Comparing the Effects of Integrated and Nomadic Navigation Systems on Road Traffic Safety

SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE OF MASTER OF SCIENCE

Lucas R. Johnston
10564888

MASTER INFORMATION STUDIES
Information Systems

FACULTY OF SCIENCE
UNIVERSITY OF AMSTERDAM

Submitted on 16.06.2023

1st Examiner
Dr. Frank Nack
Informatics Institute, University of Amsterdam

2nd Examiner
Dr. Hamed Seiied Alavi PhD
Informatics Institute, University of Amsterdam

ABSTRACT

To write.

Repository: https://github.com/lrjohnst/master-thesis-is.

Keywords:

1 INTRODUCTION

~~Recent years,~~ progress of technology has brought smartphones and other extremely versatile devices to the dashboards of cars. These devices can be quite helpful, such as voice assistants and navigation systems. Applications may also have practical uses not related to driving, like messaging and dialing. Additionally, some applications have purely the function of entertainment. These developments have had an impact on the way people drive and subsequently the safety of driving and traffic as a whole. Most countries have a ban on non-handsfree telephone use while driving. Given almost 160-thousand violations of this ban in the Netherlands in 2021 [1], it can be established that many drivers have a tendency to get distracted by their phones while driving. Being distracted by for instance the car radio, intense emotions, or using a smartphone decreases the focus toward the rest of traffic, and increases the probability of disrupting traffic or causing an accident. A large scale 2019 naturalistic study in the USA by Dingus et al [2] reports a 3.5 odds ratio of getting into a car crash while using a cell phone, over a baseline of driving without distractions.

In a sense, the car has become an information unit. Traffic today without information distribution is nearly unthinkable. There are various data the driver needs, delivered by assistance and control systems. There are also data that the driver does not primarily need, but is nevertheless provisioned such as entertainment and communication. Car infotainment systems divided by Kandemir et al into Nomadic (external, such as smart phones) and integrated devices [3] consist a broad range of applications like messaging, radio and navigation. Modern cars contain a larger variety of infotainment systems. These systems by themselves change rapidly, for example further integration of smartphones with native car infotainment systems, adding to the variety, versatility and complexity of tasks related to these systems. Also the interactions themselves have changed. Notably, many existing or new functions of car infotainment systems are controlled by touch screen, where before this may have been done by knobs and buttons.

One issue that drivers encounter is the potential for an excess of information provided during certain stages of driving that may not be relevant or necessary for the task at hand. Studies have shown that drivers can become overwhelmed when presented with too much information, leading to increased stress, cognitive load, and ultimately reduced safety [4][5]. Navigation systems may have a positive effect on traffic safety, as it prevents unnecessary searching and detours, but under condition that destination is entered into the system before starting to drive [6]. While navigation can be a necessary and useful tool to assist drivers in reaching their destination, the selection, presentation and timing of this information can be critical to avoid information overload and distraction. Therefore, finding the right balance between providing necessary information and avoiding unnecessary distractions is crucial for ensuring safe and efficient driving.

This study hypothesizes that the use of smartphones for navigation introduces numerous distractions, such as pop-ups and notifications, and these systems are not specifically designed with traffic safety in mind, posing greater danger to drivers compared to navigation systems that are integrated into the car. In this study, the impact of using an integrated navigation system on road safety was investigated in comparison to using a smartphone for navigation while driving. The results of this study poe rovide empirical grounding for future designs of car navigation systems and related regulations, ultimately to improve traffic safety. The following questions are formulated to which the answers provide these insights.

1. What are the specific distractions introduced by smartphone navigation systems that impact road safety?
2. What are the specific indicators of road safety that are relevant to the use of navigation systems?
3. How can these indicators be ranked in terms of their importance for evaluating the safety impacts of nomadic (smartphone) versus integrated car navigation systems?
4. Is there a statistically significant difference in specific road safety indicators between drivers using smartphone navigation systems versus those using integrated navigation systems?
5. How do nomadic (smartphone) navigation systems and integrated car navigation systems differ in terms of their impact on road safety indicators?
6. How can the results of this study be used to inform the design of future car navigation systems and regulations around their use to improve road safety?

Altogether this research will provide an answer to the following main research question: What is the impact of smartphone navigation systems versus integrated car navigation systems on road safety, specifically in terms of the distractions they introduce, relevant indicators of road safety, differences in impact on road safety indicators, and implications for the design of future car navigation systems and regulations around their use?

