory task. Based on previous findings, we hypothesize that:

1. The more complex the driving task, the less complex is the secondary task and vice versa.

2. The more complex the driving task, the shorter is the duration of the secondary task and vice versa.

### 3.5.2 Method

**Participants**

Data used for the analysis is based on the data query of May 2nd, 2017. Altogether, 898 secondary task events were identified. The sample of drivers is comprised of 44 male between the ages of 21 and 77 years (M = 48.55, SD = 14.37) and 40 female drivers between 20 and 66 years (M = 40.4, SD = 10.32).

**Procedure for identifying driving manoeuvres**

All of the 898 segments of secondary task engagements are used. Within the secondary task, the driving manoeuvre or the sequence of driving manoeuvres is determined based on the definitions provided by Schießl (2009). In order to determine manoeuvres at the beginning of the secondary task as well as the end of the secondary task as accurately as possible, 30 seconds before the beginning of the secondary task and 10 seconds after the secondary task are included in the determination of the manoeuvre.

**Variables**

In order to determine the influence of driving and secondary task complexity on the engagement in secondary tasks, ten driving manoeuvres were defined. The complexity of the driving manoeuvre depends on the modalities (manual, visual, auditory) that are required to complete a manoeuvre; therefore, the higher the number of modalities, the more complex is the driving task.

Driving task complexity is further moderated by environmental factors such as road type, weather condition, and road surface condition. Map-matching data provided information on road type. Information on weather condition and road surface condition were annotated.

In a second step the complexity of the environment was categorized. For weather and road surface each adverse condition was seen as increasing the driving complexity by 1 (Table 3-9). The environment complexity was calculated as the sum of weather, road surface condition and road type (range 0–4).

**Table 3-9: Classification of complexity of environment (x <= 2: low; x > 2: high. / x = sum of 3 dimensions of environment per secondary task)**

| Environment | Type | Code |
|---|---|---|
| Weather | Mist light rain | 1 |
| | Raining | 1 |
| | No adverse conditions | 0 |
| | Snowing | 1 |
| Road surface condition | Sand oil dirt | 1 |
| | Dry | 0 |
| | Snowy | 1 |
| | Wet | 1 |
| Road type | Country motorway | 0 |
| | Rural | 1 |
| | Slip road | 1 |
| | Undefined | 0 |

| | | |
|---|---|---|
| | Urban motorway | 0 |
| | Urban road (city) | 2 |

For road type, different complexities were used. Driving in the city usually is much more complex than driving on a rural road or on a highway. In cities, the road traffic system is used by different road users. The degree of traffic regulation differs depending on the location; therefore, roads with high traffic density are strongly regulated. Intersections, for example, require paying attention to multiple directions in a short time frame to avoid a collision with pedestrians, cyclists, PTWs, or other vehicles. Thus, the complexity of urban roads was rated as high (+2). Highways, on the other hand, are arguably the safest roads. Even though the travelling speed is high, lanes are wide and physical barriers in place separating travel directions. Crowding is reduced by having multiple lanes in one direction reducing drivers' effort. Accordingly, highways were rated as having a low complexity (0). Rural roads are in between these extremes. They usually have fewer intersections than cities and less road users crossing the street. However, lanes are not divided and relatively narrow resulting in small safety margins between vehicles. They were rated as having a complexity of medium (+1; see Table 3-9).

**Table 3-10: Classification of complexity of secondary task (x = 1: easy; x = 2: medium; x = 3: difficult)**

| Type | Manual | Visual | Auditory | Code |
|---|---|---|---|---|
| Electronic device | X | x | | 2 |
| Food | X | | | 1 |
| Grooming | X | x | | 2 |
| Phone | X | x | X | 3 |
| Read or write | X | x | | 2 |
| Smoking | X | | | 1 |
| Talking or singing | | | x | 1 |

Secondary tasks were categorized based on the type of distraction induced. Eating, for example, is a motor task, while texting is a visual-cognitive-motor task. Based on the categorization complexity of secondary task was determined (Table 3-10). The lane change manoeuvre was included in the analysis, but no instances of drivers changing lanes and engaging in a secondary task were found. Thus, it was not rated for involved modalities. For each secondary task up to a maximum of five driving manoeuvres during the duration of the secondary task were identified. The complexity of the driving manoeuvre for a given secondary task was calculated by finding the sum of all observed driving manoeuvres (Table 3-11).

**Table 3-11: Classification of complexity of driving manoeuvre (x < 4: easy; 4 <= x <= 9: medium; x > 9: high. / x = sum of 5 manoeuvres per secondary task)**

| Manoeuvre type | Manual | Visual | Auditory | Code |
|---|---|---|---|---|
| Approaching intersection | x | x | | 2 |
| Approaching object | | x | | 1 |
| Crossing intersection | x | x | x | 3 |
| Following road | | x | | 1 |
| Following vehicle | | x | | 1 |
| Parking | x | x | x | 3 |
| Stopping behind obstacle | | x | | 1 |
| Stopping | | x | | 1 |
| Turning left or right | x | x | | 2 |
| Changing lanes | *Not observed during secondary task* | | | |
| Undefined | | | | 0 |

The distribution of complexity categorization showed that most identified secondary tasks (97%) were performed while driving in easy environment conditions (Table 3-12). The distribution of the complexity of the manoeuvre was more diverse. Most secondary tasks were observed under manoeuvres of medium complexity (54%) followed by the high complexity (26%). Secondary tasks were mostly categorized as easy (45%) followed by difficult (32%) and then medium (23%) complexity (Table 3-12).

**Table 3-12: Frequencies of observed secondary tasks of different complexity categories.**

| | Complexity of environment | | Complexity of manoeuvre | | | Complexity of secondary task | | |
|---|---|---|---|---|---|---|---|---|
| | Low | High | Low | Medium | High | Easy | Medium | Difficult |
| Observed secondary tasks | 874 | 24 | 179 | 488 | 231 | 404 | 207 | 287 |

### 3.5.3 Results

**Distributions of environmental levels during secondary task performance**

The total time driven engaged in secondary tasks was 110236.1 s. When drivers performed secondary tasks, the road surface was dry in over 98% of the times (Figure 3-16). A wet road surface was only observed in about 2%. Snowy and other road conditions influencing driving quality (Sand, oil, dirt) were observed in less than 1% during secondary task engagements.

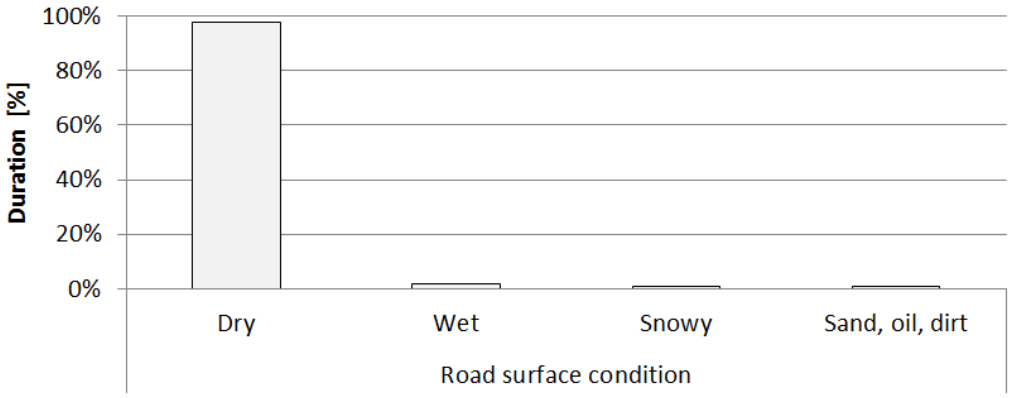



**Figure 3-16: Distribution of time driven with specific road surface conditions during secondary task performance**

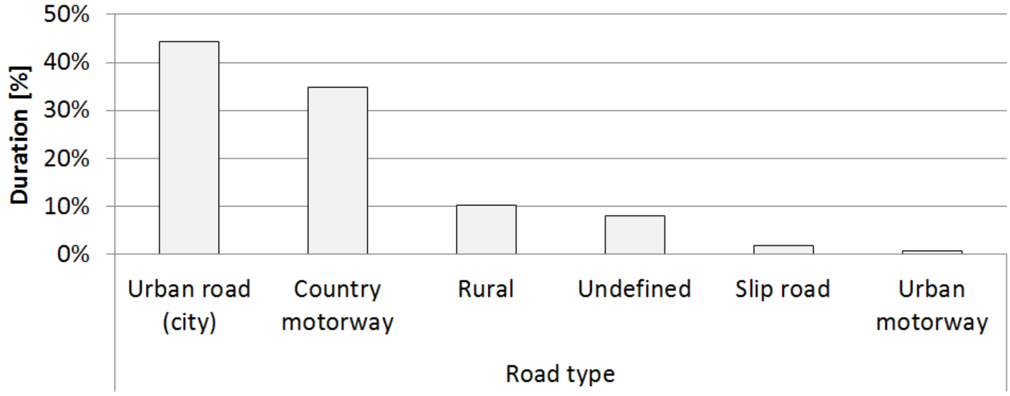



**Figure 3-17: Distribution of time driven on specific road types during secondary task performance**

The distribution of road type shows drivers performed secondary tasks mostly while travelling on urban roads (44% regular urban streets, 1% urban motorway). The second most frequent road type was country motorway driven on in about 35% of the times. Little secondary task activity was observed while driving on rural roads (10%) and slip roads (2%; see Figure 3-17).

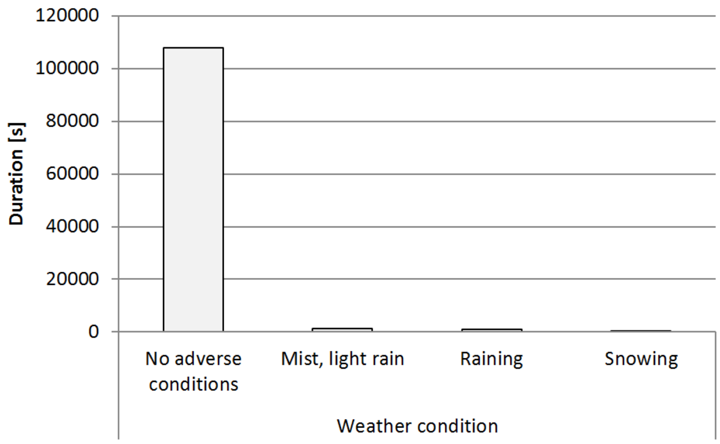



**Figure 3-18: Distribution of weather conditions during secondary task performance**

With regard to the weather conditions, drivers performed secondary tasks mostly in non- adverse conditions (97%, see Figure 3-18).

Driving manoeuvres during performance of secondary task engagements were more equally distributed. Following a road, approaching an intersection, following a vehicle, and crossing an intersection were all prominent in the data and performed about 15% of the times. Turning (8%), stopping behind another vehicle (6%), just stopping (3%), and approaching an object (2%) occurred relatively rarely. The remaining 19% of secondary tasks were performed during a driving episode that could not be related to one of the investigated driving manoeuvres (Figure 3-19).

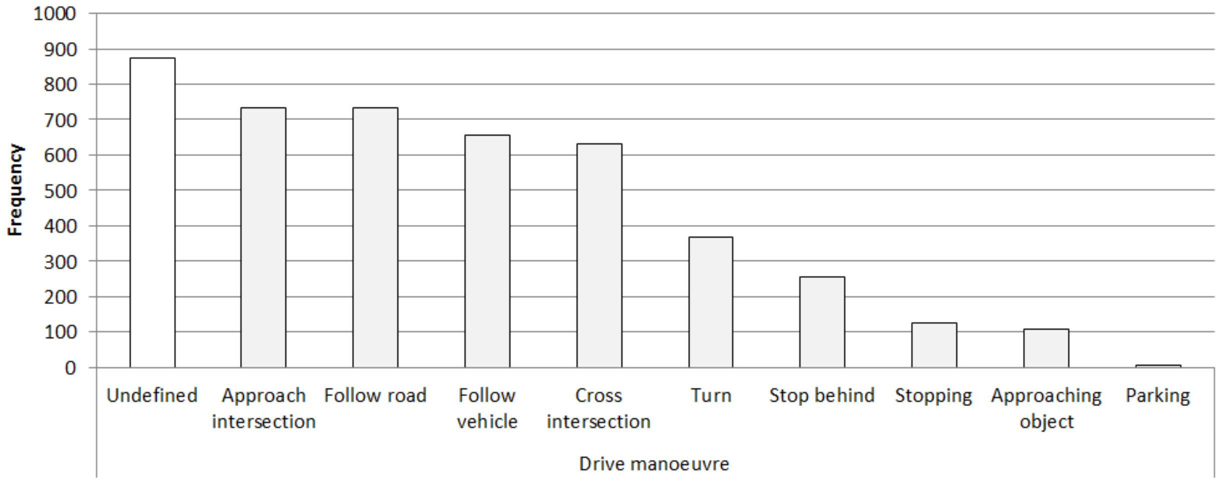



**Figure 3-19: Distribution of driving manoeuvres during secondary task performance**

**Duration of the secondary task in different environments**

The mean duration of secondary task execution was observed in relation to road surface, road type, and weather. With regard to weather, the longest secondary tasks were performed during mist or light rain (M= 174 s) followed by performance without any adverse weather conditions (123 s). During rain and snow, drivers engaged in secondary tasks the shortest (Figure 3-20). Depending on the road surface condition, drivers' secondary tasks engagement was longest on dry roads (M= 125 s) and shorter in more adverse conditions (M= 51-74 s, Figure 3-20). With regard to road type, secondary task engagement was longer on country motorways (M= 156 s) and normal urban roads (M= 111 s) compared to rural roads (M= 103 s), slip roads (M= 90 s), and urban motorways (M= 62 s).

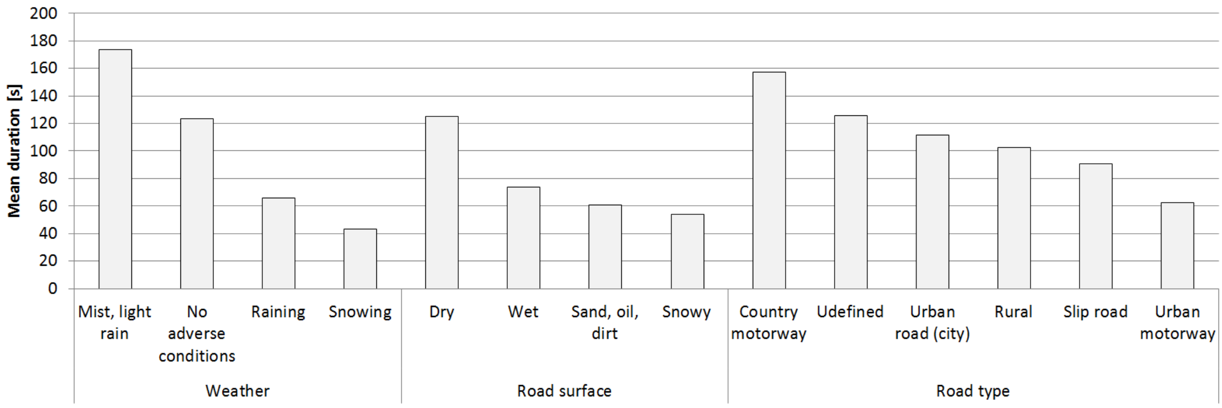



**Figure 3-20: Mean duration of secondary task performance in relation to environmental variables**

**Statistical analysis**

It was hypothesized that a more complex driving task results in engaging in a less complex secondary task. To investigate this hypothesis, the driving task was split up based on the complexity of the environment and the driving manoeuvres. While driving manoeuvres are mostly externally dictated, secondary task engagement is a conscious decision of the driver. Following the driving task difficulty model by Fuller (2005) it would be expected that drivers perform less complex secondary tasks in complex driving conditions. To investigate the differences in the distribution of secondary task complexity under the easy and difficult environment conditions, the relative frequency per environment level was calculated. This also helped to overcome the large differences in available cases per environment complexity level.

