

SESSION 7 : NATURALISTIC DRIVING STUDIES

TRYING TO VALIDATE SUBJECTIVE REPORTS WITH NATURALISTIC DRIVING DATA – A CASE AGAINST QUESTIONNAIRES AND SURVEYS TO QUANTIFY DRIVER DISTRACTION

Tibor Petzoldt¹ & Fabian Utesch²

¹ Technische Universität Chemnitz, Germany

² DLR Braunschweig, Germany

ABSTRACT: The effects of different forms of driver distraction on driving performance have been studied for years and are comparatively well understood. How often drivers actually engage in different distracting activities, however, is less clear. Available methods are either not able to provide a complete picture, or are extremely expensive. Post-drive questionnaires and surveys might provide a cheap solution to the problem. As part of a naturalistic driving pilot study, we tried to validate a post-drive survey/questionnaire that is intended to capture the occurrence and duration of different secondary tasks. However, for a variety of reasons, this attempt was unsuccessful. It became clear that there was a huge discrepancy between the drivers' naïve understandings of secondary tasks (what is it, what is part of it, how long is it, etc.) and scientific definitions of the same concepts. Further problems arose from the fact that even though questioned directly after the trip, many accounts appeared to have been reconstructions, rather than recollections of secondary task engagement. We conclude that subjective accounts of secondary task engagement are largely inappropriate to quantify driver distraction.

1 INTRODUCTION

The fact that certain distracting activities have a negative influence on driver behaviour and driving performance is well established through a large number of laboratory and on-road studies (for an overview see Regan, Lee, & Young [1]). However, how frequently or for how long drivers actually engage in these activities is often unknown. The analysis of crash data [2] cannot provide the required information, experimental studies cannot capture natural user behaviour, and observations from outside the vehicle [3] are limited in what can be observed. Unfortunately, large scale naturalistic driving studies [4], which are able to overcome most of the issues of other methods,

Naturalistic Driving Studies are extremely time and resource consuming. Drivers' own reports, either in form of surveys or questionnaires, might provide a cheap solution to the problem.

However, few efforts have been made to use questionnaires or surveys to systematically assess how often or for how long drivers engage in secondary tasks and other activities. Feng, Marulanda and Donmez [5] proposed a "Susceptibility to Driver Distraction Questionnaire (SDDQ)", in which participants are asked (among other facts) about their engagement in distracting activities. However, the possible answers are "never; rarely; sometimes; often; very often", and hardly allow for a good estimate of prevalence or duration (which, admittedly, is not the declared goal of the questionnaire). Similarly, the American "National Survey on Distracted Driving" [6] asked participants by phone how often they would engage in certain activities while driving, again only with answer categories such as "always; almost always; sometimes; rarely; never".

McEvoy, Stevenson and Woodward [7] assessed the prevalence of distracting activities through a telephone survey in which they asked their participants to provide "the frequency of distracting activities during the most recent driving trip" (p. 243), going through a typical list of distracting activities in the process. Here, it appears that respondents provided actual numerical frequencies (although this does not become clear from the analysis). The authors reported an estimate of one distracting activity every six minutes, however also acknowledged that "the time spent on each activity may vary" (p. 245). Huemer and Vollrath [8, 9] finally provided an attempt at capturing the actual duration, asking participants for how long they have been engaged in specific secondary tasks during the last 30 min of driving. They proposed to approach drivers at parking lots directly after they finished a trip, so the memory of their recent driving behaviour would still be rather accessible.

Unfortunately, it appears that so far, there has been no attempt to validate such a questionnaire or interview on naturalistic driving data. It is unclear how well reported frequency or duration of secondary task engagement reflects actual objective engagement in these activities. As part of a German naturalistic driving pilot study, we tried to find out if there is any correspondence between reported and observed distraction.

2 METHOD

We instrumented a van (Volkswagen T5) with an improved version of a data acquisition system that had been developed in a previous project [10]. The van was part of the German Aerospace Center (DLR) car pool and used by 15 different drivers during the study (all of which had consented to being recorded). The main goal of this pilot study was to test the system under naturalistic driving conditions. As part of the project, however, we also intended to showcase the whole process, including actual data analyses. Driver distraction was identified as an ideal pilot research topic, as its assessment required extensive video annotation (which is a vital part of naturalistic driving data analysis).

We used the Huemer and Vollrath [9] survey and administered it as a questionnaire. The questionnaire contained a list of activities that is based on a review of studies (many of them naturalistic approaches) and statistics in which different categorisations of distracting activities could be found [11-13]. We directly transferred this approach to a post-drive questionnaire, keeping all the categories, explanations and scales identical (Table 1). Drivers were asked whether they had engaged in any of the listed activities in the last 30 min of their drive, and if so, for how long. While our drivers were made aware when they consented to participate that they might be asked certain questions about their trips at some point, they did not know in advance when they would have to report on their activities, nor were they aware of the nature of the eventual questions. As drivers filled in the questionnaire with experimenters close by, any questions and problems while completing the questionnaire were recorded as well.

To compare our drivers' subjective reports of secondary task engagement with their actual behaviour, we annotated the video material from all their drives, including the 30 min of driving that were covered by the questionnaire. To allow for a comparison, we designed our video annotation scheme based on the questionnaire. We used the same secondary task categories that were included in the questionnaire, and added explanations and guidelines for the annotators following Stutts et al. [13] and VTTI [14].