2 RELATED WORK

This section aims to provide an overview of previous studies on navigation-assisted driving and its impact on road safety, including the types of distractions and interface design, as well as the indicators of road safety used in previous studies.

Grahn and Kujala conducted a study in 2020 that aimed to compare the degree of visual distraction caused by smartphone-based applications to that caused by a specialized application for cars (Carrio). The study involved two different experiments (n=97) conducted in a driving simulator. Visual distraction was measured in terms of distance driven with occluded vision (occlusion distance). According to their findings, the specialized application caused less visual distraction due to its specialized user interface design, the division of tasks into subtasks, and, to a lesser extent, the size of the screen ~~[1]~~. The same study found that task structure, specifically how tasks are divided into subtasks, is important. People tend to switch tasks at subtask boundaries, such as between words, and this has implications for reducing distraction in car information systems [1].

In an analysis of distraction by car infotainment systems, a team from University of Utah tested differences in cognitive load between various functions and interface components for car infotainment systems (IVIS) [2]. In this naturalistic study (n=120), distraction was measured using the ISO standardized Detection Response Task (DRT) measure, and by a set of subjective measures, gained by a questionnaire after each driving session. The research found significant differences between various applications (like navigation, entertainment, messaging or dialing) and various components of user interfaces with respect to driving performance. Simulation studies such as Grahn and Kujala [1] or Jun Ma [3] are the most frequently used methodology in similar studies, followed by naturalistic studies as the second most common approach [4].

2.1 Types of distractions

Use of navigation systems may introduce various types of distractions, classifiable as cognitive, visual, manual, and auditory [3][5], which can impact driving performance and road safety.

Numerous of the reviewed articles focus on visual distraction, which tends to be measured in terms of duration or frequency of glancing, or similarly fixation count or duration. ~~Visual distraction seems to be highly represented in previous studies into distracted driving.~~ It is also a core concept in the US National Highway Traffic Safety Administration 2013 driver distraction guidelines for in-vehicle electronic devices [6] which as reported by Kujala and Salvucci suggests three main guidelines to minimize: (1) individual glance duration, (2) mean glance duration, (3) total glance time [7]. Additionally, it should be noted that the same study notes that glancing and visual distraction are not necessarily equivalent.

While cognitive distraction by itself is difficult to measure, the adverse effect it has on driving performance has been observed in lab studies ~~[8]~~. A benchmark of four measures of driver workload by McDonnell et al. observed Task Interaction Time to be most sensitive to work load differences between 40 tested cars, followed by DRT Miss Rate, NASA-TLX and DRT Reaction Time. Furthermore: the latter two measures were found to require a sample size larger than the sample size in their study (n=173) to have sufficient power [8].

While voice control allows the driver to keep the eyes on the road, a tradeoff is that voice control tends to cause higher cognitive load compared to manual interaction (excluding touch screen). Steering wheel button control in combination with voice control have been found a beneficial combination for the more basic tasks [3]. Mitigating high cognitive load by full text visual feedback in turn causes high visual load and time pressure, which in turn may be mitigated by visual feedback in the form of keywords and icons [5].

2.2 Relation between interface design and driving performance

Comprehensive literature review by Oviedo-Trespalacios, et al resulted in an extensive list of secondary in-vehicle tasks such as conversing, reaching, answering calls, dialling, browsing, reading, texting and typing [4]. These IVIS-related tasks mentioned in this study can be considered unhelpful distractions, together with adjusting the radio, entertainment systems, dealing with irrelevant (navigation) data, and specifically to nomadic systems: popups and notifications by for instance social media or disruptions and interruptions of the navigation application.

User interface design for cars has been mentioned as ‘a community’ [5], pointing at the fact that it has a certain maturity as a field of study. This literature review has observed improvement of car safety to be a key driver of this field. Car navigation systems for consumers have been around since Mazda introduced them in 1990, at the time as a system integrated with the car [9]. Since, the market has additionally seen dedicated navigation devices (like a Garmin or TomTom device), smartphone navigation apps (like Google Maps), and more recently, the linking of smartphone navigation apps to the car IVIS for instance by cable or Bluetooth. At least iOS (Apple CarPlay) and Android (Android Auto) currently support such features. The latter mentioned feature may be considered a sort of hybrid between nomadic navigation and navigation via the car integrated IVIS. It has the advantage of staying up to date automatically, contrary to other dedicated navigation devices or integrated IVIS navigation applications, which must be updated manually. Given that not everyone updates their navigation system [9], and given that an updated navigation system improves the user experience and potentially even safety, the mentioned hybrid system potentially has an advantage.