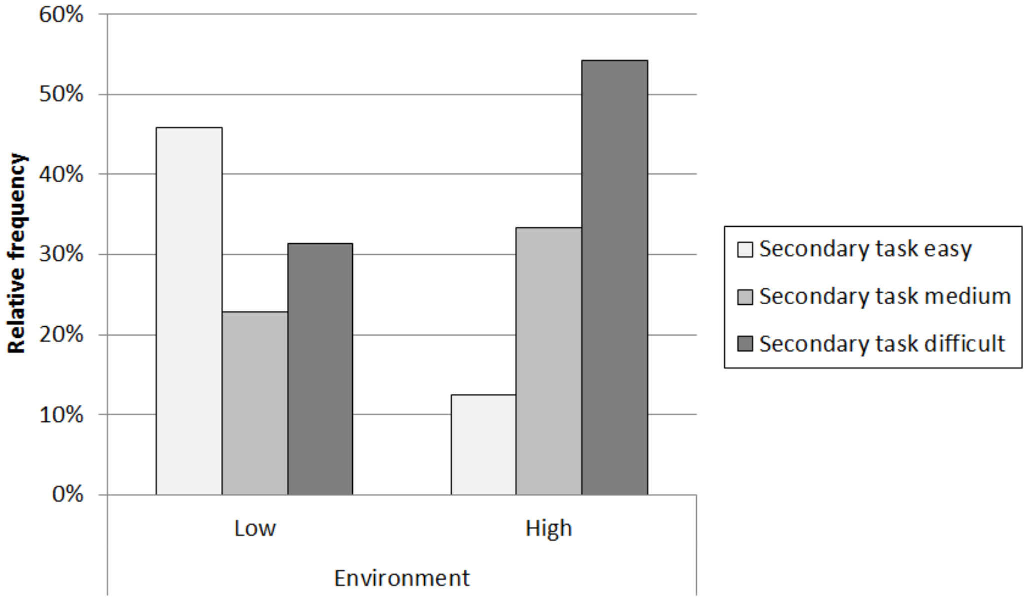



**Figure 3-21: Relative frequency of secondary tasks split by environment complexity levels**

The data showed that under difficult environmental conditions a greater share of demanding secondary tasks were performed. The differences under easy environment conditions were smaller and the share of easy secondary tasks was higher than medium or difficult tasks (Figure 3-21). A Chi$^2$ test revealed a significant association between the complexity of the environment and the complexity of the secondary task, $\chi^2(2) = 10.72$, $p = 0.005$, indicating that the factors are not independent of each other.

The same comparison was done between secondary task and driving manoeuvre complexity. Small differences between the two factors (Figure 3-22) were observed. Independent of manoeuvre complexity, easy secondary tasks were always the most frequent (42-47%) followed by difficult secondary tasks (30-33%) and then the ones with medium difficulty (22-26%). A Chi$^2$ test revealed no significant association between the complexity of the driving manoeuvre and the complexity of the secondary task, $\chi^2(4) = 1.54$, $p = 0.819$, indicating independence of the factors.

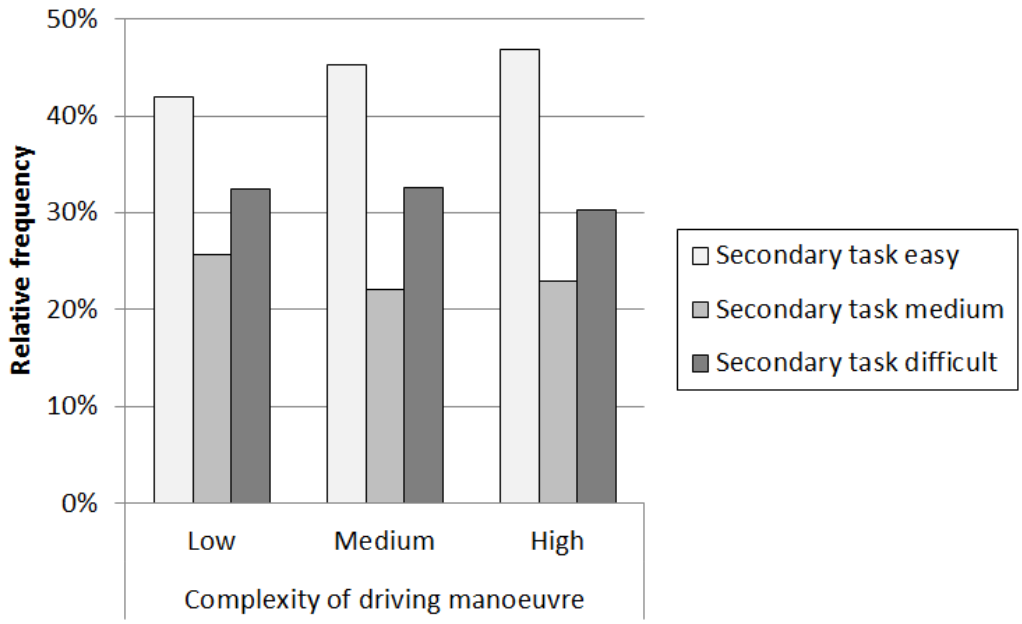



**Figure 3-22: Relative frequencies of secondary tasks under different driving manoeuvres**

It was also expected that a more complex driving task coincides with shorter secondary task duration and vice versa. This hypothesis was investigated by comparing the duration of the secondary task between the complexity levels of the environment and the driving manoeuvre.

The data showed that more than 90% of the secondary tasks were performed with an environment complexity of less than 2 (low). To make a direct comparison possible, we choose an equal number of drivers, who only drove in a low (or high) environment complexity and built a new table of independent groups. Both groups were compared using a Wilcoxon-Rank sum test. The duration of the secondary task under low environment complexity (M = 127, Md = 58.3) did not differ significantly from the conditions of an environment with a high complexity (M = 103, Md = 64.5), *W=54785, p = .726* (Figure 3-23).

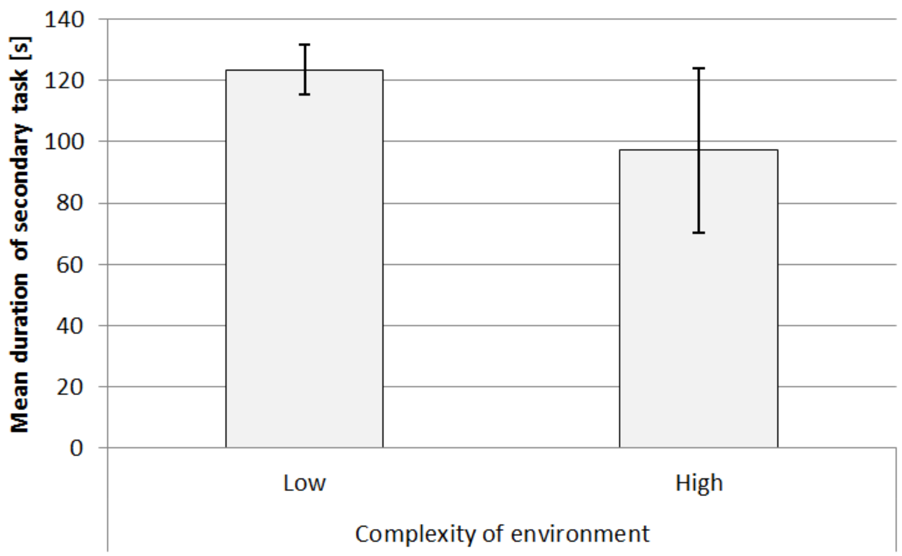



**Figure 3-23: Mean duration of secondary task under complexity levels of the environment. Error bars indicate the standard error.**

Close inspection of the data showed that not every driver experienced all three complexity levels of the driving manoeuvre (low, medium, and high). To be able to compare the three levels, drivers without exposure to all three levels were excluded from analysis. A homogeneity test indicated that the variances between the levels were not equal. Thus, a Friedman-Test was utilised to compare secondary task duration of all three complexity levels of the driving manoeuvre. The duration of the secondary task differed significantly between three complexity levels, $\chi^2(2)=7.95$, $p = .019$ (Figure 3-24). The duration of the secondary task was longer in the low complexity driving condition (M = 271.5, Md = 86.9) then in medium (M = 88.5, Md = 58.9) or high complexity driving condition (M = 79.3, Md = 51.8).

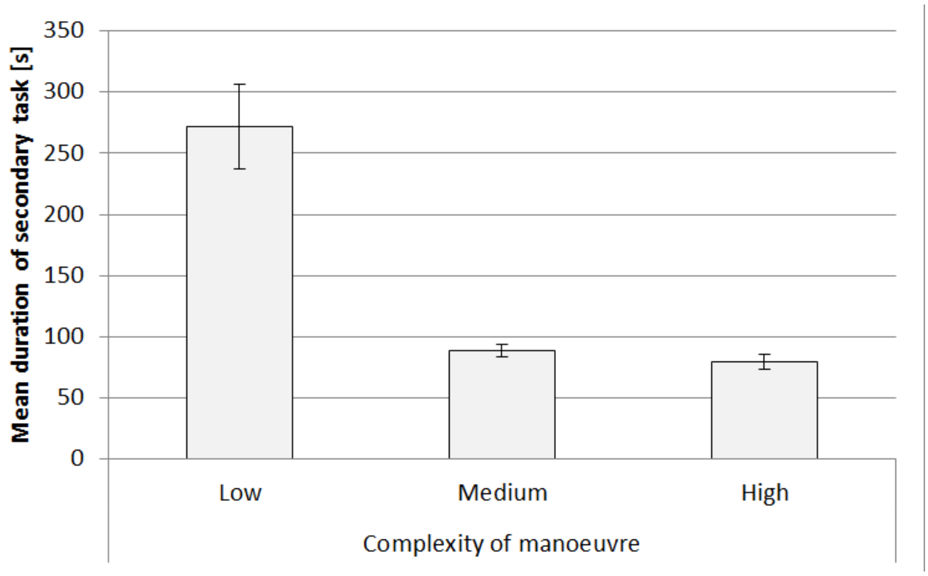



**Figure 3-24: Mean duration of secondary task under complexity levels of the driving manoeuvre (error bars indicate the standard error)**

### 3.5.4 Discussion

This analysis focused on investigating driver engagement in secondary tasks in relation to the complexity of the driving task. The complexity of the driving task was split up into the complexity of the environment and of the driving manoeuvre.

Drivers appear to perform fewer easy and more medium and difficult secondary tasks in environments with a high complexity. Regarding the secondary task engagements in relation to the complexity of the driving manoeuvre, no indication for dependence was found. Drivers performed all secondary task complexity levels independently of the driving manoeuvre complexity. Together with the results of environment complexity, the hypothesis that less complex secondary tasks coincide with more complex driving tasks cannot be confirmed. While driving manoeuvre did not affect engaging in secondary tasks, environment influenced the engagement contrary to the hypothesis.

This is unexpected since drivers are generally able to schedule demanding secondary tasks to moments of low task complexity (Schömig et al., 2011; Lerner and Boyd, 2005). Several explanations are possible. One issue is that Schömig et al. (2011) collected their data in a driving simulator. The type of secondary tasks performed and their complexity may be perceived differently in the simulator compared to real world driving. On the other hand, in experimental studies, the secondary tasks are chosen by the experimenter, for example, by level of complexity; while, in NDS, drivers decide on what type of secondary task to engage in while driving. Level of complexity may be perceived differently by experimenters and drivers. Complexity in the simulator might be moderated by unfamiliarity of the environment, the simulator, and the driving task; while, in NDS, most trips are commutes on familiar routes. Further, in the study by Lerner and Boyd (2005) drivers' engagement in secondary tasks was not observed, but drivers were asked to report their willingness to engage in them. Taking into account the probability of a response bias, subjectively reported engagements may differ from secondary task involvement observed in naturalistic driving studies. This may explain the different results.

Another explanation may be the used definition and categorization of the complexity level does not reflect the perceived complexity level. As described by Fuller (2005) drivers adapt their behaviour to the perceived driving task difficulty. The complexity definition used for the analysis was done by researchers and could not be validated without violating the NDS principal to minimize the interaction with the drivers. In future work, the categorization needs to be validated and, if necessary, refined and results recalculated with the existing data.

Finally, it has to be considered, that the results presented here were calculated based on the data available by 2 May 2017 representing about 50% of the collected data. Analysis of the entire dataset may lead to different results. Thus the analysis should be repeated when the whole dataset is available to verify the effects.

The second part of the analysis focused on the hypothesis that a more complex driving task coincides with shorter secondary task duration. It is relevant to note that the intention of the investigated hypothesis was to find out whether drivers adapt their behaviour according to the experienced situation and maintain a safe driving style. The majority of drivers performed secondary tasks only in non-adverse road surface conditions. This is evidence that drivers were aware of the driving environment. Results of secondary task engagement in adverse weather conditions indicate a small tendency in accordance with the hypothesis, even though the analysis did not reveal a significant effect between complexity levels of the environment. However, this tendency is supported by the results found in the complexity of the driving manoeuvre. Drivers spent significantly less time engaged in secondary tasks during medium and difficult manoeuvres compared to easy ones. This does seem to indicate that drivers are aware of the danger that accompanies secondary task execution under challenging driving conditions.

In summary, results indicate that drivers adapt their behaviour depending on the driving conditions with regard to the duration of secondary task performance but not to secondary task complexity.

# 4 Truck driver engagement in secondary tasks

In this chapter, analysis of truck drivers' engagements in secondary tasks while in their trucks is reported. Four research questions are addressed, each reported in a separate subsection.

Engaging in secondary tasks while driving can be considered normal from a truck driver's perspective, and it is partly related to activities meant to facilitate work-related duties. For example, a driver can engage in calls towards customers in order to contact them just before a delivery, can use electronic devices (such as a satnav) to easily reach a destination, or can look at the delivery journal to plan for the next destination. Examples of other activities, not strictly work-related, are eating/drinking, smoking, talking/singing, and personal grooming. All the above-mentioned activities are summarized in Figure 4-1, showing the total time spent by all the drivers in such activities (both as a percentage of total time and in terms of the total number of hours per task). The global number of hours of driving analysed was 83. The total time spent in all the secondary tasks sums up to about 20% of the total driving time that has been annotated. Therefore, as an example, it is possible to deduce that the time spent in phone-related activities is about 5% of the total annotated driving time.