Table 1: Questionnaire with distracting activities for participants to complete

	Yes/No	For how long (in min)?
Eating & drinking		
Smoking		
Grooming & clothing related activities <i>e.g., manicure, change clothes</i>		
Operation of vehicle functions & built-in devices <i>e.g., adjust seat, mirror, tune the radio</i>		
Operation of nomadic devices <i>e.g., operate iPod, mobile phone</i>		
Activities related to passengers <i>e.g., converse, gesture, hand over objects (e.g., food)</i>		
Other activities <i>e.g., pet related, search for objects, read / write, tidy up / clean the car</i>		
"Self-initiated" activities <i>e.g., soliloquy, sing, think about something, look at something intensely (inside vehicle)</i>		
Distraction from outside the vehicle <i>e.g., route related (e.g., work zone), look at something (e.g., pedestrian, billboard), listen to something (e.g., music from another vehicle, horn)</i>		

3 FINDINGS

Initially, we planned a statistical comparison of subjective and annotated data. However, already when our drivers completed the questionnaire, it became apparent from their inquiries that this might prove difficult. This impression was corroborated very early in the analysis process, when we found substantial differences between the two datasets already during a first inspection of the raw data. As a consequence, we decided to instead focus on these differences and investigate them on a descriptive level. Soon, it became clear that trying to validate a questionnaire with naturalistic driving data is not such a good idea after all, for a variety of reasons.

One central issue is the fact that the overlap between scientific definitions of distraction and secondary tasks (as they are used for video annotation) and the common understanding of what a secondary task is (as it would turn up in a questionnaire) is limited. This results not only in problems determining what a secondary task is and what not, but also causes difficulties for the

assessment of secondary task duration (when does it start, when does it stop?) and the segmentation of secondary tasks (is it one long task, or several short ones?).

For example, the SHRP2 data dictionary [14] lists the categories “reaching for food-related or drink-related item”, “eating with utensils” and “eating without utensils” as the three categories under which some eating-related activity would be filed. From the definitions, it becomes clear that the annotators’ main concern lies in the motor activity that is associated with either locating or holding the item, whereas chewing food (with both hands on the wheel) would, at least from the definition, not be considered a distracting activity. Based on an understanding of motor distraction (mainly hands-off-the-wheel) as a major factor in incidents and crashes, this appears to be fully reasonable from a scientific perspective. However, it hardly follows an everyday understanding in which chewing (and finally swallowing) is the very essence of eating. In our data, we found instances in which participants reported eating for 15 min during their 30 min drive, which was not reflected in our annotations. Reinspection of the video revealed that participants were just chewing gum all the time, and apparently considered this an activity that had to be labelled eating.

The eating category causes other problems as well. For Stutts et al. [13], eating “starts when food is brought to mouth (or mouth to food) and stops when food or hand is removed from mouth” (p. 25). In our dataset, we found a participant eating from a bag of chips. According to the definition, every single instance of reaching for a chip and bringing it to the mouth would be annotated as a separate episode of eating (again with chewing not being part of the annotation). Understandably, our participant expressed difficulties when requested to quantify the amount of time he spent eating during the last 30 min of driving. He might have been even more troubled when asked about the number of separate eating episodes had we followed the approach of McEvoy et al. [7].

Such issues are not limited to the activity of eating. We found several instances in which participants reported to have been involved in a conversation for 30 min, i.e., for the whole trip segment that was covered. Again, this was not reflected in our annotations. Stutts et al. [13] consider a

conversation as active “as long as someone is responding within ~10 seconds” (p. 26), which is a somewhat artificial criterion, that does, especially for longer trips, result in a considerable number of separate conversations (with a substantial amount of time that is not labelled as conversation). However, when our drivers were asked directly about their trip, they hardly differentiated into separate conversations when they were travelling with the same passenger for the complete trip. Again, this would have been even more problematic had we asked our participants not about the duration, but about the number of conversations during the trip. It can be suspected that even if a differentiation in separate conversations would be made, it would rather occur along the lines of conversation topics (e.g., conversing about private issues first, and then talking about business - two conversations), and not follow some seemingly arbitrary criterion that is based on the duration of silence between the drivers’ and passengers’ contributions.

The estimation of task duration is also a problem in secondary tasks that are comparatively short (e.g., tuning the radio). The questionnaire asks for the duration of certain activities to be reported in minutes, which appears to be reasonable, as a driver cannot be expected to be much more precise anyway (it has to be acknowledged, however, that some participants reported fractions of minutes). In contrast, video annotation is done frame by frame. This can lead to serious overestimations of total task time in the survey data for the shorter tasks. Several of our participants reported to have been operating vehicle functions and built-in devices for as long as five minutes. Our annotations showed that the participants indeed had operated the radio several times, but total duration did not even come close to the reported five minutes. Especially when there are several short interactions, it might be suspected that participants just infer total task duration based on a subjectively generated mean task duration and the number of individual operations. Our participants often seemed to reconstruct, rather than actually recall, task duration.

Non-observable aspects of a task might add to the distortion, as they are not part of the annotated data set, but might be included in the drivers’ post hoc duration estimate. Deliberate thought processes that might precede an observable distracting activity (e.g., trying to remember a phone number

before actually placing the call, thinking about the route before starting to interact with the navigation system) could be included in a subjective assessment of the activities' duration, but cannot be annotated based on the video recording, which, again, can lead to discrepancies between reported and observed task duration. For tasks that are purely cognitive, the aspect of observability is an even bigger problem, as the actual cognitive activity that is suspected to occur cannot be observed directly. In a naturalistic driving study, all that can be done is to infer such a form of distraction based on observable behaviour. Consequently, in their analysis of the SHRP2 dataset, Victor et al. [15] did not include the annotated activities "lost in thought", "looked but did not see" and "cognitive, other", "as they were believed to be questionable categories" (p. 43). Drivers, on the other hand, might well be able to report instances of purely cognitive distraction.

The category of distraction from the outside is equally problematic. In video annotation, the analysts have to rely on observed head position and glance behaviour, at times aided by one of the outside views. A judgment on what glance can be "considered to be part of the driving task" [14] is usually difficult (even more so when there is no usable information from the outside views available), and may differ between the clearly defined driving task on a highway and the highly dynamic and variable urban environment. Especially in an urban environment, it might be a challenge to define what is part of the driving task in the first place (e.g., is taking the eyes of the road when looking for parking spot part of the task?). The driver, on the other hand, will tend to report instances in which he clearly recalls to have been specifically captured by an outside stimulus, which might or might not have resulted in an observable change in gaze direction. He might have been looking straight ahead, but report that his attention was captured, for example, by some object in the vehicle he was following.