Interaction modalities: Haptic feedback can help alleviate visual distraction and allow the driver to focus on the road [10]. Audio feedback, such as a "read aloud" feature, can also be helpful, although it may not be as effective in some situations and can still cause cognitive distraction [1]. Different input modalities for certain tasks, or different mixes of modalities are likely to have an effect on cognitive, visual or manual distraction. A 2022 study by Jun Ma et al suggests a well-designed touch screen may be more suitable for certain complex secondary tasks, compared to knobs and buttons, despite the fact that knobs and buttons are by themselves more simple to operate [3].

Multiple studies have identified navigation destination entry as highly demanding [10][11], and in at least two instances it was even identified by direct experiment as the most demanding secondary task [2][3] among other common tasks such as text messaging, dialing and radio volume adjustment.

2.3 Driving performance indicators

Besides in-vehicle tasks, Oviedo-Trespalacios created an inventory of “Human Machine Systems” (HMS) performance metrics: headway, lateral position (lane position), speed, crashes, and workload [4]. The mentioned metrics may be considered synonymous, or closely related to driver performance indicators.

Analysis reveals that the design of the IVIS interface affects driving speed. Engaging in activities such as conversing, dialing, or texting while driving leads to a decrease in driving speed and an increase in headways [4]. This is a well-known effect and named by Young and Regan as “compensatory or adaptive behavior” [11]. Lane position has been found to be impacted by visual and manual load. Also voice control that generates cognitive load has been found to affect departures from the lane center (more so than on speed control). Still voice control seems to distract less than operating a touch screen. [3]. Furthermore, it has been shown that voice control with full text visualization leads to higher headway variability, attributed to higher total glance durations [5].

The positive association between secondary tasks while driving and decreased driving performance seems to be moderated by environmental factors that impact the complexity of driving tasks [12]. Also minding the interdependencies of distraction variables, Kandemir, et al. propose the existence of “toxic” task combinations in which certain tasks, while not overly burdensome on their own, may surpass a certain threshold when performed in conjunction with more complex tasks, such as dialling while simultaneously braking at a red light [13]. In a similar sense, Oviedo-Trespalacios have approached what they called “Mobile Phone Distracted Driving” as a human-machine system. They have focused their observations not just on distractions by certain tasks, but also by conflicts that occur between combinations of tasks [14].

2.4 Wiener Fahrprobe

Talk about WFP methodology here.

3 OVERVIEW OF METHODOLOGY

The data collection and analysis consisted two main phases, first a requirements elicitation phase in which a theory was formulated about relations between distraction while driving and navigation systems. The requirements elicitation phase consisted a survey (section 4) and an expert interview (section 5). In this theory, specific indicators of driving performance as well as common distractions relevant to the use of navigation systems were identified, as well as variables or contextual factors that may influence the relation between distractions and driving performance. Secondly, this theory was deductively tested by applying the Wiener Fahrprobe method in a naturalistic driving experiment.

F	Full sample including S, I and Z, dedicated and other
S	Sample of smartphone navigation app users
I	Sample of integrated car navigation system users
Z	Sample of streamed navigation from smartphone to board computer
f	Frequency
$\bar{x}$	Sample mean
α	Significance level, set at 0.05 for all tests

Table 1: Variable names used throughout this document

4 SURVEY

4.1 Goal of the survey

A survey was conducted with the purpose of attaining first hand data about navigation assisted driving behaviour. This data ~~would~~ then ~~be~~ analysed to be able to describe the way people use navigation systems, how they might or might not get distracted by them, what the role of the navigation interface would be, and how potential distractions might affect driving performance.

4.2 Survey design

The survey was conducted online, was anticipated to take 10 to 15 minutes to complete and no incentive was offered to the ~~anonymous~~ participants. The participants answered questions concerning: (1) their use or of navigation systems while driving (e.g. how, what for, preferences, frequency), (2) how they may or may not have found they were distracted by navigation systems, (3) how have they found those distractions to impact their driving performance, (4) how distractions may be linked to navigation user interface. The resulting raw qualitative data was the main source of indicators such as events, distractors and their perceived effects on safety. A data dictionary containing the original Dutch survey questions, their associated data types, and a concise reference string for each question is available: https://github.com/lrjohnst/master-thesis-is/link-to-file. The data dictionary provides an overview of the survey questions and their respective variables.