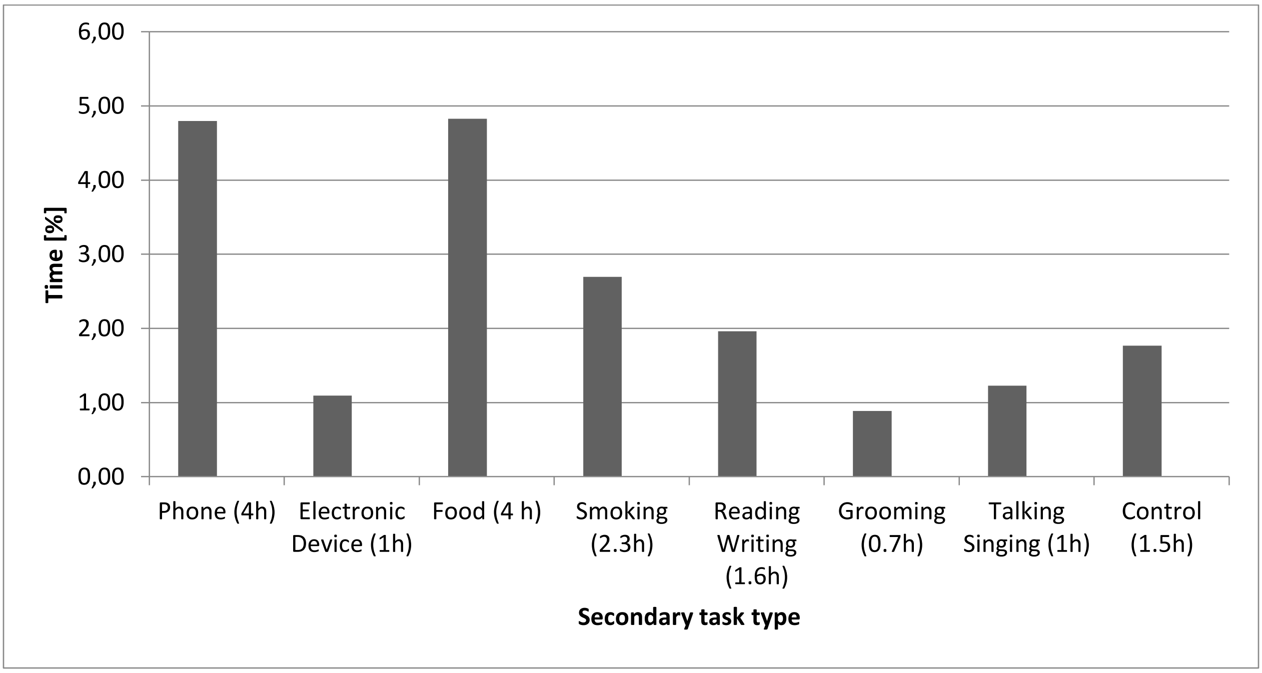



**Figure 4-1: Percentage of the total driving time spent on tasks**

The UDRIVE data is unique since it provides a large amount of naturalistic driving data that can be exploited for very different analysis. The data is of three main categories:

- CAN bus data
- Videos
- Self-report questionnaires

The combinations of these three elements enables analysis which cannot be performed on many other existing datasets. Moreover, the subjects were observed in their natural behaviour, during everyday driving, without any manipulation by the observer which could potentially bias or influence the collected data.

The UDRIVE truck dataset is composed of 44 drivers (of whom 24 were annotated), distributed over four truck fleets. The driver with the most available collected data had, at the time of writing this report, 1024 hours of driving data available in the UDRIVE database. All the fleets drove only in The Netherlands and delivered goods to customers both in urban and non-urban areas. A total of 1445 secondary tasks were annotated for trucks in UDRIVE, with a total duration of 83 hours from 264 annotated trips. Figure 4-2 shows the distribution of trip durations across the 83 hours and 264 annotated trips. For a description of the annotation procedure, see section 2.3, page 5. In short, an initial manual annotation of driver video included the marking of the point in time each task started and when it ended – see UDRIVE Deliverable D41.1. The duration between these two times is considered to be the total task time (TTT). Additional annotations was performed to break down the phone tasks that had been identified in the first round of annotation, into sub-tasks, and annotation was also performed to identify intersections and other contextual variables.

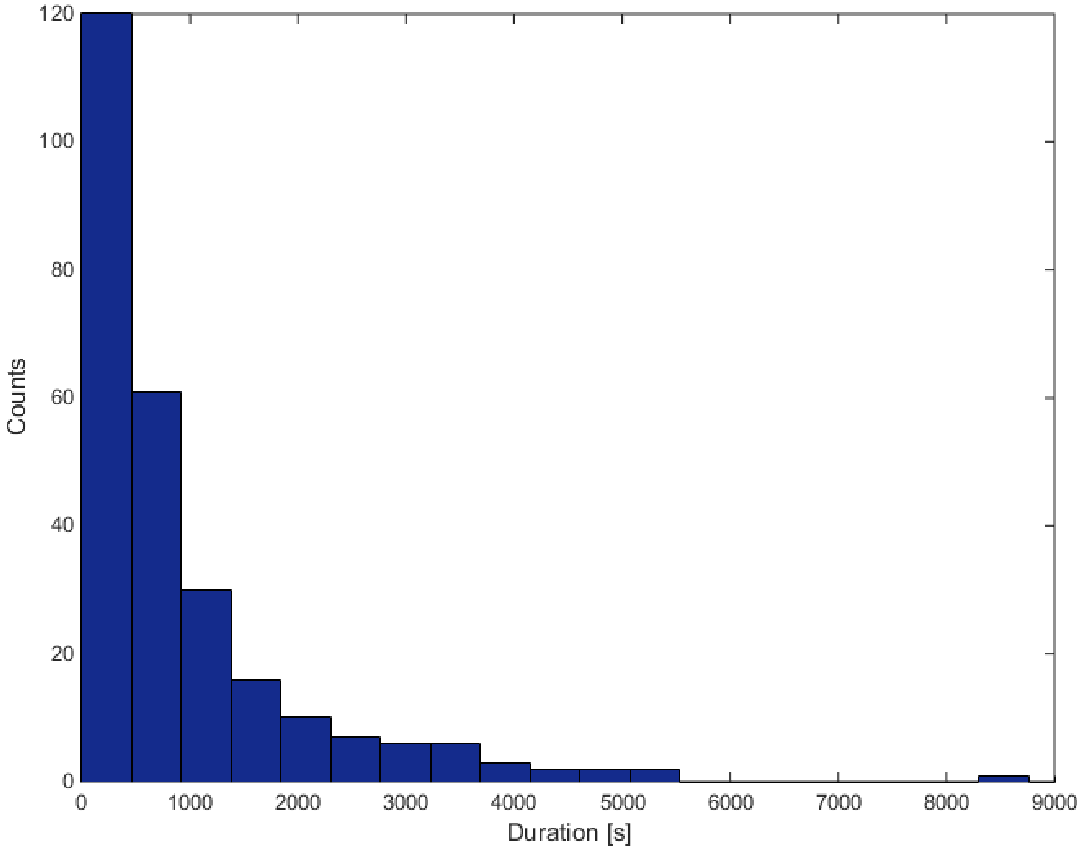



**Figure 4-2: Histogram of the durations of the annotated trips across the 83 hours of the total trip duration**

*Note: To make the figure easier to read, trips longer than 3 hours (three trips in total of 3.5, 5.7 and 12.8 hours) have been removed.*

## 4.1 What was the prevalence of secondary task activity?

This research question relates to how often drivers initiate a variety of secondary tasks while in their vehicle and how long those secondary tasks last.

### 4.1.1 Total task time by task type

Figure 4-3 shows a bar plot of the average TTT across all drivers for each task. Note that to avoid drivers with many tasks influencing the average disproportionately (bias results) the mean TTT was first calculated for each driver, after which the overall average was calculated. To evaluate the effect of this imbalance, Figure 4-3 shows both the average of averages (bias removed), and also the simple overall average task duration (average across all tasks). Inspection of the data revealed that there was one very long-lasting occurrence of a smoking task (30 minutes in total) which led to the large difference in duration of the smoking task between the two types of average that are shown in Figure 4-3. Also note that only one-third (eight) of the drivers smoked at all while driving. Thus the total number of smoking instances was low, even if the average duration was long.

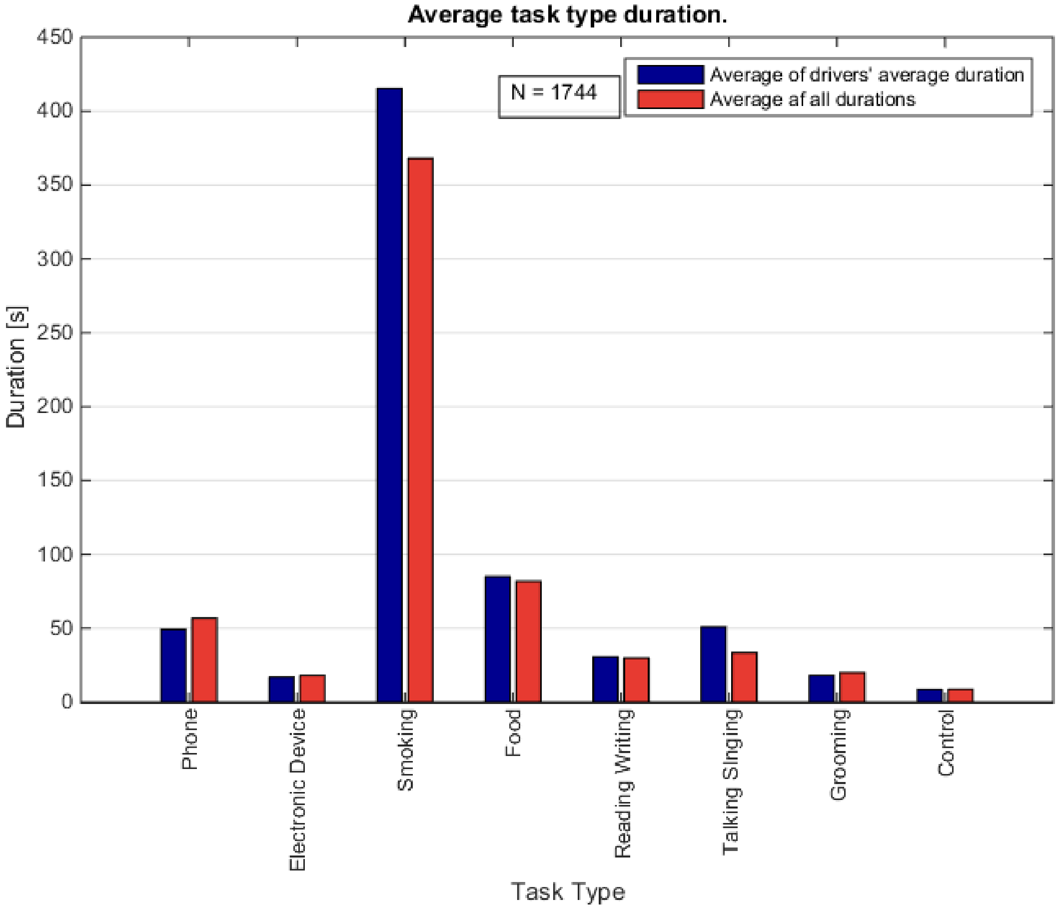



**Figure 4-3: Average task duration by task**

Figure 4-4 through Figure 4-8 show box plots of the individual drivers TTT for a variety of secondary tasks. One figure per task provides a quick overview of the variability between and within drivers. Note that the driver numbers in the plots are not the DriverID in the original dataset, but a random ID assigned to each driver.

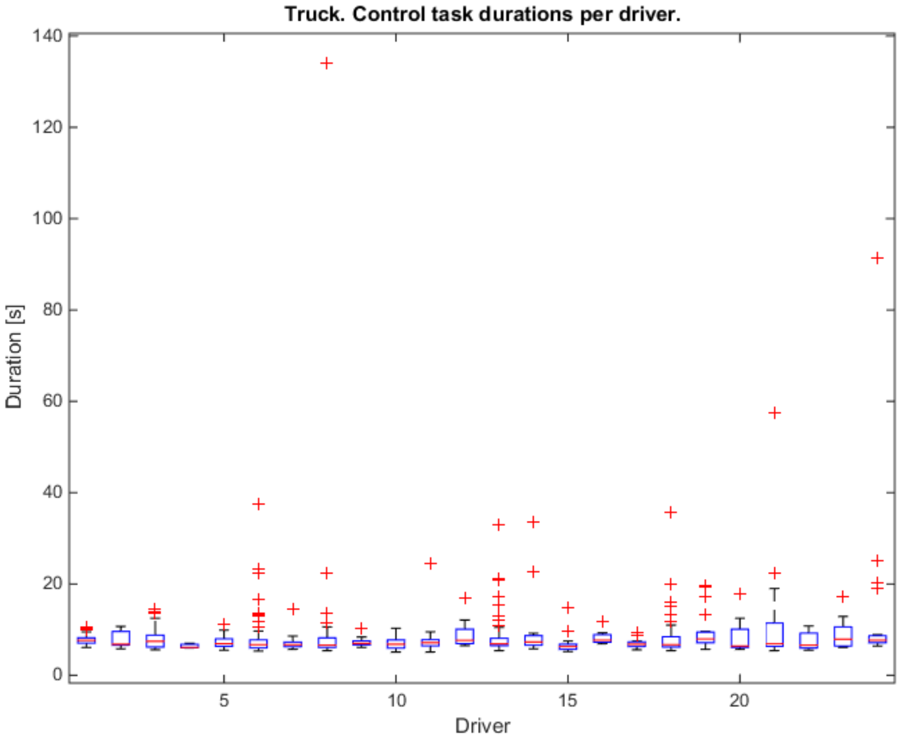



**Figure 4-4: Box plots of the duration of the control task (typically interacting with the truck buttons and levers)**[1]

[1]The plot shows medians (red horizontal line), 25th/75th percentiles (box limits), max/min excluding outliers (whiskers), and outliers (+). For definitions of outliers, see Matlab™ R2015b documentation.