Further issues arise from the fact that some activities cannot be easily ascribed to one of the categories, or are composed of sub-tasks that belong to different categories. In a survey or questionnaire, for reasons of economy, drivers are confronted with very broad categories of secondary tasks and provided with examples that are instructive, but not exhaustive. For example, in the Huemer and Vollrath [9] survey, there is the example "search for

Naturalistic Driving Studies objects” listed under “other activities”, whereas “operate mobile phone” is an example for “operation of nomadic devices”. Now, in most cases, operating a mobile phone requires the driver reaching for the mobile phone in the first place. But when does “reaching” for the phone become “searching” for the phone? While in video annotation, some arbitrary distinction can be made and consistently followed (e.g., the SHRP2 dictionary [14] has the category “Cell phone, locating/reaching/ answering”, which is distinct from locating other objects), drivers will follow their individual understanding of the distinction between the categories.

Finally, a very basic, but nevertheless crucial issue is the time window from which participants are required to report. While for annotation the required segment (be it 30 min or any other duration) can be easily selected, it is difficult for drivers to recall when they did what during their trip. Especially after longer trips, routine activities that drivers might engage in repeatedly (e.g., tuning the radio) might be misplaced in time. Our participants often reported to recall that they had engaged in a certain activity during their overall drive, but admitted having considerable difficulties in remembering whether the respective activity had occurred during the last 30 min of their trip, or sometime earlier.

4 CONCLUSIONS

All the mentioned aspects made the validation of questionnaires through naturalistic driving data difficult, if not impossible. No driver can be expected to follow a scientific approach in the assessment of his own activities while driving. At the same time, it is hardly possible to translate the naïve understanding of distraction and secondary tasks into a format that is usable for scientific purposes. To some degree, it might be argued that this is not really a disadvantage, as both approaches can be valuable. After all, aspects that are difficult to observe, such as cognitive distraction or distraction from the outside, might be covered more appropriately by drivers’ subjective accounts of their driving behaviour. Yet, it appears that, overall, drivers’ assessment of their engagement in secondary tasks cannot be trusted. Estimates of frequency and duration of secondary task engagement must be suspected to be severely biased for a variety of reasons. Scientific definitions

of secondary task categories and task elements often do not follow the everyday definitions of driver distraction. This is not necessarily the fault of the drivers. Very often, categorisations are relatively artificial, and not always follow common sense. Yet, most of the time, scientific definitions used in video annotation at least are consistent and clearly described, and follow underlying theories of the effects of driver distraction and secondary task engagement. Annotation guidelines and detailed information on categorisation decisions are usually accessible for everyone, allowing not only for an understanding of the data, but also for a judgment of its value and validity. Such guidelines and decisions are mostly implicit when a driver is asked for his judgment, and may differ not only from the scientific understanding, but also from driver to driver. Although it would be possible to give the same elaborate guidelines to drivers explicitly, this is hardly feasible, and also cannot really be expected to increase data quality. Drivers' reports will still be retrospective accounts, and must be assumed to rely substantially on reconstruction based on experience, rather than actual recollection. Therefore, subjective accounts of secondary task engagement might provide information about what drivers believe they are doing, but should not be understood as a means to actually quantify driver distraction.

5 ACKNOWLEDGMENTS

This paper is based on results of a research project carried out at the request of the German Federal Ministry of Transport and Digital Infrastructure, represented by the Federal Highway Research Institute, under project No. FE82.0400/2010. The authors are solely responsible for the content.

6 REFERENCES

- [1] Regan, M. A., Lee, J. D., and Young, K. L.: Driver distraction: Theory, effects, and mitigation (Boca Raton, FL: CRC Press 2009)
- [2] Stutts, J., Reinfurt, D., Staplin, L., and Rodgman, E.: The role of driver distraction in traffic crashes. (Washington, DC: AAA Foundation for Traffic Safety 2001)

- [3] Pickrell, T. M.: Driver electronic device use in 2012. (Traffic Safety Facts Research Note. Report No. DOT HS 811 884, Washington, DC: National Highway Traffic Safety Administration 2014)
- [4] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. et al.: The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car field experiment. (Report No. DOT HS 810 593, Washington, DC: National Highway Traffic Safety Administration 2006)
- [5] Feng, J., Marulanda, S., and Donmez, B.: Susceptibility to Driver Distraction Questionnaire (SDDQ): Development and relation to relevant self-reported measures. Proceedings of the Transportation Research Board 93rd Annual Meeting. Washington, DC, 2014
- [6] Schroeder, P., Meyers, M., and Kostyniuk, L.: National survey on distracted driving attitudes and behaviors – 2012. (Report No. DOT HS 811 729, Washington, DC: National Highway Traffic Safety Administration 2013)
- [7] McEvoy, S. P., Stevenson, M. R., and Woodward, M.: The impact of driver distraction on road safety: Results from a representative survey in two Australian states, *Injury Prevention*, 2006, 12, pp. 242–247.
- [8] Huemer, A., and Vollrath, M.: Driver secondary tasks in Germany: Using interviews to estimate prevalence, *Accident Analysis & Prevention*, 2011, 43, pp. 1703-1712.
- [9] Huemer, A., and Vollrath, M.: Ablenkung durch fahrfremde Tätigkeiten – Machbarkeitsstudie [Distraction by non-driving activities – feasibility study]. (Berichte der Bundesanstalt für Straßenwesen, Mensch und Sicherheit M225, Bremerhaven, Germany, Wirtschaftsverlag NW 2012)
- [10] Mosebach, H., Lietz, H., Schomerus, J., Petzoldt, T., Baumann, M., Henning, M. et al.: Methodische und technische Aspekte einer Naturalistic Driving Study [Methods and technology issues in naturalistic driving] (FAT-Schriftenreihe 229, VDA 2011)
- [11] Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J., Ramsey, D.