At first, questions ~~are asked~~ to characterize the participant, for instance by frequency of car use or type of navigation system used. An introductory question was asked to get the participant to think about the topic. The participant was then asked to name up to five distractions specifically related to their most used navigation system. This open elicitation was followed by a five-point Likert-scale question asking the participants to rate in how far they felt they a defined distraction (based on distractions learnt from literature) occurred. The next two questions were similarly first an open question to name five negative effects of navigation-related distraction, followed by a Likert-scale mentioning common scenarios from the literature. The intention with these four questions was to look for potential relations between navigation related distraction, and driving behaviour. The same structure (an open question and then a Likert-scale question was then applied to gather data about perceived "bad instructions, and "interruptions". The survey concluded with two open questions about (1) what the participant would like to see changed with regard to the interface of their navigation system, and (2) what the participant would like to further mention.

4.3 Sampling

The larger proportion of participants to the survey were recruited from the network of the researcher and a smaller proportion by distribution of flyers at fueling stations in and around Utrecht, Netherlands. A total 80 of people participated, 13 of whom were disqualified due to not owning a driver's license, not using any navigation system or incomplete answers. The first survey response was on April 29 and the last response was on June 8, with a mean completion time per participant of 1086 seconds. Tables UV and XU describe the reported used navigation systems among participants. The overall mean frequency of weekly car use is displayed in table TC. A Mann-Whitney test suggests there is a significant difference between the means of groups *S* and *I* (p=0.0327), possibly increasing a confounding effect of variables.

Navigation system type (n=67)	*f*	*f/n*
Smartphone navigation (S)	33	0.49
Car-integrated (I)	16	0.24
Stream from smartphone to board computer (Z)	15	0.22
A dedication navigation device	2	0.03
Another navigation system	1	0.01

Table UV: Frequencies navigation system type used

Navigation app used (n=33)	*f*	*f/n*
Google Maps	25	0.76
Apple Maps	5	0.15
Waze	2	0.06
Flitsmeister	1	0.03

Table XU: Frequencies navigation apps used

	F	*S*	*I*	*Z*
Mean	4.2	3.8	5.1	4.3
St. Dev.	1.9	2.0	1.8	1.6

Table TC: Descriptive statistics of reported weekly frequency of car use.

4.4 Survey set-up

The survey was built, visually designed and published by use of an online tool named SurveyLegend (surveylegend.com). Filling in of the survey by participants and exporting the data took place on SurveyLegend. Participants were referred to the survey by either a shared URL, or by QR-code. Participants were explicitly told the survey results were anonymous, which was honored among other measures by restricting recording of participant IP-addresses. The final resultset was downloaded as CSV-file on 31st May 2023, to be edited using Excel and subsequently analysed by use of Excel and PHP and Python based scripts written by the author.

4.5 Analysis methods

For the four open questions (1) "Name five distractions while driving related to your navigation system" (nav_distraction), (2) "Name five negative effects on your driving performance resulting from navigation system related distractions" (nav_behavior), (3) "Name a few examples of unhelpful information or instructions by your navigation system" (bad_instructions), and (4) Name a few examples of interruptions of your navigation system while you are trying to use it" (interruptions), participants had ~~five optional~~ text fields to fill in. Open coding was performed on the answers. While coding the researcher made the translation from the mostly Dutch responses to English codes. Specifically for 'nav_distraction' a second coding session was done to refine insights after the first coding session. For 'nav_behavior', 53% of responses were codable to terms of driving behaviour, but 47% did not and were discarded. Only the results of the second coding session of 'nav_distraction' were used in further analysis. Following the coding, frequencies of each response code were determined per group S, I and Z and relative frequency (f/n) was calculated as the ratio of the observed frequency to the total number of codes assigned to the particular group. Additionally, in order to determine if there were significant differences in the reported codes, a series of one-way chi-square tests was conducted for each combination of the variables S, I, and Z with individual codes. To address differences in sample sizes, weighted frequencies were used instead of absolute frequencies. Weighted frequencies were calculated by dividing the absolute frequency by the number of participants in each group and then multiplying this value by the number of participants in the smallest group (f(Z)=15). This conservative approach reduced the risk of extrapolation or inflated values based on a smaller sample. By employing the smallest group as the weighting factor, the analysis aimed to enhance rigor and mitigate biases arising from uneven sample sizes. It ensured that the weighted frequencies were proportionate to the number of participants, thereby accounting for the varying sample sizes across the groups. In cases where the resulting p-value turned out higher than the predetermined significance level (α=0.05) or the expected frequency turned out smaller than 5, the corresponding result was disregarded. The tests aimed to assess whether the distribution of codes varied significantly across the different groups. For each of the 15 individual Likert-scale questions, the results were coded into numbers starting from 1 (never applies), up to 5 (always applies) and descriptive statistics were compiled. The 15 result sets were tested for equalness of variance by F-test, normalcy by Shapiro-Wilk test and means across all permutations of the 15 variables were compared by Mann-Whitney test. T-tests were performed but disregarded as sample sizes were too small in most cases and only 2 out of 15 variables had normally distributed frequencies. Furthermore, Pearson's correlation coefficients (r-values) were calculated for all combinations of the 15 Likert-scales, within groups F, S, I and Z (appendix **HV**). For each resulting r-value, a confidence interval (CI) on 95% confidence level was calculated using the Fisher transformation method [**literature**]. Each variable pair that corresponded to a CI lower bound value equal to or greater than 0.3 was analysed further by drawing correlation diagrams for each group F, S, I and Z (appendix **HB**).