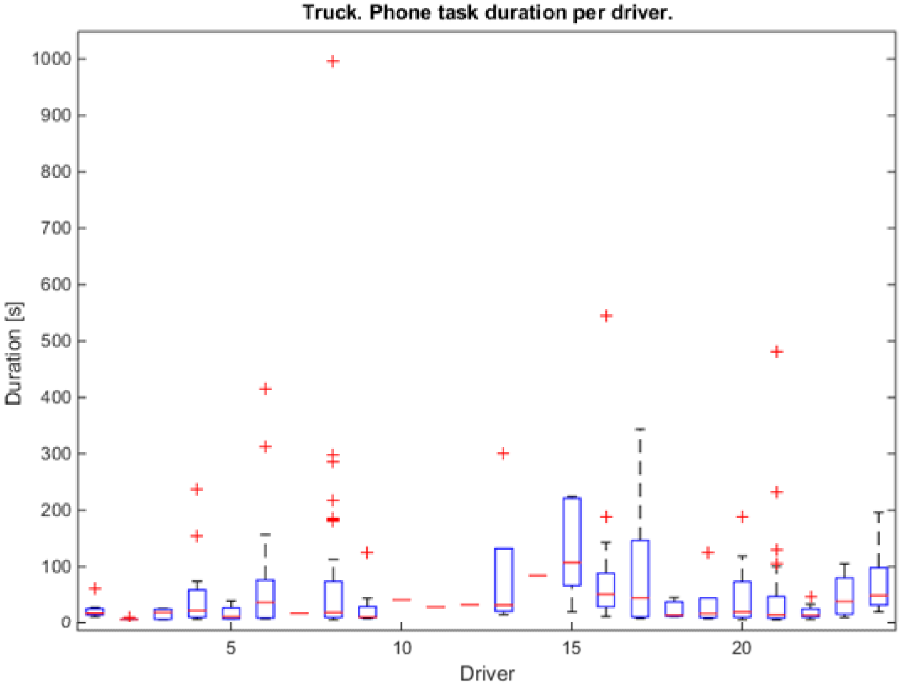



**Figure 4-5: Box plots of the duration of the phone task**

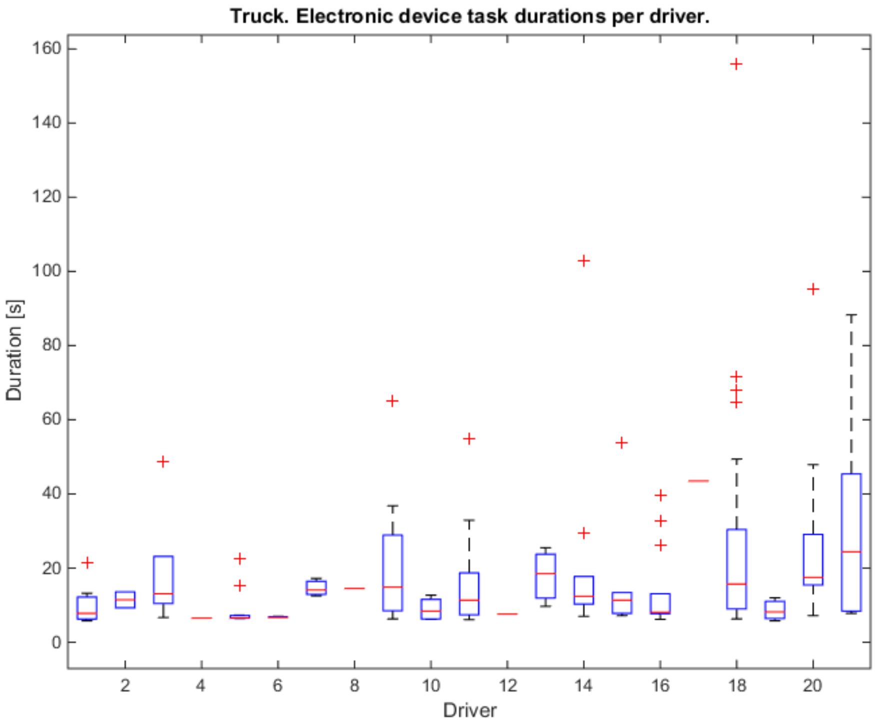



**Figure 4-6: Box plots of the duration of the electronic device task**

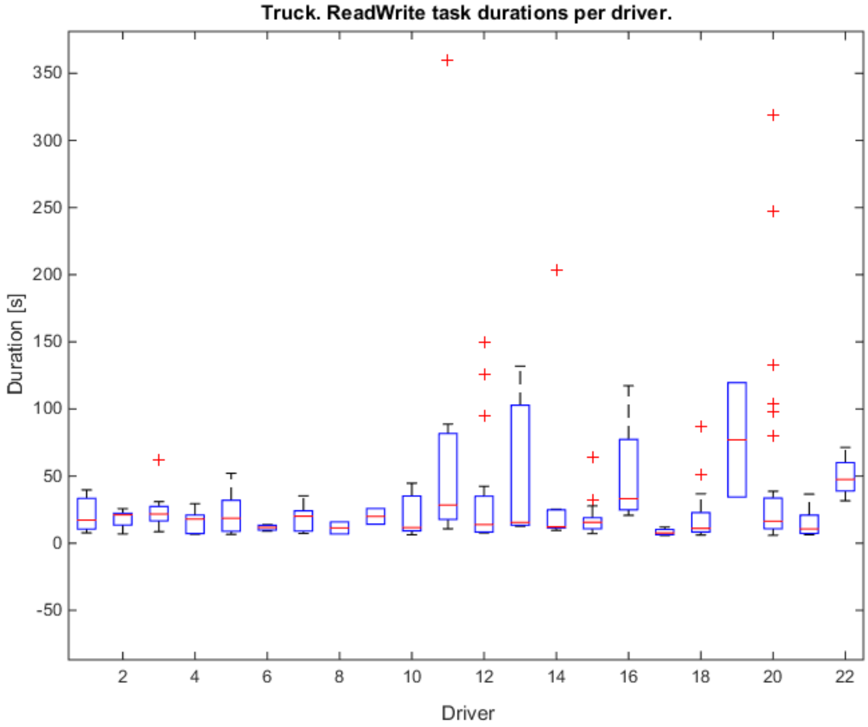



**Figure 4-7: Box plots of the duration of the reading/writing task**

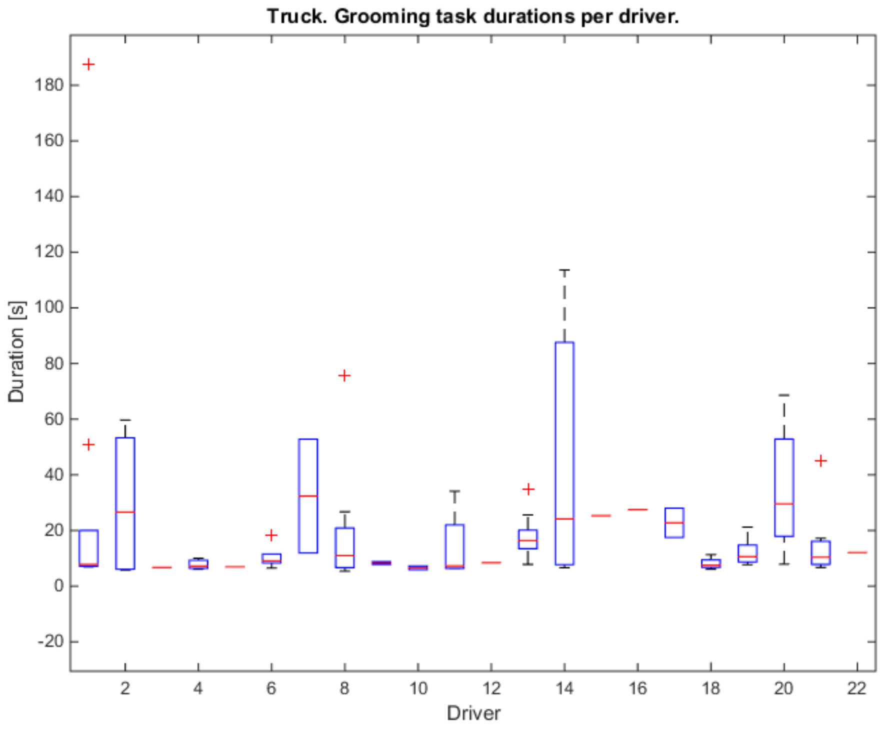



**Figure 4-8: Box plots of the duration of the grooming task**

A manual inspection of Figure 4-4 through Figure 4-8 indicates the following:

- Reading/Writing: this task shows relatively low variability between the drivers. The longest task was 320s long, the shortest 6s long.
- Phone Tasks: these show a relatively low variability between the drivers but a fairly large variability within driver. The longest task was 1000s long, the shortest 5s long.
- Grooming: these show a considerable variability in the task duration both across the different drivers and within the same driver. The longest task lasted 185 seconds, while the shortest was 10 seconds.
- Electronic device: these show a considerable variability in the task duration both across the different drivers and within the same driver. The longest task lasted 160 seconds, while the shortest was 5.5 seconds.
- Control: these show a relatively low variability in the task duration both across the different drivers and within the same driver. They are generally very short compared to the other categories. The longest task lasted 140 seconds, while the shortest was 5 seconds.

### 4.1.2 Phone use breakdown in comparison with other secondary tasks

Many studies of phone tasks are categorized into text messaging (SMS) and talking (Dingus, 2014; Dingus et al., 2016, Victor et al., 2015), with the latter typically being divided into hands-free and handheld talking. In this study we did not have a distinct category *texting*, as it was difficult to distinguish texting from other interactions (other types of writing interactions with the cell phone), and we wanted to have a term that could be used across both phone and electronic device task visual/manual interaction. We used the term interaction. For all definitions of tasks, see Appendix A in UDRIVE Deliverable D41.1. Figure 4-10 and Figure 4-11 show the distribution of the engagement in phone-related secondary tasks, divided in its sub-categories. The most frequent task was handheld interaction which accounted for about 2% of the total driving time annotated and for about 35% of the phone sub-tasks. Note that hands-free talking was hard to distinguish from the talking/singing category in the manual annotation. However, the annotators separated these categories to the best of their ability, i.e. coded the task as hands-free when they could observe reliably that the driver was making a phone call.

There were a total of 226 phone tasks in this study. An example of task durations for one sub-category — the 96 interactions with handheld phone annotated in the project — is shown in Figure 4-9. Few tasks are very long, and the vast majority of such tasks are shorter than 250 s.

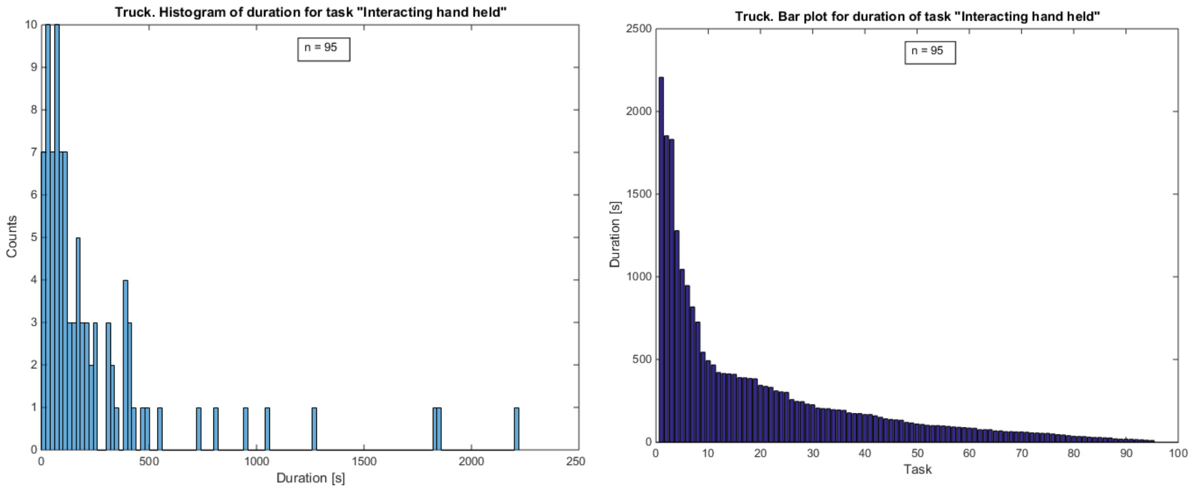



**Figure 4-9: Histogram of all "interaction handheld" tasks in the dataset (left), the same data sorted from longest to shortest TTT duration (right)**

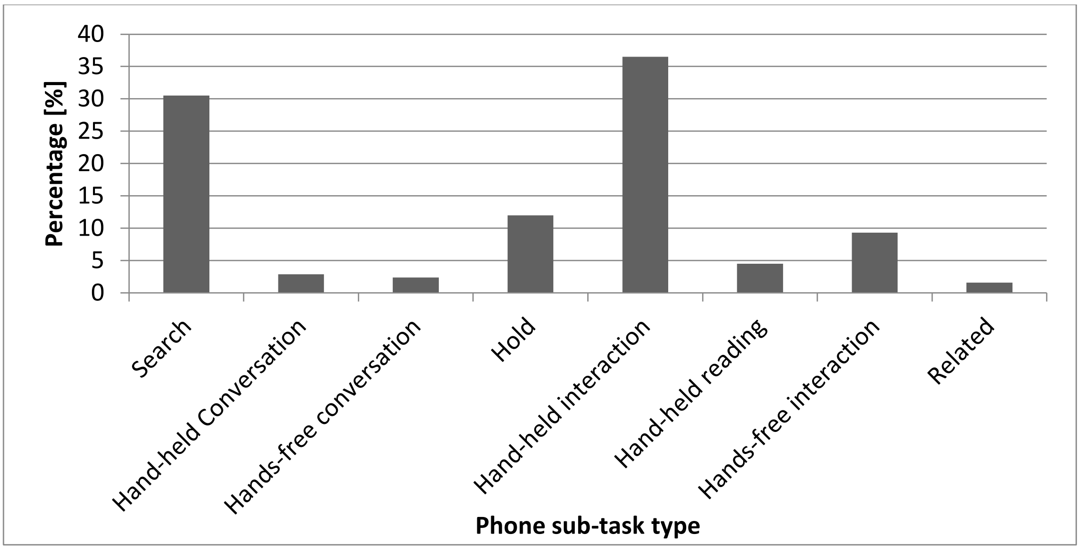



**Figure 4-10: Distribution of the sub-phone tasks**

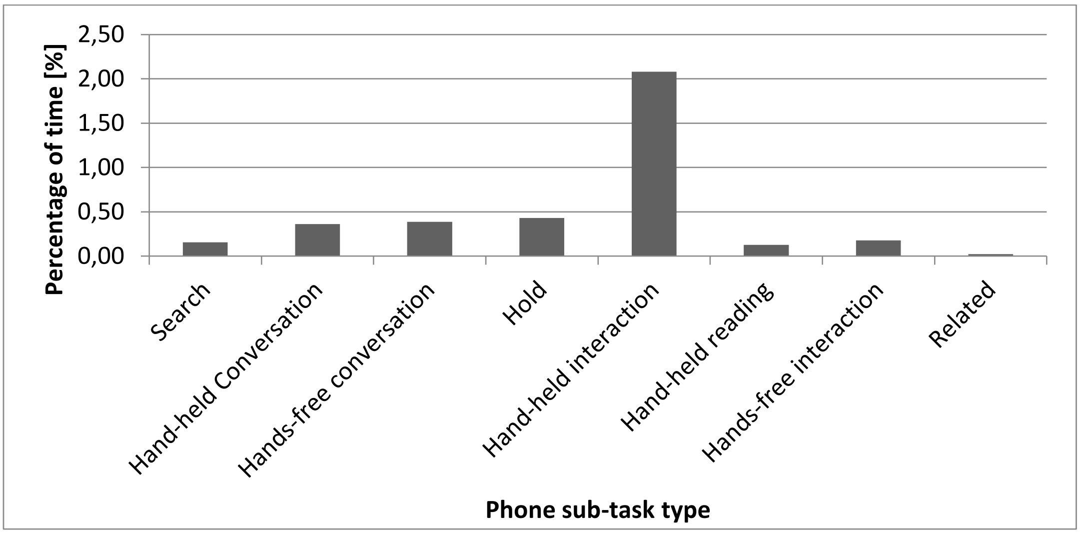



**Figure 4-11: Percentage of engagement time by type of sub-phone tasks**

An overview of all annotated task-types in UDRIVE, visualised together with the four phone sub-categories, is shown in Figure 4-12. Smoking has by far the longest average duration, while the control tasks (e.g. button presses) are on average short but by far the most frequent. The phone task (any phone task) is the most frequent task after the control task. However, it should be noted that reading/writing occurs more than twice every hour on average, and has a duration on average of approximately 20 s. Note also that initiating both handheld and hands-free phone calls was relatively rare in this dataset, compared to interacting with the phone, where handheld interaction was the most common task after the control (e.g. button presses) task.

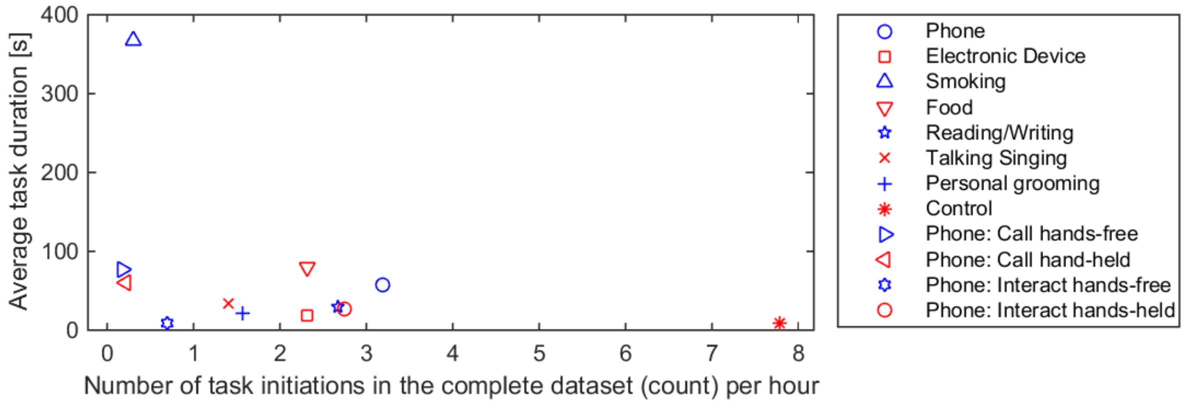



**Figure 4-12: Task duration versus task frequency**

*Note: For task duration, the mean was calculated for each driver and then the mean of means was calculated.*

## 4.2 Does driving task complexity influence the decision to engage in secondary tasks?

This section addresses the question of the relationship between driving complexity. Analysis include studying a) the speeds when initiating secondary tasks compared to speed of all driving, b) when, in relation to the end of trips, drivers initiated phone calls, c) when drivers initiated phone tasks in relation to exiting intersections, d) the frequency of tasks initiated in intersections compared to the frequency of phone tasks initiated outside (before/after) intersections, and e) how/if drivers choose to do different secondary tasks in parallel.