J.: The impact of driver inattention on near-crash/crash risk: An analysis using the 100-Car Naturalistic Driving Study Data (Report No. DOT HS 810 594, Washington, DC: National Highway Traffic Safety Administration 2006)

- [12] Sayer, J. R., Devonshire, J. M., and Flannagan, C. A.: The effects of secondary tasks on naturalistic driving performance (Report No. UMTRI-2005-29, Ann Arbor, MI: The University of Michigan, Transportation Research Institute 2005)
- [13] Stutts, J., Feaganes, J., Rodgman, E., Hamlett, C., Meadows, T., Reinfurt, D. et al.: Distractions in everyday driving. (Washington, DC: AAA Foundation for Traffic Safety 2003).
- [14] Virginia Tech Transportation Institute (VTTI): SHRP2 researcher dictionary for video reduction data. Blacksburg, VA: Virginia Tech Transportation Institute (2014)
- [15] Victor, T., Dozza, M., Bärghman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J.D., and Markkula, G.: Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk (Report S2-S08A-RW-1 2015)

DIGITAL AND VIDEO ANALYSIS OF EYE-GLANCE MOVEMENTS DURING NATURALISTIC DRIVING FROM THE ADSEAT AND TeleFOT FIELD OPERATIONAL TRIALS – RESULTS AND CHALLENGES

Karthikeyan Ekambaram, James Lenard, Steven Reed

Loughborough Design School, Loughborough University, UK

ABSTRACT: The EU projects ADSEAT (2009-2013) and TeleFOT (2008-2012) both included components of work involving naturalistic driving trials in instrumented vehicles. Of specific interest to this paper was the use of video recordings and digital eye-tracker readings to monitor eye-gaze behaviour. The aim of the study was to describe the results and challenges of applying these two methodologies under real-life driving conditions based on nine subjects from the ADSEAT project and ten from the TeleFOT project. It proved possible to detect the effect of navigation devices on driver attention as reflected in eye-glance behaviour through manual review of video recordings. This procedure was however very labour intensive. While the digital eye-tracker produced reliable measurements of head movements through real-time image processing and recognition of facial features, it generally failed to provide meaningful data on eye-gaze movements. There was however several minutes of remarkably accurate eye-gaze readings found within hours of recording that proved the technology could work if the experimental methodology were perfected. This potentially opens the way to cost-effective analysis of eye-gaze behaviour by the application of computerised algorithms to digital files.

1 INTRODUCTION

The naturalistic driving results reported in this paper derive from two EU projects funded under the 7th Framework. The ADSEAT project ran from October 2009 to March 2013. The objective of the study was to provide guidance on how to evaluate the protective performance of vehicle seat designs in reducing whiplash-associated disorders. As part of this study, driving trials were conducted for nine subjects examining head position as a risk factor for whiplash in rear impacts. The TeleFOT project ran from June

Naturalistic Driving Studies
2008 to June 2012 and constituted the largest field operational test of functions provided by in-vehicle aftermarket and nomadic devices conducted to date. This study collected a large amount of data through in-vehicle data loggers and participant questionnaires. One of the aims of the TeleFOT study was to examine the distraction caused by the presence of navigation devices as indicated by changes of eye-gaze behaviour. Of particular relevance to this paper is a series of detailed field operational trials conducted in Great Britain, of which ten subjects are reported in detail.

In both the ADSEAT and TeleFOT studies, it proved quite challenging to obtain the desired quantity and quality of eye-gaze data. Two approaches were adopted, (a) a manual review of video recordings of the driver's face and eyes to identify the object or field of attention and (b) computerised analysis of the digital readings of an eye-tracking device. The purpose of this report is to give an indication of the results that were obtained and the challenges that were encountered in collecting this type of data.

2 MATERIALS AND METHODS

2.1. *Vehicle Instrumentation*

The vehicle used for the trials, a 2010 Ford Mondeo sedan, was fitted with three main test instruments: a data logger for vehicle speed, acceleration and GPS location, a FaceLAB™ eye-tracker for head position and eye-gaze direction, and a four-track video system (Figure 1).



Figure 1 Instrumented vehicle with eye-tracker cameras.

2.2. Driving Routes

The route for the driving trials included urban and suburban regions of Leicester, a city with a population of over 300,000 (Figure 2). The drivers' behaviour at nominated intersection manoeuvres was studied in detail.



Figure 2 Route for driving trials: Loughborough and Leicester, Leicestershire UK.

Volunteers drove the vehicle for around 30–60 minutes through the designated route accompanied by a researcher in the front passenger seat who set up the test instrumentation and monitored it for correct functioning during the trial. In a first series of trials, travel directions were provided verbally by the accompanying researcher while in a second series directions were provided by a portable navigation device mounted in the central region of the upper dashboard.

2.3 Subjects

The nine subjects from the ADSEAT sample comprised five men and four women aged between 23 and 53 years. An impression of the typical seating

postures adopted by participants while driving is provided by the snapshots in Figure 3.



Figure 3 Seating posture while driving for nine ADSEAT subjects.

3 Results

3.1 Video review of eye-glance movement

The results in this section had the target of eye movements identified through manual review of the in-car video recording of ten subjects from the TeleFOT project. The objects of attention for the video review were categorized as 'forwards', 'outside', 'right or left mirror', 'rear-view mirror', 'passenger', 'instrument panel' and 'interior (nfs)' as pictured in Figure 4.