### 4.2.1 The speeds when initiating secondary tasks compared to speed of all driving

Figure 4-13 compares the speeds at phone task initiation to the speed distribution across all annotated trips. This can be seen as a representation of when drivers feel more comfortable engaging in secondary tasks, compared to everyday driving (not performing a task).

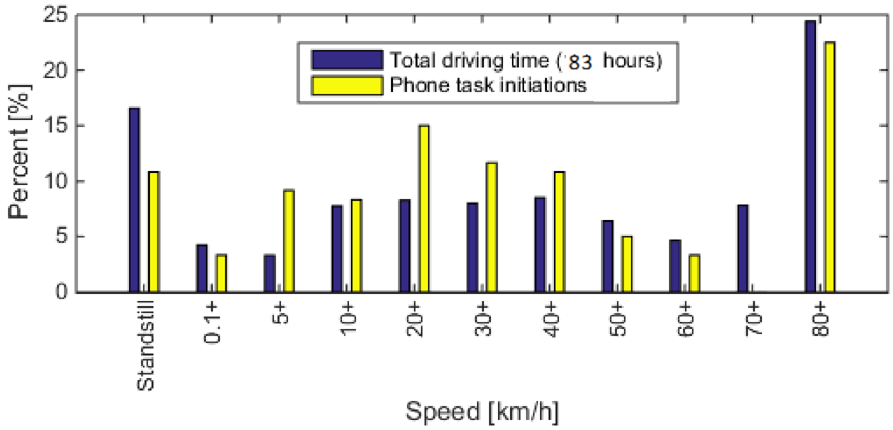



**Figure 4-13: The speed at initiation of phone tasks (all forms) compared to the speed distribution across all data used in the analysis**

Figure 4-13 shows that the majority of the phone tasks were initiated at low speeds (at or below 30 km/h) or when standing still. The speed range with the highest frequency of phone task initiations (but also most time spend driving) is 80 km/h and above. This may be when the environment is less complex and the driver even may have support systems, such as cruise control, engaged. Drivers in this study initiated phone tasks less often at standstill and at 70 km/h, and more often at 5 and 20 km/h, compared to everyday (no-phone-task) driving. Note the large difference in the proportion of phone task initiations at standstill for cars and trucks. Cars drivers were at standstill in 56% (section 5.1.2) of visual-manual tasks, which are the majority of all truck tasks (see Figure 4-11). It is further noteworthy that in a study by Tivesten and Dozza (2015), the standstill frequency of phone tasks was also significantly higher than in everyday driving, while in this study it was less frequent for truck drivers. The Tivesten and Dozza study only included physical phone interactions, but the proportion of pure voice interactions for truck drivers in the current study was small. Thus the studies should be comparable.

### 4.2.2 Timing between phone task initiation and the end of the trip

An analysis was made to study when, in relation to the end of a trip, drivers typically initiate a phone task. It is not obvious which metric to use to demonstrate this, but in this study, we chose to plot the trip duration versus the time when a phone task was initiated (coupled to each trip; see Figure 4-14) and calculate the proportion of task initiations in each 14.3 degree slice (1/7 = 14.3 degrees per slice). The results show that drivers tend to initiate phone tasks earlier (rather than later) in trips (approximately 44% of the phone task initiations were in the first 28.6% of the trip), and not so much in the middle of trips. This may be as a result of the drivers contacting the delivery location (saying they are on their way and possibly asking for delivery plan confirmation), although the initial hypothesis was that that would have make drivers initiate tasks closer to the end of the trip.

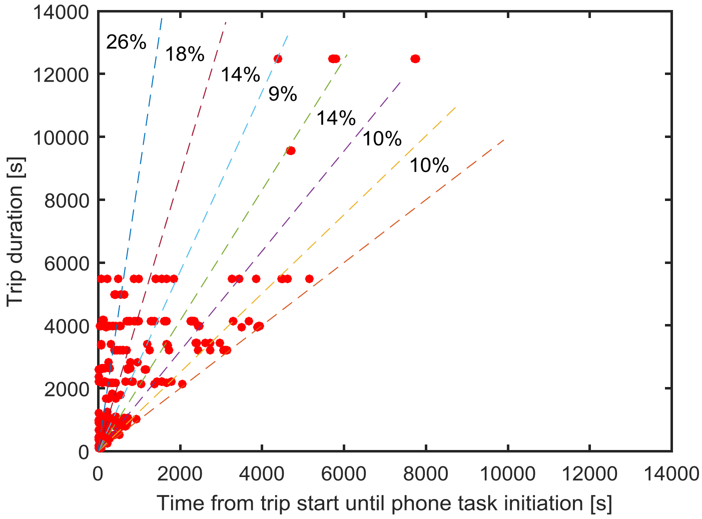



**Figure 4-14: Duration of individual trips vs. the time from start of the individual) trip until the initiation of a phone task**

*Note: The percentages in each slice shows the proportion of phone tasks initiations in that slice, out of the total.*

### 4.2.3 When do drivers initiate phone tasks in relation to exiting intersections?

The study by Tivesten and Dozza (2015) indicate that car drivers in Sweden wait until they have just passed an intersection before they initiate visual-manual phone interaction. In this study we did a similar analysis, but used all phone tasks (and not only visual-manual) as our sample size did not allow for such stratification. We calculated the time from the exiting an intersection until the start of a phone task. If no phone task was initiated before the next intersection started, the former end was consequently ignored. Thus, only durations from the exiting of an intersection until the start of a phone task, without any intermediate intersection starts, were considered in the analysis. Figure 4-15 illustrates this definition. Figure 4-16 then shows a histogram of durations from exiting of an intersection until the start of a phone task (given this definition. Note, however, that the duration from exiting an intersection until initiating a phone task is affected by the distribution of gaps between intersections, as the time from the end of an intersection has an upper bound in the gap between intersections. This may have an impact on the interpretation of results. Therefore Figure 4-15 shows a histogram of the time from the end of an intersection until the start of another, as available throughout the 83 hours of data used in this analysis. Further analysis should consider methods that takes the exposure into account directly in the evaluation. However, time constraints did not permit that in our analysis.

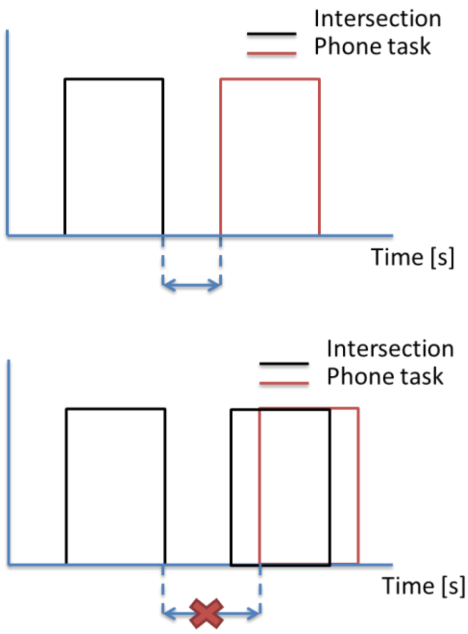
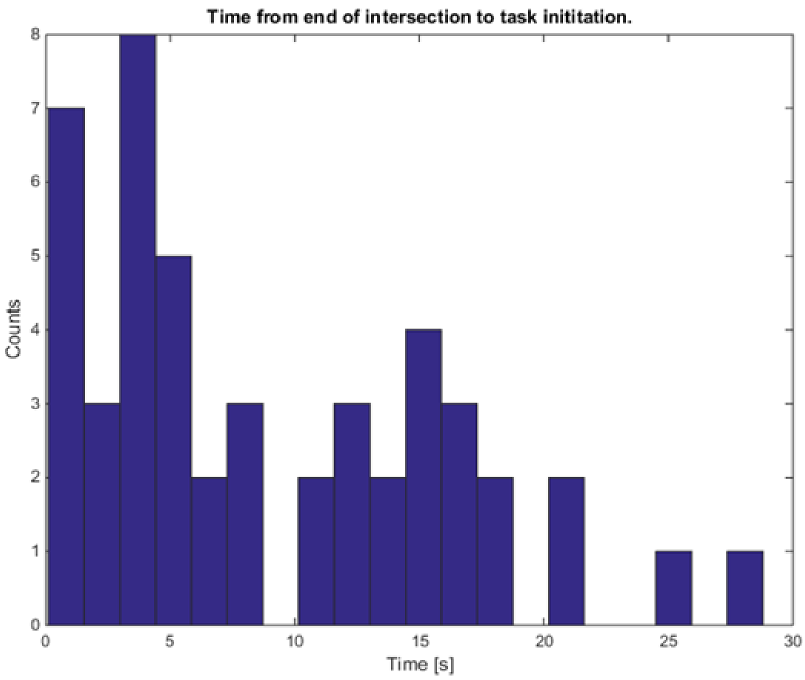



**Figure 4-15: An illustration of the definition of the time from the end of a task until the start of a phone task. The lower panel shows that if there is any subsequent intersection start between an intersection end and a phone task initiation, that measurement is not used in the analysis.**



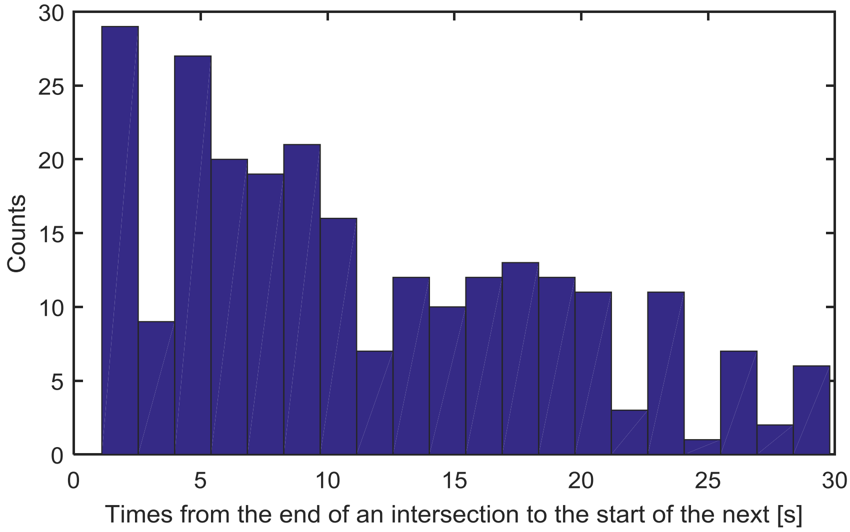



**Figure 4-16: Time from end of intersection to task initiation (top), and a histogram of durations from ends of intersections to the start of the next intersection (bottom). Both figures have the same bins sizes.**

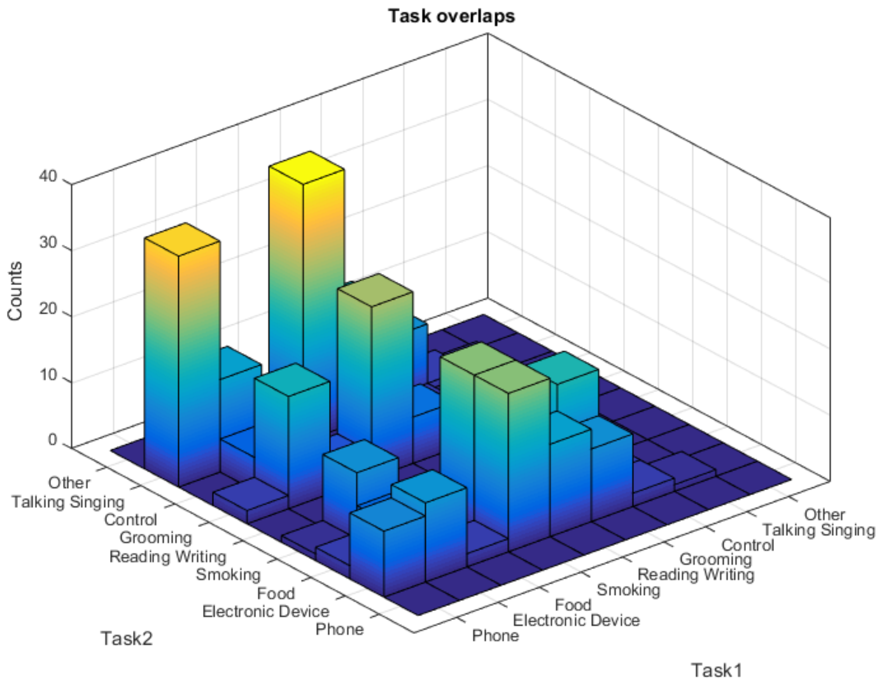
*Note: For the bottom plot. all bins to the right of 30 s are simply discarded to facilitate the comparison of the x-axes of the two plots.*

### 4.2.4 Parallel secondary tasks

Drivers engaging in parallel tasks need to task-switch not only between one task and driving, but between driving and two (or more) tasks. Here the question was how often drivers engage in parallel secondary task activities. That is, we have analysed how different secondary tasks overlap with each other in terms of actual task overlaps (from start of task to end of task).

Figure 4-17 shows the nine by nine combinations of tasks that can appear in parallel (two or more tasks at the same time) in the UDRIVE annotated data. The most common overlapping activities are talking/signing and phone task, and talking/singing and food task (eating/drinking). The former is quite natural as talking is a form of phone task, while eating/drinking while talking/singing is not as obvious.  Smoking is the task that has the most overlap with other tasks – also as expected.

It is difficult to draw firm conclusions with respect to impact on safety of the driver performing parallel secondary tasks. More studies are needed to investigate the actual safety effect of parallel tasks in different combinations.



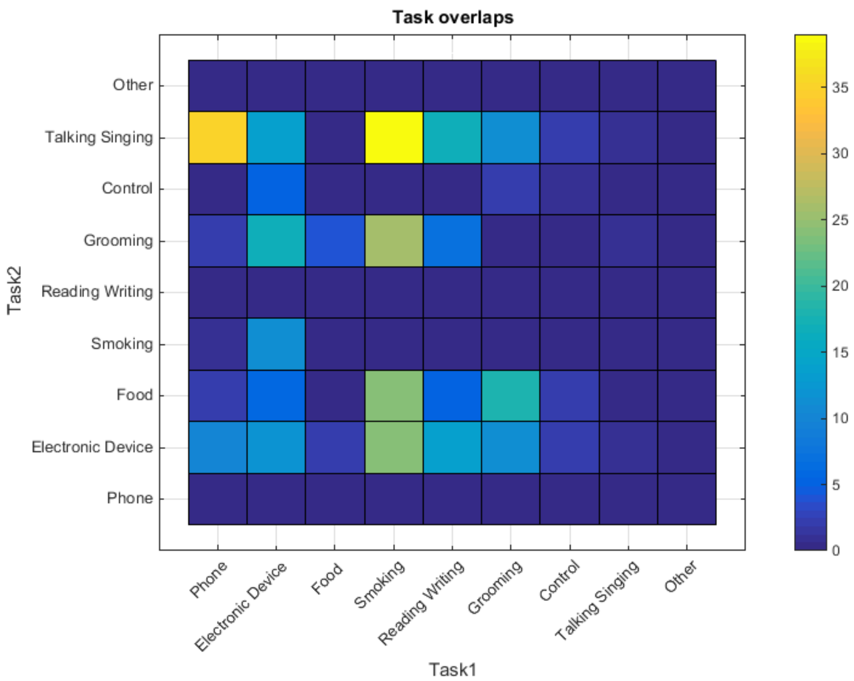




**Figure 4-17: An illustration of combinations of parallel task as annotated in the UDRIVE truck data. Top: 3D visualization; bottom: top view**

## 4.3 Conclusions

The analysis of drivers' engagement in secondary tasks using naturalistic driving data reveals habits that would not be possible to capture with other types of studies. In particular, from the analysis it emerged that drivers spend about 20% of the driving time performing secondary tasks (such as using the phone, reading, eating, smoking, etc.). The two most common types of tasks were eating/drinking (5% of the driving time) and control (5% of the driving time). Some tasks are performed more often than others and for different average durations. For example the longest task was smoking, followed by calling with the phone. Even though these tasks were the longest, they were not very common. Contrary to these, i.e. frequent but short, are the control tasks. Visual-manual interactions with a phone are one of the most frequent and long tasks, and therefore lead to a high exposure.

From the video annotations and analysis, it appears that drivers feel more comfortable in initiating a task at low speeds (below 30km/h) or at very high speeds (more than 80km/h) — the latter possibly due to the possibility of using the cruise control in free-flow traffic. However, the proportion of phone task initiations was lower at standstill than the overall proportion of standstills in the data. This is the opposite from what has been reported in previous studies on car drivers (Tivesten and Dozza, 2015). Further, tasks are more frequently initiated at the beginning of the trip (e.g., to contact the customer at the delivery point) and soon after an intersection had been passed (that is, in line with Tivesten and Dozza, 2015, drivers seems to wait to initiate phone tasks until after passing intersections).

With respect to parallel tasks, the most common parallel tasks are talking/signing tasks in conjunction with phone tasks, and talking/singing tasks together with food tasks (eating/drinking). The former is quite natural as talking is a form of phone task, while eating/drinking while talking/singing is not as obvious.  Smoking is the task that has the most overlap with other tasks, also as expected.

# 5 Do drivers adapt their safety margins before, during and after performing secondary tasks?

## 5.1 Car

The research question here is whether, in undertaking secondary tasks, drivers adapt their safety margins. Because of the interest is mobile phone use while driving, the focus is on phone-related tasks.

### 5.1.1 Method

The aim of this study is to tell if secondary task has a significant effect on driving in particular in terms of safety margin. Base on the AIDE recommendations (Ostlund, 2005), three types of safety margin were chosen: speed, position in the lane and time headway. On account of the sensors used to acquire data, specific characteristics of each of these performance indicators are described below. The 10hz time series were used.

**Speed**

This variable came directly from the CAN data. Errors in the recording can lead to infinite values which were filtered to obtain a “clean” speed signal.

**Position in the lane**

This variable was measured by the Mobileye camera. When no lane was detected, the position took a “Nan” value. As in other studies, lane position was not processed on urban roads. To study the impact of secondary task activity on the position in the lane, we chose to use Standard Deviation of the Lane Position over a time window that is described below (Ostlund, 2005).

**Time headway**

This variable was also measured by the Mobileye camera. It had a “NaN” value when no vehicle was detected in front of the car (if the preceding vehicle was too far away for example) or when the car was stopped.

For secondary tasks, the available signal in the data base was manually annotated for a random 962 selected trips. These trips were reviewed in their entirety and manual annotation provided start and end of several type of secondary task. It is widely known that different types of secondary task vary in their impact on driving (e.g. Sayer et al., 2005; therefore this study focused on phone-related activity. In particular, visual-manual tasks have been observed to have a higher risk in term of road safety (e.g. Rosenbloom, 2006). The list of visual- manual tasks is provided in Table 5-1.

**Table 5-1: List of phone related visual-manual tasks**

| Annotation title | Meaning |
|---|---|
| **P_Search P_Related** | Searching for a phone, dealing with the phone |
| **P_HHeld_Conversation** | Hand held conversation |
| **P_Hold** | Holding the phone |
| **P_HHeld_Interaction** | Interacting with the phone |
| **P_Handheld_reading** | Reading on the phone |

The first step consisted in creating the visual-manual signal by merging the corresponding manual annotations into an indicator variable. This was coded with 2 values: 0 no task, 1 visual-manual task ongoing.

In the UDRIVE data, a desynchronization was detected between the video images and the CAN signal. This discrepancy is particularly critical in our case since we are looking for a difference in driving parameters during the task. To be sure the impact of the task is included in the sequence of interest, a padding of 2s before and after the task is added to the visual-manual signal. Figure 5-1 shows the padded visual-manual signal.

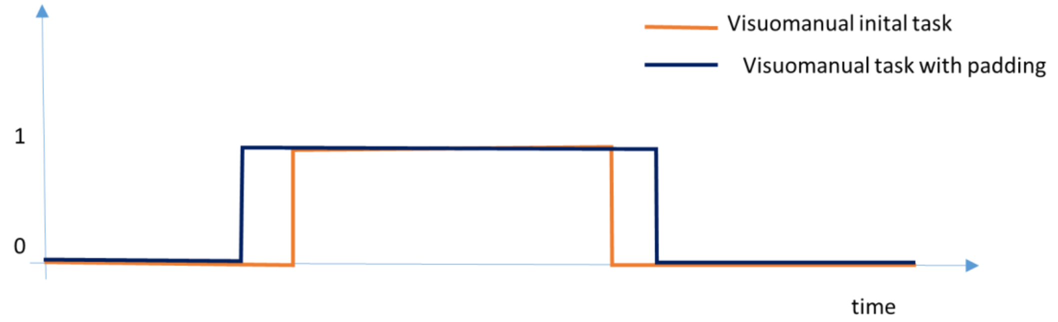



**Figure 5-1: Padding for visual-manual task**

Each task was then matched with its baseline. The baseline of a task was defined as a period of the same duration as the secondary task (padding included) and ending at the beginning of the task (Figure 5-2). When another secondary was present in the baseline window, the task was not kept for our analysis.

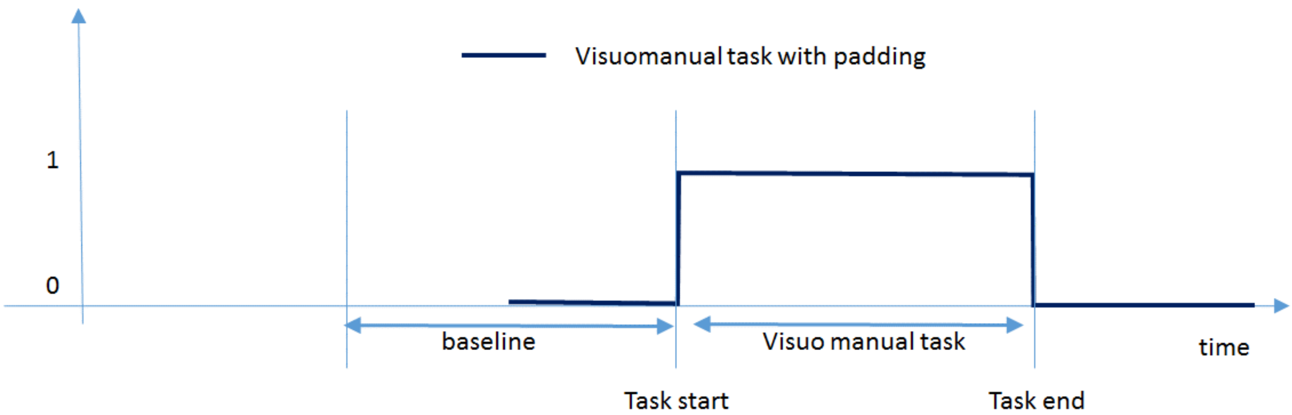



**Figure 5-2: Visual-manual task and its baseline**

Mean speed, SDLP and mean time headway were processed over these two matched time windows as well as the difference in Speed, SDLP and THW between task and baseline. For SDLP and time headway, if the variable was a NaN during more that 80% of the time window (either task or baseline), the variable mean also takes a NaN value. For the analysis, the pairs (baseline-task) with mean speeds equal to 0 were removed because in that case, safety margin is not relevant.

### 5.1.2 Results

Altogether, 291 visual-manual tasks were available from the manual annotations. After the filtering described above, 271 matched pairs were available for analysis.

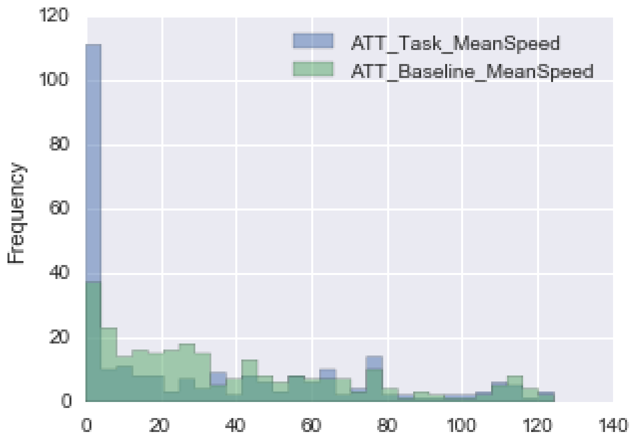



**Figure 5-3: Distribution of speed before and during secondary task**

Figure 5-3 shows the distribution of speed for task periods and baseline periods. Note that these are two separate distributions, superimposed on each other. The rather striking result is that 56% of the visual-manual phone related tasks were performed at 0 km/h speed and that compares with much less stationary or near-stationary driving in the baseline periods. This suggests that drivers are aware of the risk of undertaking such tasks while moving and try to engage in a safer context.

It is also possible to compare each secondary task episode with its matching baseline, focussing on the speed difference between a visual-manual task and its baseline. Results are discussed below.

**Impact of secondary task on speed**

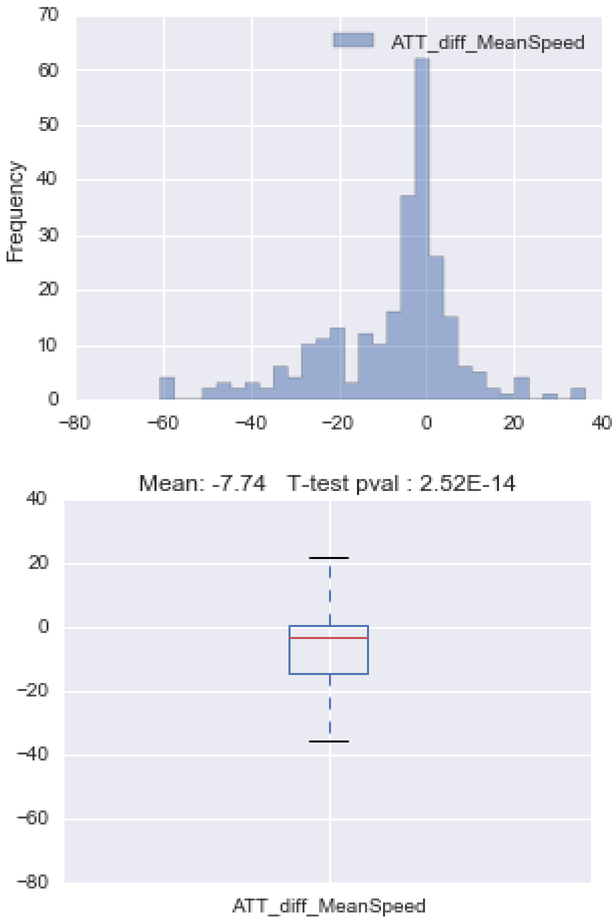



**Figure 5-4: Distribution of speed difference between visual-manual tasks and their baselines**

Figure 5-4 shows the distribution of speed difference. Driver behaviour during VM tasks was spread between accelerating and reducing speed. Nevertheless, the distribution is centred around 0 and the mean difference in speed is −7.74 km/h between baseline and paired task, indicating that drivers tended to slow down while performing the tasks.

Driver behaviour can be split on whether they accelerated (27% of the tasks) or decelerated (73% of the tasks). As regards speed reductions, the mean deceleration was around −13 km/h with a significant difference from 0 (Figure 5-5).

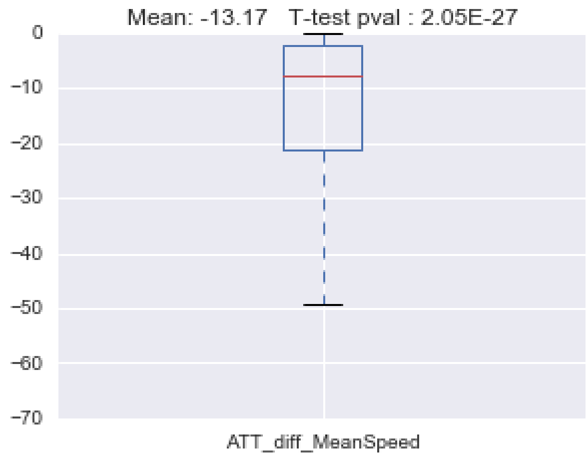



**Figure 5-5: Boxplot of difference in mean speed between task and baseline for drivers who decelerated**

It must be noted that a high proportion of drivers who decreased speed were actually stopping before beginning the secondary task. The task was not necessarily performed entirely while stopped but the mean speed is quite low during the task: mean speed during task is only around 20 km/h. Drivers thus seem to increase their safety margin while performing visual-manual secondary tasks by stopping before or during the task.

Concerning the drivers who increased their speed, the mean increase was around 7 km/h compared to the counterpart baselines (Figure 5-6).

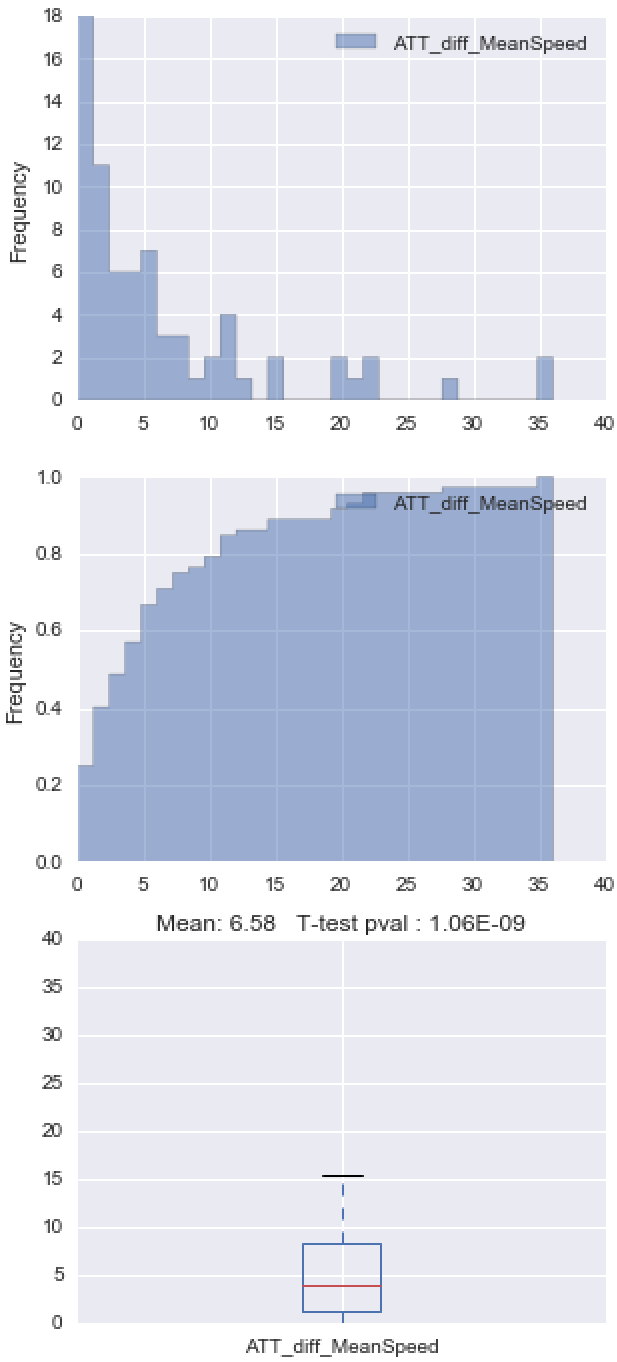



**Figure 5-6: Accelerating behaviour distribution and boxplot**

This value was lower than the mean deceleration and less frequent. Anyway, those secondary tasks represent an interesting sample to study and the reason for accelerating should be studied. One explanation might be that the task was started while stationary, and the traffic in front then moved off.