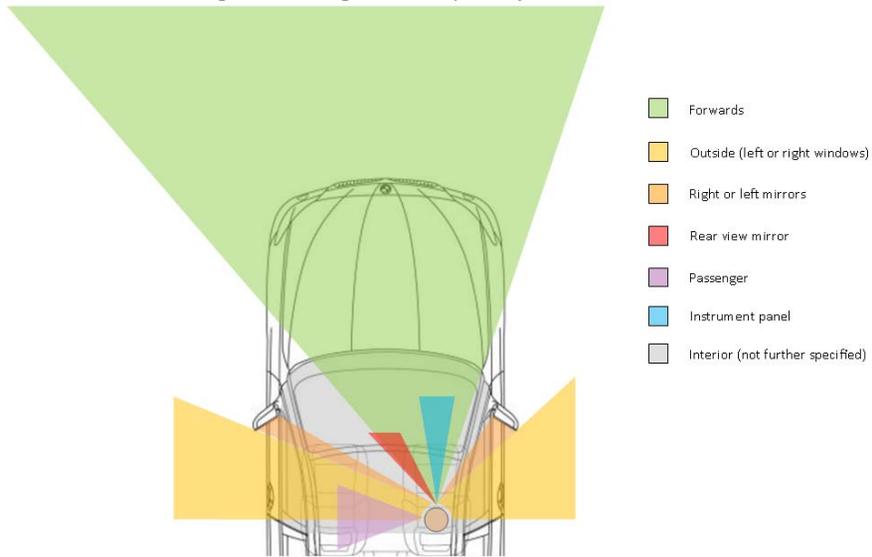


Figure 4 Description of eye glance surfaces/areas

Figure 5 shows the distribution of glances to all recorded objects or fields of attention as a proportion of total glances made by each participant. The largest group of glances was recorded as 'forwards', this category can be described as an eye/head position towards the direction the vehicle is travelling and is bounded by the vehicles nearside and offside A-pillars. This category offers a range of head movement of around 45 degrees, although this is not equally split due to the offset right driving position (see Figure 4). As such, this group may include some head rotation but it is generally recorded when the driver is not looking at any other definable feature. As expected 'forwards' glances account for over 90% of glances for most participants with no driver falling below 85%. To eliminate the variance in glances to objects or fields of attention other than forwards, Figure 6 uses the same data as Figure 5 but for all participants combined, therefore creating an average of all the glances made to each location.

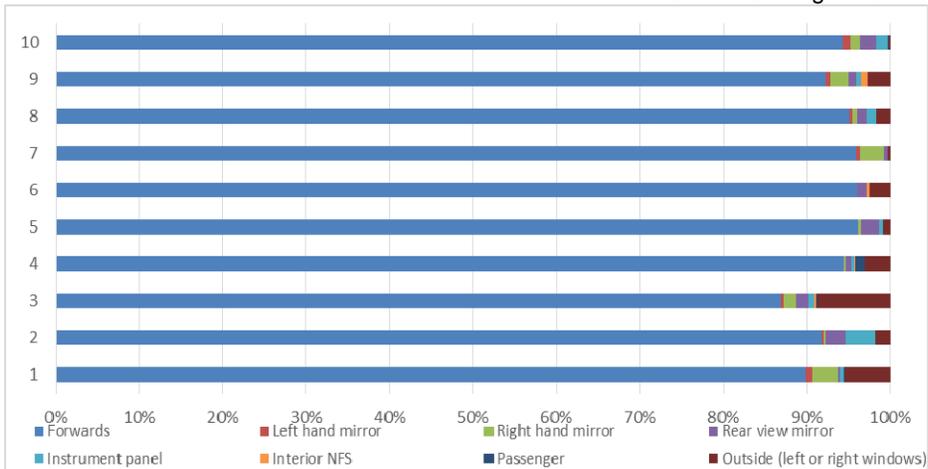


Figure 5 Distribution of glances in normal driving (by participant).

Figure 6 shows that apart from ‘forwards’ glances the next most common glance location is to the outside through either the offside or nearside door windows. A glance of this type will induce considerably more rotation of the drivers head, particularly with respect to the nearside door window, than a glance to the forward roadway. Glances to the right-hand door mirror and rear-view mirror are also relatively common but result in much less head rotation and in some cases, depending of driver stature and seating position, can be glanced at with eye movement alone.

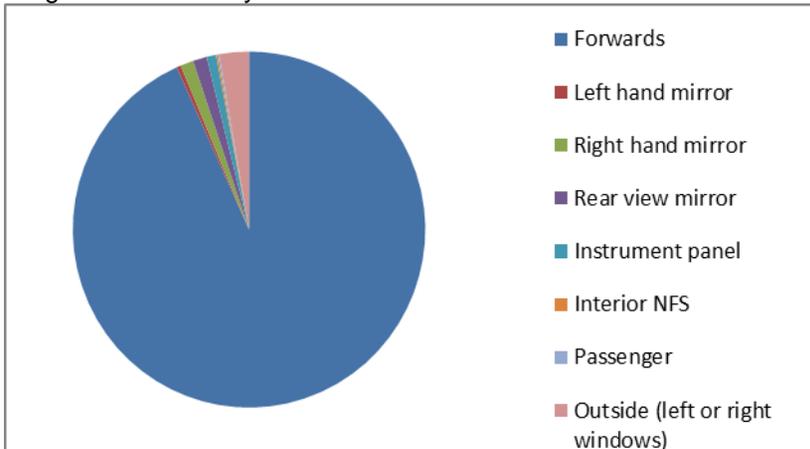


Figure 6 Distribution of glances in normal driving (aggregated).

In order to identify the effect of the other glance locations Figure 7 and Figure 8 have the ‘forwards’ category removed to give better clarity on the different objects or fields of attention. The chart is ordered to show the largest groups

first, namely the areas outside through the side windows followed by rear-view mirror and right-hand mirror (Figure 6). This figure shows clearly that areas that involve extreme head turning, looking at a passenger or around the interior of the vehicle are relatively uncommon and represent the lowest proportion of glances. Data for Figure 7 is for all participants combined but data shown in Figure 4 indicates that there is variance between different drivers. Figure 8 shows the eye glance data for each participant.

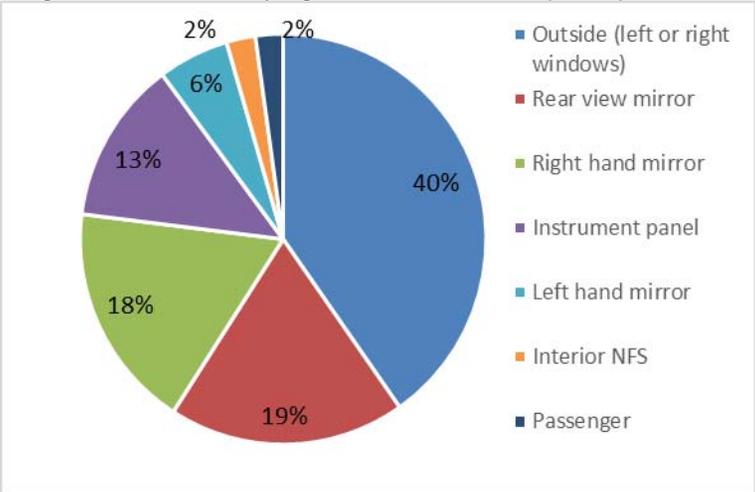


Figure 7 Proportion of glances to different areas during driving excluding 'forwards' (aggregated).

Data for some participants illustrates (Figure 8) that their glance behaviour could influence their head position much more than for other participants. For example participant 4 exhibits a larger proportion of glances towards the passenger and interior which indicates more extreme head turning, whereas participant 7 distributes glances towards objects with lower head rotations such as the rear-view mirror, right-hand mirror and instrument panel.

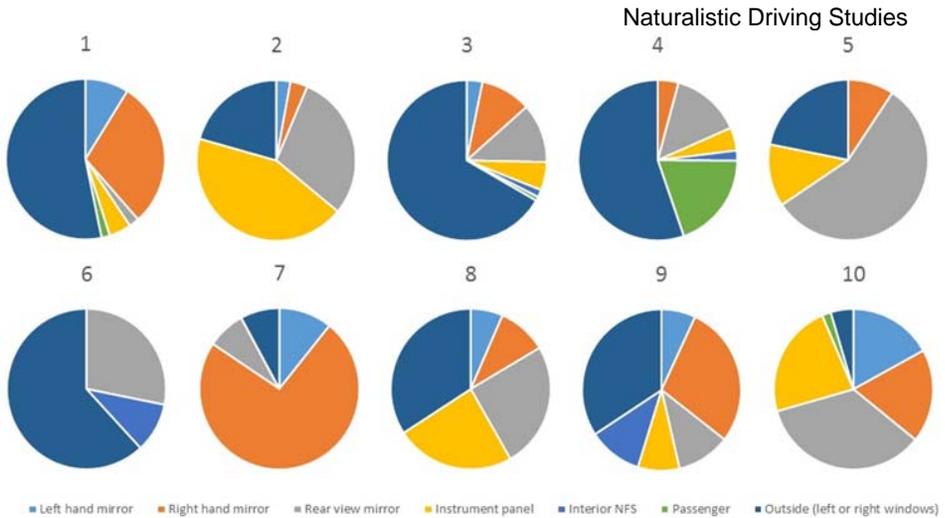


Figure 8 Proportion of glances to different areas during driving excluding 'forwards' (by participant).

Data presented in this study is for periods in the video data where the vehicle was moving as this was deemed to be the condition under which distraction (or 'eyes off road' glances) was most risky. Glance analysis was also conducted for periods where the vehicle was stopped and data for this shows some small but possibly significant differences. Although average glance duration of around 0.8 seconds (to all locations combined) was only around 14% greater than the average glance duration in the moving data (0.7 seconds) the clearest change between the moving and stationary periods is the increase in longer glances (greater than 1 second). These longer glances also tended to be to areas outside of the side windows or to the passenger; both of which involved more extreme head turning.

3.2 Digital analysis of eye-tracker readings

The results in this section report the outcome of analysis of the eye-tracker digital readings of nine subjects. The duration of nine driving trials is shown in Figure 9. Periods of missing readings (when the eye-tracker was not able to fix on facial features to assess head position) are outlined at the top of each bar and shaded in yellow where the video was reviewed manually to identify occupant behaviour. The proportion of missing readings ranged widely from almost negligible in case 2 to over half in case 6.

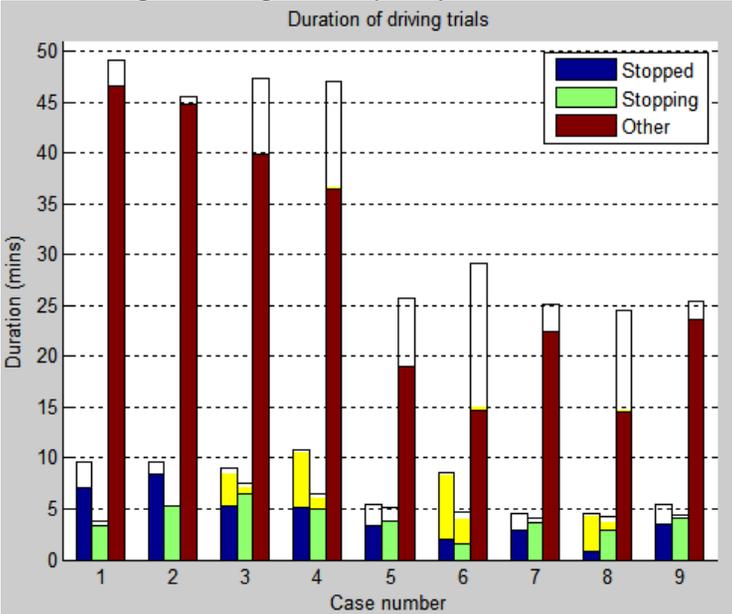


Figure 9 Duration of driving trials and vehicle movement (by participant).