**Impact of secondary task on SDLP**

For this analysis, matched pairs where SDLP was not measured (no lines detected) for either baseline or task were removed. This left 110 events available for study .

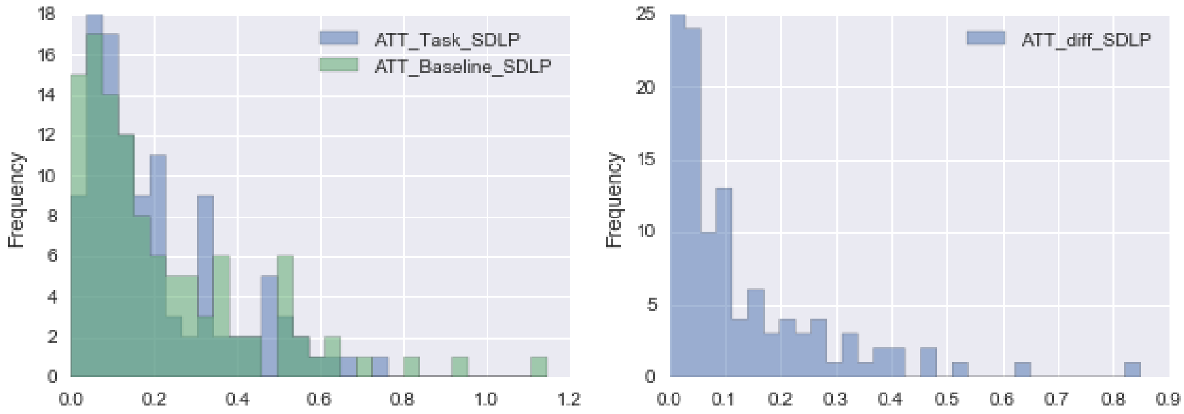



**Figure 5-7: Distributions of SDLP during VM tasks and baselines (left) and distribution of difference in SDLP between tasks and baselines (right)**


bb

**Figure 5-8: Boxplot of SDLP difference between task and baseline**

Figure 5-7 compares the distribution of SDLP across the task and baseline periods, while Figure 5-8 shows the range of differences between matched pairs of task and baseline episodes. Figure 5-8 shows a standard deviation different from 0. This deviation represents a degradation in driving and potentially a reduction in road safety. This result shows that even if drivers tended to adapt their safety margins by decreasing their speed or engaging in tasks when stopped, they were actually unable to manage the secondary task as safely as they perhaps believed.

**Impact of secondary task on THW**

This metric could be studied with only 65 samples because most of the tasks were initiated at 0 speed when no THW is available. Most of the remaining samples, did not include any preceding vehicle.

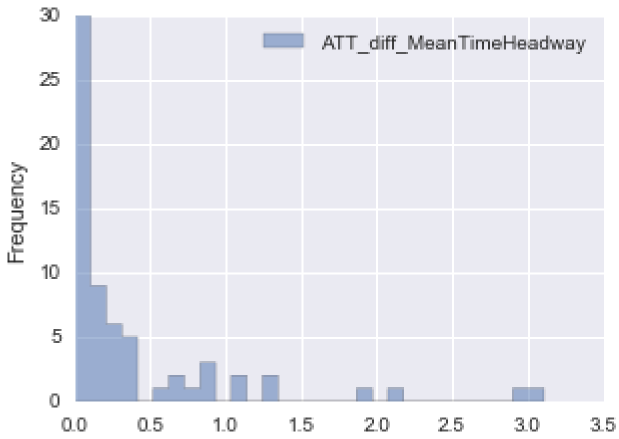



**Figure 5-9: Distribution of THW difference between task and baseline**

Figure 5-9 shows the two distributions. With such a small sample, achieving statistical significance on comparisons is difficult, but the mean difference in time headway between task and matched baseline was 0.27 sec (higher during task), as shown in Figure 5-10. This is indicative of a potential increase in safety margin.

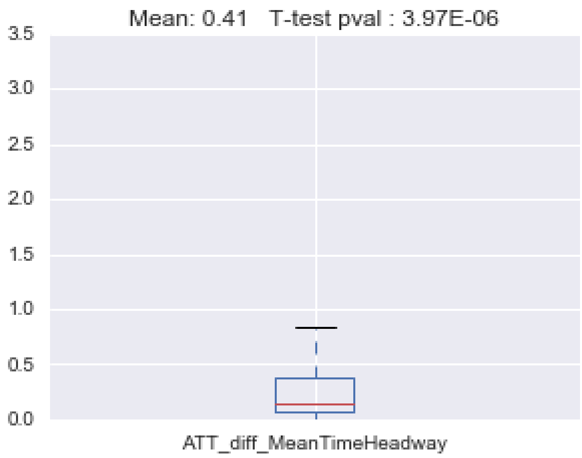



**Figure 5-10: Boxplot of THW difference between task and baseline**

Globally speaking, drivers seem to be aware of the danger of engaging in a secondary task. They tend to stop before beginning a task or they choose to engage when they are stopped. When moving, mean speed decreased during the visual-manual task by approximately 15 km/h and time headway increased. Nevertheless, SDLP shows a degradation of driving during secondary task with an increase of 0.1m.

The number of samples available for this study was quite small and more secondary tasks should be studied when available. When the number of tasks is sufficient enough, it would be interesting to study more driving parameters such as maximum and minimum speed and the kind of manoeuvres drivers initiate before engaging.

## 5.2 Truck

In this analysis a set of variables' values were extracted for seven distinct point in time. Here after these points are jointly called analysis points (APs). The APs are shown in Figure 5-11. The analysis in this section was highly inspired by Tivesten and Dozza (2015) in the choice of both the APs and the metrics for analysis. For each of the seven APs, speed and time-headway (THW) was retrieved, and the differences between the points were analysed. This analysis was done partially though simple t-tests between the different points (Figure 5-11), and partially through the use of generalized linear models (GLM).

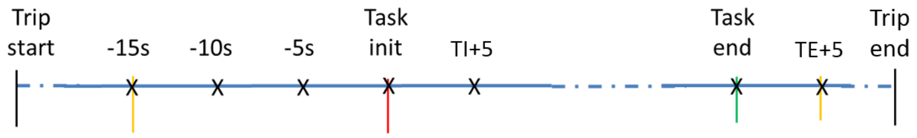



**Figure 5-11: A visualization of the seven analysis points used in this analysis, as well as the start and end of the trip**

*Note: The dotted sections of the time-line represents a variable length, while the solid lines represents 5s equal separation between points.*

### 5.2.1 Speed and THW differences between analysis-points – tabulation

To understand how (if) drivers adapt their safety margins when engaging in secondary tasks the speed differences between the seven analysis-points described above were analysed. Table 5-2 shows the speed difference between adjacent APs (e.g., between −15s and −10s). Note that for each data point the average speed and THW is shown. Difference for each driver was calculated first, then the average was taken across drivers (and the SD was calculated across the 24 drivers' averages). For each combination a student t-test was performed. The p-value for each combination is included in Table 5-2. Grey cells with bold *p*:s are significant with an alpha of 0.01 (the reader is encouraged to reflect on the multiple test-issue, and interpret the statistics accordingly). Figure 5-12 shows the same data graphically.

Note that Table 5-2 shows the results of the analysis of drivers' (potential) speed and THW adaptation when initiating and performing both phone tasks, and all annotated tasks. Initially, only phone tasks (any type) were analysed, but the relatively low number of observations provided low power. However, when all 1800 tasks were analysed, several comparisons (cells) showed significant results, even with a relatively strict significance criterion (alpha=0.01). The results showed relatively large effect sizes.

For example, for phone tasks there was a significant increase in speed between a) before and up to five seconds into the task, and b) the end of or after the task. That is, drivers seem to have self-regulated by increasing their speed after the phone task ended, rather than reducing it before the task was initiated. There may, of course, be alternative explanations, for example, as drivers typically initiate phone tasks shortly after having passed an intersection (see Tivesten and Dozza, 2015; and Figure 4-16), at least in traffic-light-controlled intersections, there may be a speed increase between before and after the intersection (and the corresponding −15s and onward). Also, as task interactions are common at the start of trips (Figure 4-14) they may start on smaller (lower speed-limit) roads, moving into larger (higher speed-limit) roads – resulting in an increase in speed between before and after the phone tasks. Further studies are needed to understand if the speed increase is an actual effect or an artefact of such confounders.

When all tasks were considered together, results showed a significant speed decrease between the first two time-points: −15 s and −10 s before task start (−.21 and −1.07 km/h, respectively) and task start. This may be due to drivers self-regulating, but further analysis showed that stopped vehicles affected the results. In Table 5-2, when considering all tasks, all secondary tasks are used, including those where the driver has pulled in and stopped the vehicle. When the secondary tasks (of "all task") where at least one analysis point had a speed of zero were removed (Table 5-3), a similar trend as for phone-task-only appeared — no decrease in speed between before the task and task start, but instead an increase in speed from −15 s before the task until task end, and 5 s after the task end. That is, drivers did not significantly decrease the speed before or during the task, but instead increased their speed after it was completed. Removal of standstill task was done also for phone tasks. The trends were the same as with the standstills included. One way of interpreting the results of speed differences over the seven points between with and without standstill, and in relation to the low proportion of task initiations at stand-still, is that drivers who stop and interact with their phones slow down, but do not stop, until after they have initiated the phone interactions. Note that, as described above, there still may be confounders related to context (which in turn affects self-regulation) for both all tasks and phone-task-only that influenced the results.

There were no significant differences in the THW for any combination of time-points. This was not expected. We expected THW to follow speed in significance. If there truly was no difference in THW when there was a speed difference between points (Figure 5-11), the lead-vehicle would need to slow-down/speed up at the same time as the host vehicle. There are, however, two other main potential rationales to this unintuitive result. First, due to limitations in range measurements, THW is available much less often than speed (lower N). Second, and similarly lowering N, to calculate THW there needs to be a lead-vehicle. Time-points when no lead-vehicle was present were ignored. Consequently, analysis in terms of comparing significance between speed and THW should be made with care, if at all.

**Table 5-2: The speed and THW difference for the 6x6 combination of analysis points for phone tasks and all tasks**

| | -10s | | | -5s | | | Task start | | | Task start +5s | | | Task end | | | Task end +5s | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Phone | All tasks | | Phone | All tasks | | Phone | All tasks | | Phone | All tasks | | Phone | All tasks | | Phone | All tasks | | |
| | Speed | Speed | THW | Speed | Speed | THW | Speed | Speed | THW | Speed | Speed | THW | Speed | Speed | THW | Speed | Speed | THW | |
| -15s | 0.48 | -0.16 | 0.37 | 0.19 | **-0.77** | 0.1 | 0.11 | **-1.21** | -0.064 | 0.45 | -1.35 | -0 | 3.35 | -0.8 | -0.12 | **5.30** | -0.012 | 0.11 | *Mean* |
| | 6.48 | 6.03 | 1.85 | 11 | **10.72** | 2.18 | 15.2 | **14.56** | 2.26 | 17.3 | 17.49 | 2.4 | 22.6 | 18.59 | 2.46 | **22.3** | 19.75 | 2.19 | *SD* |
| | 0.28 | 0.25 | 0.06 | 0.8 | **0.005** | 0.25 | 0.91 | **6E-04** | 0.5 | 0.72 | 0.026 | 0.97 | 0.033 | 0.17 | 0.23 | **0.001** | 0.84 | 0.8 | *p* |
| -10s | | | | -0.23 | **-0.58** | -0.037 | 0.22 | **-1.07** | -0.26 | -0 | -1.22 | -0.08 | 2.89 | -0.58 | -0.39 | **4.60** | 0.21 | -0.006 | *Mean* |
| | | | | 6.4 | **6.59** | 1.27 | 11.76 | **11.57** | 2.04 | 14.7 | 15.25 | 2.81 | 20.45 | 16.73 | 2.34 | **20.52** | 18.15 | 2.11 | *SD* |
| | | | | 0.6 | **9E-04** | 0.89 | 0.78 | **5E-04** | 0.034 | 0.99 | 0.021 | 0.39 | 0.04 | 0.34 | 0.018 | **0.002** | 0.49 | 0.31 | *p* |
| -5s | | | | | | | 0.86 | -0.44 | -0.076 | 0.53 | -0.4 | 0.092 | 3.3 | 0.81 | -0.32 | **4.82** | 0.0021 | -0.001 | *Mean* |
| | | | | | | | 6.9 | 6.81 | 1.51 | 10.85 | 11.37 | 2.25 | 18.87 | 16.11 | 2.28 | **19.25** | 5.82 | 2.38 | *SD* |
| | | | | | | | 0.86 | 0.021 | 0.43 | 0.52 | 0.41 | 0.66 | 0.011 | 0.034 | 0.13 | **4E-04** | 0.55 | 0.99 | *p* |
| Task start | | | | | | | | | | 0.63 | 0.002 | -0.22 | **3** | **0.95** | -0.003 | **4.32** | **1.48** | --0.11 | *Mean* |
| | | | | | | | | | | 6.08 | 5.81 | 1.81 | **17.18** | **12.73** | 1.65 | **18.15** | **14.6** | 1.85 | *SD* |
| | | | | | | | | | | 0.16 | 0.55 | 0.36 | **0.009** | **0.002** | 0.89 | **0.013** | **0.0002** | 0.65 | *p* |
| Task start +5s | | | | | | | | | | | | | 3.2 | **1.85** | 0.44 | **4.80** | **2.72** | 0.55 | *Mean* |
| | | | | | | | | | | | | | 17.06 | **16.26** | 2.1 | **17.78** | **17.72** | 2.43 | *SD* |
| | | | | | | | | | | | | | 0.013 | **0.002** | 0.31 | **5E-04** | **0.0002** | 0.24 | *p* |
| Task end | | | | | | | | | | | | | | | | 0.68 | -0.004 | -0.95 | *Mean* |
| | | | | | | | | | | | | | | | | 5.97 | 5.95 | 1.34 | *SD* |
| | | | | | | | | | | | | | | | | 0.1 | 0.92 | 0.71 | *p* |

*Note: All secondary tasks, including those where the drivers were standing still, are included in this table.*