Approximately 23 minutes of video were manually reviewed for the four drivers with the highest proportion of missing readings while their vehicle was stopped or stopping (Figure 10). This video review clarified the activity of drivers during the periods of missing data within the resources available for the work. Two types of activity were observed to provide the main explanation for the missing data: firstly, rotation of the head beyond the measurable range of the eye-tracker and, secondly, rotation of the head rapidly from side to side, not necessarily beyond the range of measurement of the eye-tracker, but too fast for it to maintain continuous, real-time image processing. These are described as 'Extreme head turning' (7 minutes) and 'Repeated head turning' (13 minutes) in Figure 10. The explanation for missing readings in the remaining 2–3 minutes was either 'Other types of head movement' or 'Unknown'.

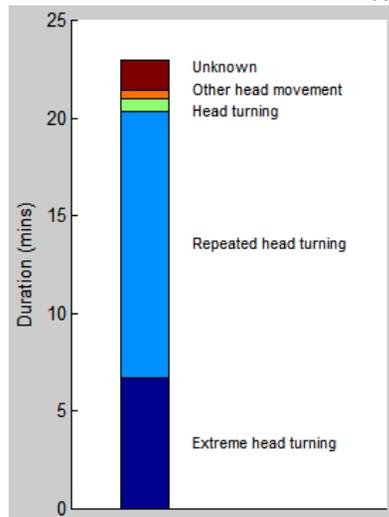


Figure 10 Activity of drivers obtained while vehicle stopped or stopping and digital instrument readings not captured.

Figure 11 show a rare period of perfect eye-tracking recording for a single subject in a continuous transition from one manoeuvre to another. Each glance target has a characteristic shape: instrument panel ('Inst'), internal rear-view mirror ('Mir C'), external right mirror ('Mir R'), front passenger ('Face') and the left and right exterior ('Ext L', 'Ext R'). The upper trace showing head rotation is synchronised with the lower trace showing eye-gaze movement. Small breaks in the eye-gaze traces accurately record eye blinks, interestingly on the return from glances to the rear-view mirror.

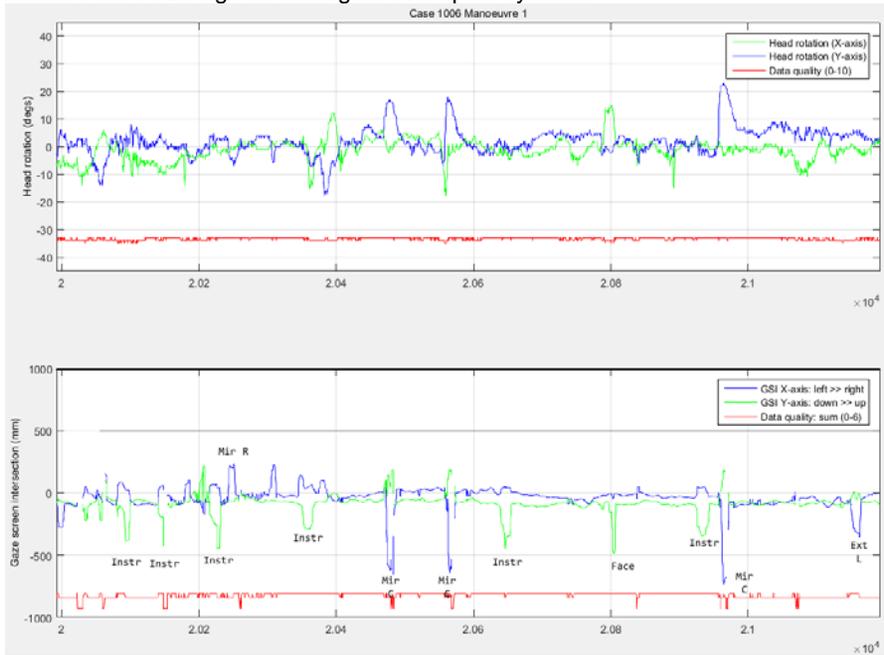


Figure 11 Successful tracking of eye-glance movement (subject 1006).

The eye-tracker featured an in-built function for assessing the quality of its readings based on real-time image processing. In Figure 12 most of the readings received the highest gaze quality assessment (red line) but in fact bore no useful relationship to reality.

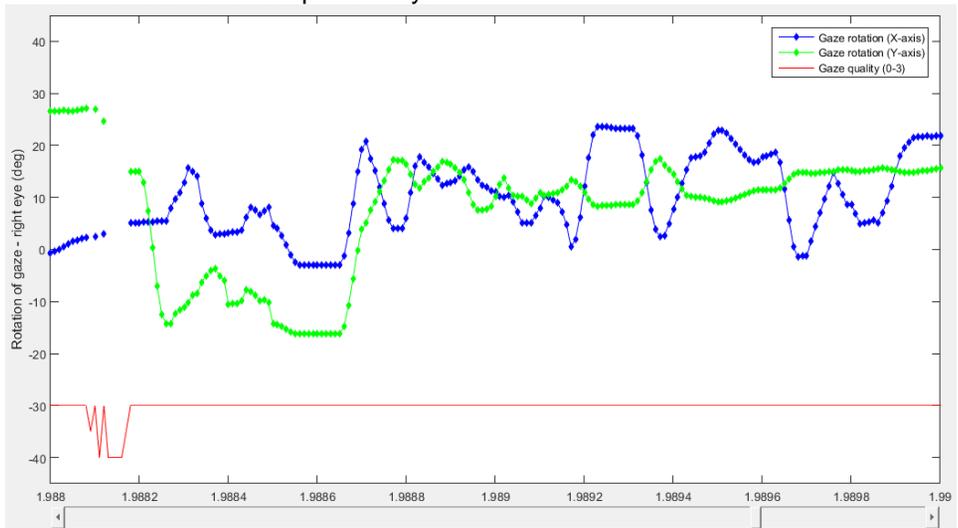


Figure 12 Tracking of eye-glance movement reported as successful but not actually meaningful (subject 2063, sample 1).