**Table 5-3: The speed difference between for the 6x6 combinations of analysis points for all tasks, with all secondary tasks, where at least one analysis point had zero speed, removed**

| | Speed | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | -10s | | -5s | | Task start | | Task start +5s | | Task end | | Task end +5s | | |
| | Phone | All tasks | Phone | All tasks | Phone | All tasks | Phone | All tasks | Phone | All tasks | Phone | All tasks | |
| -15s | 1.25 | 0.40 | 1.94 | 0.73 | 2.97 | **1.34** | 3.56 | 1.73 | **7.00** | **2.16** | **8.13** | **1.79** | *Mean* |
| | 6.53 | 6.13 | 11.17 | 10.69 | 14.73 | **13.91** | 16.73 | 16.71 | **21.84** | **17.57** | **21.79** | **19.14** | *SD* |
| | 0.024 | 0.029 | 0.040 | 0.022 | 0.018 | **0.0014** | 0.024 | 0.017 | **0.00020** | **0.000045** | **0.000021** | **0.0022** | *p* |
| -10s | | | 0.76 | 0.36 | 1.85 | **0.97** | 2.36 | 1.18 | **5.79** | **1.77** | **6.84** | **1.40** | *Mean* |
| | | | 6.58 | 6.51 | 11.12 | **10.89** | 14.18 | 14.71 | **19.62** | **15.59** | **19.79** | **17.42** | *SD* |
| | | | 0.17 | 0.066 | 0.048 | **0.0028** | 0.077 | 0.063 | **0.00054** | **0.00015** | **0.000069** | **0.0080** | *p* |
| -5s | | | | | 1.17 | **0.71** | 1.94 | 1.14 | **5.21** | **1.54** | **6.21** | 1.18 | *Mean* |
| | | | | | 6.19 | **6.36** | 10.32 | 11.21 | **18.86** | **13.50** | **19.06** | 15.57 | *SD* |
| | | | | | 0.023 | **0.00019** | 0.043 | 0.018 | **0.0010** | **0.00013** | **0.00014** | 0.012 | *p* |
| Task start | | | | | | | 1.02 | **0.44** | 3.74 | **1.11** | **4.46** | 0.69 | *Mean* |
| | | | | | | | 6.18 | **6.05** | 18.32 | **11.82** | **19.05** | 13.96 | *SD* |
| | | | | | | | 0.070 | **0.008** | 0.012 | **0.0013** | **0.0046** | 0.096 | *p* |
| Task start +5s | | | | | | | | | 3.56 | 1.62 | **4.53** | 1.21 | *Mean* |
| | | | | | | | | | 18.53 | 14.98 | **18.86** | 16.80 | *SD* |
| | | | | | | | | | 0.035 | 0.010 | **0.0094** | 0.090 | *p* |
| Task end | | | | | | | | | | | 0.12 | **-0.69** | *Mean* |
| | | | | | | | | | | | 6.18 | **6.44** | *SD* |
| | | | | | | | | | | | 0.81 | **0.00027** | *p* |

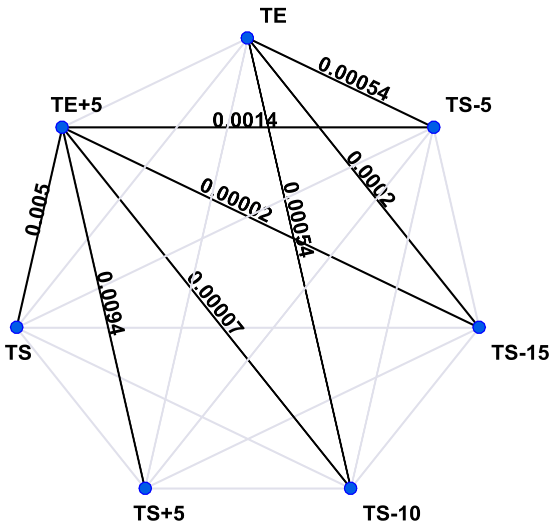



**Figure 5-12: A visual representation of the t-test combinations in Table 5-3, with significant comparisons shown in black**

## 5.2.2 Speed difference between tasks – modelling

Table 5-2 and Table 5-3 above demonstrate the overall adaptation of safety margins made (or not) by drivers when performing secondary tasks in the UDRIVE trucks. However, there are several potential confounding variables that would be valuable to study further. In this section two modelling approaches were evaluated. The first was the use of a generalized linear model (GLM), and the second the use of linear mixed effect models (LME) with driver as random effect (in the latter). In both modelling approaches, age and gender were modelled to predict the difference in speed between the two analysis points “Task start −15s” and “Task end”. The intent was to include more predictors in the model, such as weather, light conditions, closeness to intersection, lead-vehicle present, time-headway, absolute speed, etc., but limitations in annotations and time constraints for the analysis prohibited such analysis.

However, finally, the following models were used to fit the data with the Mixed Effect Model approach and, without the random effect, with the Generalized Linear Model:

1. Speed ~ Gender + (1 | DriverID)
2. Speed ~ Age + (1 | DriverID)
3. Speed ~ Age + Gender + (1 | DriverID)

None of the above models showed any statistically significant difference, meaning that neither Gender, nor Age can be used to predict the speed difference between the two selected time instances. A possible influence on the results can be attributed to the big disparity between the number of samples of the categories in “Gender”. In particular, the number of samples related to female was too low to be valuable to use (only 2 samples). Age was not expected to have any predictive effect on the speed choice.

### 5.2.3 Conclusions

The analysis performed in this chapter aimed to understand if and how drivers adapt their safety margins when engaging in secondary tasks. For phone tasks, there was a significant increase in speed between before and up to five seconds into the task, and the end of or after the task. That is, drivers seems to self-regulate by increasing their speed after the phone task ended, rather than reducing it before the task is initiated. There may, however, be several confounders affecting this result.

Contrary to speed, THW was not significantly different across the analysed time-points. However, this is probably due to limitations in range measurements; THW is available much less often than speed is (thus fewer samples were available). Moreover, the need of a leading vehicle is crucial for THW calculation. Time-points where no lead-vehicle was present were ignored. Generalized Linear Model (GLM) and Linear Mixed Effects Model (LME) were included in the analysis with the purpose to investigate if variables like weather, light conditions, closeness to intersection, lead-vehicle present, time-headway, absolute speed could have been useful to predict speed variations. For several reasons, only age and gender were finally modelled to predict the difference in speed between the two analysis points “Task start -15s” and “Task end”. As expected, neither gender nor age were significantly affecting the speed difference.

# 6 Implications

## 6.1 Who does it most?

There are considerable differences in activity between the car and truck drivers observed in UDRIVE. The car drivers were engaged in secondary tasks for about 10 percent of their driving time, while the truck drivers were engaged in secondary tasks for about 20 percent of their driving time. The activity for the latter group includes a substantial amount of smoking by a minority of the drivers. The amount of mobile phone use was similar between the driver groups — about 5 percent for the truck drivers and 4 percent for the car drivers.

Female car drivers spent more time than males in hands-free conversation and also in handheld interaction, while males were more active in interaction with a handheld phone, which might be setting or adjusting a navigation app. There were also substantial differences by country among the car drivers, with the Polish drivers engaged in secondary tasks for 19.4 percent of their driving time and using the mobile phone for 9.8 percent of their driving time. The German drivers had the least activity, with the French and British drivers in between.

It can also be observed that not everyone carries out the more demanding activities and that personality and habitual behaviour has a considerable impact on willingness to engage in secondary tasks: a composite personality score had a clear relationship to whether or not car drivers were engaged or did not engage in a number of activities, including mobile phone use.

## 6.2 Policy aspects

There are a range of countermeasures to distracted driving available. Campaigns have some merit in that it is generally that they can help persuade drivers of the road safety case for compliance with the law, and potentially encourage drivers to consider safety when deciding whether or not it is appropriate to engage in secondary tasks. But while campaigns may soften up drivers to be more conscious of safety, they are clearly not sufficient on their own to change behaviour. This can been seen, for example, in the general tendency (apart from the German car drivers) to carry out visual-manual tasks while driving as evidenced in the behaviour of the Polish car drivers and to a lesser extent the French and British car drivers. The truck drivers have been observed to have a particularly high rate of activity, stimulated no doubt in part by work-related interaction, since they were driving in a delivery operation. The resistance of some drivers to safety-related messages can also be seen in the personality analysis — those who are less safety conscious are considerably more prone to various activities.

Thus campaigns need to be supplemented by legal restrictions, penalties and police enforcement. Unfortunately, detection of illegal mobile phone use (e.g. handheld calling and texting) is difficult, and drivers are probably aware of this. This most likely limits the impact of stiffer penalties, although it not an argument against having strict penalties, including points on the license. In the discussion on countermeasures at the UDRIVE Final Event, there was substantial support from the audience for technological countermeasures as a means to prevent distracting activities. The most obvious such countermeasure would be mobile phone software that prevents handheld calling and interaction, including texting. Another possibility is only permitting interaction through the vehicle, i.e. requiring drivers to connect to their vehicles via Bluetooth, and then having the vehicle regulate what is and what is not permitted in terms of activity. Such a solution may in any event become necessary with the advent of highly automated driving.

# 7 Conclusions

## 7.1 Conclusions

As has been stated in section 6.1, considerable differences could be observed in the level of activity and the type of activity between various driver groups. Type of driver (car versus truck), country of driver, gender of driver and personality of driver all play a role in affecting propensity to engage.

In terms of whether the drivers can be seen to be self-regulating their activities — engaging in tasks when it can be considered safer — there are definite signs that this is indeed the case. The car drivers carried out more of their visual-manual interactions with a handheld mobile phone while stopped (56%) For the truck drivers only around 10 percent of all phone task initiations were while at a standstill, but there was still a substantial tendency for these initiations to be at lower speeds than overall driving. This same tendency could be seen for the visual-manual interaction of the car drivers, where the mean speed for activity was around 20 km/h and interaction speeds were generally lower than those in the preceding (prior to interaction) driving period.

The car drivers adapted somewhat to situational demands. Task duration went down as manoeuvring complexity went up and also reduced somewhat as environmental complexity increased. However, more detailed analysis showed that in difficult environmental conditions, compared to easy ones, a greater share of the secondary tasks that were performed were more difficult. This drivers seem to have a rather imperfect understanding of how to balance primary task demand against secondary task demand, or find the secondary tasks so compelling that they do not drop them when the situation becomes more difficult.

## 7.2 Recommendations

There are, as discussed in section 6.2, some important policy aspects. Campaigns, penalties and enforcement, may need to be supplemented by technological solutions.

It is also important that there be further analysis of the UDRIVE database in this very important area. The analysis reported here has not been carried out for all the drivers, and notably the Duct car drivers were not included. It has also been carried out fairly rapidly, meaning that the richness of the database, has not been fully exploited. It is very much to be hoped that the data will be used extensively for further investigation of how, when and where drivers engage in secondary tasks and how their activities are adapted to the dynamics of the traffic situation.

# 8 References


Arnett, J. (1994) Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, 16(2), 289–296.

Brookhuis, K.A., de Vries, G. and De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, *23*(4), 309-316.

Carsten, O. and Merat, N. (2015). Protective or not? *Proceedings of the 4th International Conference on Driver Distraction and Inattention*, Sydney, November.

Consiglio, W., Driscoll, P., Witte, M. and Berg, W.P. (2003). Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. *Accident Analysis and Prevention*, *35*(4), 495-500.

Dellinger, F., Robert-Seidowsky, M., Bernard, E., Guyonvarch, L. and Guillaume, A. (2017). Computer vision algorithms for detecting secondary tasks in naturalistic driving studies. *Proceedings of the 5th International Conference on Driver Distraction and Inattention*, Paris, March.

Department for Transport (2005). *Evaluation of the effectiveness of the National Driver Improvement Scheme*. Road Safety Research Report No. 64. Department for Transport, London.

Dingus, T.A. (2014). Estimates of prevalence and risk associated with inattention and distraction based upon in situ naturalistic data. *Annals of Advances in Automotive Medicine*, 58: 60-68.

Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M. and Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, 113(10), 2636-2641.

Engström, J., Johansson, E. and Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 97-120.

French, D.J., West, R.J., Elander, J. and Wilding, J.M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics*, 36, 627–644.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention*, 37(3), 461–472.

Hancock, P.A., Lesch, M. and Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention*, 35(4), 501-514.

Hickman, J.S., Hanowski, R.J. and Bocanegra, J. (2010). Distraction in commercial trucks and buses: assessing prevalence and risk in conjunction with crashes and near-crashes. Report No. FMCSA-RRR-10-049, Federal Motor Carrier Safety Administration, U.S. Department of Transportation, Washington, D.C.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T.J. and Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis and Prevention*, 38(1), 185-191.

Jamson, A.H. and Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 79-96.

Kircher, K., Patten, C. and Ahlström, C. (2011). Mobile telephones and other communication devices and their impact on traffic safety – a review of the literature. VTI rapport 729A. VTI, Linköping, Sweden.

Lajunen, T., Parker, D. and Summala, H. (2004). The Manchester Driver Behavior Questionnaire: A cross-cultural study. *Accident Analysis and Prevention*, 36, 231-238.

Lerner, N. and Boyd, S. (2005). *On-road study of willingness to engage in distracting tasks.* Report DOT HS 809 863. National Highway Traffic Safety Administration, Washington, DC.

Metz, B., Landau, A. and Just, M. (2014). Frequency of secondary tasks in driving – results from naturalistic driving data. *Safety Science*, 68, 195–203.

Nagel, H.-H. (1994). A vision of 'vision and language' comprises action: an example from road traffic. *Artificial Intelligence Review,* 8, 189-214.

Östlund, J., Peters, B., Thorslund, B., Engström, J., Markkula, G., Keinath, A., Horst, D., Juch, S., Mattes, S. and Foehl, U. (2005). *Driving performance assessment methods and metrics*. D2.2.5 of AIDE.

Özkan, T. and Lajunen, T. (2005). Multidimensional Traffic Locus of Control Scale (T-LOC): factor structure and relationship to risky driving. *Personality and Individual Differences*, 38(3), 533–545.

Parker, D., Stradling, S.G. and Manstead, A. (1996). Modifying beliefs and attitudes to exceeding the speed limit: An intervention study based on the theory of planned behaviour. *Journal of Applied Social Psychology*, 26*,* 1-19.

Ranney, T.A., Mazzae, E., Garrott, R. and Goodman, M.J. (2000). NHTSA driver distraction research: Past, present, and future. In: *Driver distraction internet forum*.

Rosenbloom, T. (2006). Driving performance while using cell phones: an observational study. *Journal of Safety Research*, 37(2), 207-212.

Sayer, J.R., Devonshire, J.M. and Flannagan, C.A. (2005). *The effects of secondary tasks on naturalistic driving performance.* Report No. UMTRI-2005-29. University of Michigan Transportation Research Institute, Ann Arbor, Michigan.

Sayer, J., Devonshire, J. and Flannagan, C. (2007). Naturalistic driving performance during secondary tasks. In: *Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.*

Schießl, C. (2009). Modellierung von Belastung und Beanspruchung im Fahrkontext. Doctoral Dissertation, Technische Universität Carolo-Wilhelmina zu Braunschweig.

Schömig, N., Metz, B. and Krüger, H.-P. (2011). Anticipatory and control processes in the interaction with secondary tasks while driving. *Transportation Research Part F: Traffic Psychology and Behaviour,* 14(6), 525–538.

Tivesten, E. and Dozza, M. (2015). Driving context influences drivers' decision to engage in visual–manual phone tasks: Evidence from a naturalistic driving study. *Journal of Safety Research,* 53, 87-96.

TRL, TNO and RAPP-TRANS (2015). *Study on good practices for reducing road safety risks caused by road user distractions*. Final report. European Commission, DG MOVE, Brussels.

Victor, T., Dozza, M., Bärgman, J., Boda, C-N., Engström, J., Flannagan, C., Lee, J.D. and Markkula, M. (2015). *Analysis of naturalistic driving study data: safer glances, driver inattention, and crash risk.* SHRP 2 Report S2-S08A-RW-1. The Second Strategic Highway Research Program. Transportation Research Board, Washington, D.C.

Warner, H., Özkan, T. and Lajunen, T. (2010). Can the traffic locus of control (T-LOC) scale be successfully used to predict Swedish drivers' speeding behaviour. *Accident Analysis and Prevention*, 42(4), 1113-1117.

West, R.J., French, D.J., Kemp, R. and Elander, J. (1993). Direct observation of driving, self reports of driver behaviour and accident involvement. *Ergonomics,* 36(5), 557-567.

# 9 List of Figures

# 10 List of Tables