Digital readings were missing for many periods of recording. Figure 13 shows

an example of readings that were too intermittent to be useful for identifying eye-glance targets.

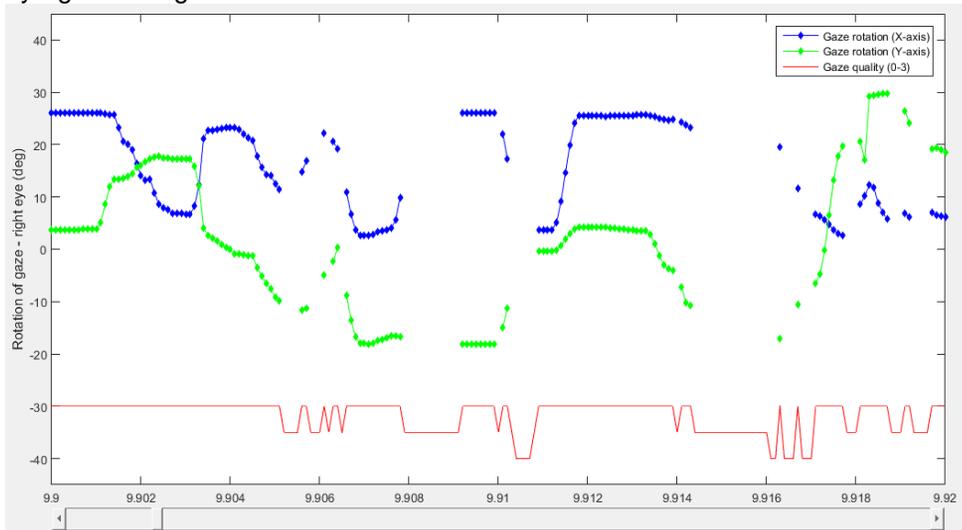


Figure 13 Intermittent recording of eye-glance movement (subject 2063, sample 2).

4 DISCUSSION

The video review data showed that for majority of the periods analysed, drivers looked ‘forward’ to engage in driving. On average, the participants had their eyes off the normal activity (looking forward) for only 7% of the total test duration, the highest being 13% and lowest 4%. The manual video review process provided good quality information, however it was very time consuming and tedious.

Unlike the digital data collected the video data provides almost 100% coverage of the interior occupant views and exterior contextual views. Although the analysed data consists of only selected sections of a much larger trial (approximately 10 minutes of analysed data from a trial lasting over one hour), these were not selected for the quality of the data but for the road type and layout they contained. This selection methodology indicates that it is likely that the remaining unexamined video data contains similarly good quality glance behaviour to that seen in the examined sections.

Deriving data from video, so called data reduction, is a relatively mature science that is well understood. One aspect of this understanding is that it is

very time consuming and the TeleFOT study was no different in this respect. Each data set from each participant contained over an hour of data of which only around ten minutes was analysed; to analyse this ten minutes and reduce the video down into a form which can be easily used took around one working day (or approximately ten hours). Understandably, issues with analysing larger data sets in this way leads to huge time constraints and many working days to achieve.

One limitation to the approach of analysing glance behaviour from video is that it provides data that is reliable but does not necessarily provide data that is of high clarity. Data for glances towards the forwards roadway for example will be very reliable—i.e. the analyst will be very sure the driver was looking forwards—however it might not be possible to determine whether that forwards glance was to the right or left of centre or what further information such as the drivers head position during the forward glance.

Analysing the TeleFOT data indicates that drivers do control their glance behaviour. Observed glance patterns for periods where the vehicle was moving differed from the patterns where the vehicle was stopping or stationary. In this latter period glances tended to be towards areas that were out of the side windows with more extreme head movements or they tended to be of much longer duration. This indicates that as the driver begins to bring the vehicle to a stop and subsequently when the vehicle is stationary, they perhaps see less risk in this glance behaviour.

Missing data from ADSEAT is predominantly in the 'repeated head turning' or 'extreme head turning' categories and is a result of the technical limitations of the eye tracking equipment. It is likely that this missing data is particularly associated with periods where the vehicle is stopping or stationary as observed in the TeleFOT video analysis.

The eye-gaze readings shown in Figure 11 were quite exceptional for their quality and rarity in the ADSEAT and TeleFOT studies. Subject 1006 recorded several minutes of perfect data under demanding driving conditions that included fluctuating light intensity, large movements of the head and upper body, and vehicle acceleration, braking and turning. Some similar

quality data was found for a couple of other subjects but only for a few seconds. In these cases the good readings occurred at the beginning of the driving trial. Most of the eye-gaze data for the tens of hours of trials was either (a) missing or excessively intermittent (Figure 13) or (b) impossible to relate to reality despite being reported by the eye-tracker device as being of good quality (Figure 12). The conditions that disrupted eye-gaze recording are not fully understood, nor is it known why no data of intermediate quality appeared—it was either perfect or entirely unusable.

The recording of head position and rotation from the eye-tracker relied on recognition of facial features but, unlike gaze direction, did not depend on real-time image processing of the eye, particularly the iris and pupil. Head position readings were far more robust than eye-gaze readings and supported a substantial analysis of digital data for the ADSEAT project. A corresponding analysis of digital eye-gaze data as an indicator of driver distraction could not however be carried out for the TeleFOT project.

5 CONCLUSION

Manual video review produces results but demands high resources (time, labour, cost). Digital processing potentially automates the analysis but perfection of experimental techniques or application of newer technology required to obtain suitable data.

6 REFERENCES

- [1] ADSEAT project, <http://www.adseat.eu/>.
- [2] TeleFOT project. <http://wiki.fot-net.eu/index.php?title=TeleFOT>.