Also talk about the coding procedure for the last few open questions: 'desired design changes' and 'further remarks'.

And now describe how the results of these previous steps were used to formulate a theory to test in the field.

4.6 Findings

Examining figure **PA** about the responses to the open question 'nav_distractions', notable differences in relative frequency can be observed, only the first of which statistically supported. Most notably and firstly it is observed that compared to group Z and S, more group I participants report distractions related to "Bad instructions or difficulty interpreting". This observation is supported by chi-square tests: p(S,I)=0.0234 and p(I,Z)=0.0056. This indicates that users of car-integrated navigation systems likely more often report cases such as "wrong route suggestions", "system outdated", or "instructions do not take into account maintenance or traffic". Secondly, distractions related to code "Conflict between other system and navigation" appear to be reported less for group I compared to the other groups. This suggests that users of streaming or smartphone based navigation systems more often reported conflicts such as: "calling interferes with navigation", "using other apps while navigating", or "other apps overlay my navigation system", the latter indicating instances where another app takes precedence while driving and using navigation. Thirdly, cases related to "Navigation interferes with driving tasks" seem to be reported less frequent within group Z, compared to group I. This suggests that users of navigation systems streamed to board computer from smartphone report more often cases such as: "Navigation requests input", "Racing against ETA", or "Too much glancing at navigation screen". Fourth and finally, group Z participants seemed to complain more often about cases related to "Route changes or suggestions", compared to group I. This would suggest that streaming navigation system users report more cases such as "Sudden rerouting", "Alternative routes popping up", or "Navigation suggests undesired alternative route".

For the open question "Identify 5 negative effects on your driving behaviour resulting from distraction caused by your navigation system" (nav_behavior), the relative frequencies of codes are represented in figure **NM**. Most notably and firstly it is observed that participants in group I report more negative effects related to "Decision making", compared to groups S and Z, which is supported by results of chi-square tests: p(S,I)=0.0239 and p(I,Z)=0.0288. Therefore it seems likely users of car-integrated navigation systems more often report cases like "Becoming insecure / desoriented", "Missing an exit", or "Making illegal or dangerous move". Secondly, there seems to be a notable difference between group S and groups I and Z regarding negative effects related to "Lane position". This suggests that smartphone navigation system users would more frequently report cases such as "Car follows my gaze to the nav on left/right", or "Swerving". **This is where I left off for now…**

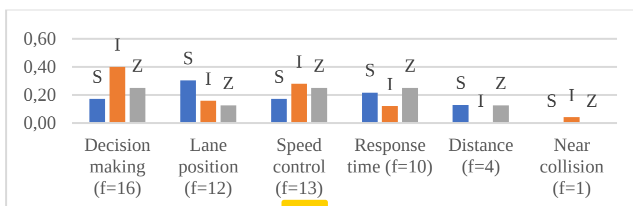

Figure NM: Relative frequency per code (x-axis), per group (S, I and Z) for reported nav. distraction-related behavior (nav_behavior), sorted descending from left to right by total number of occurrences (f).

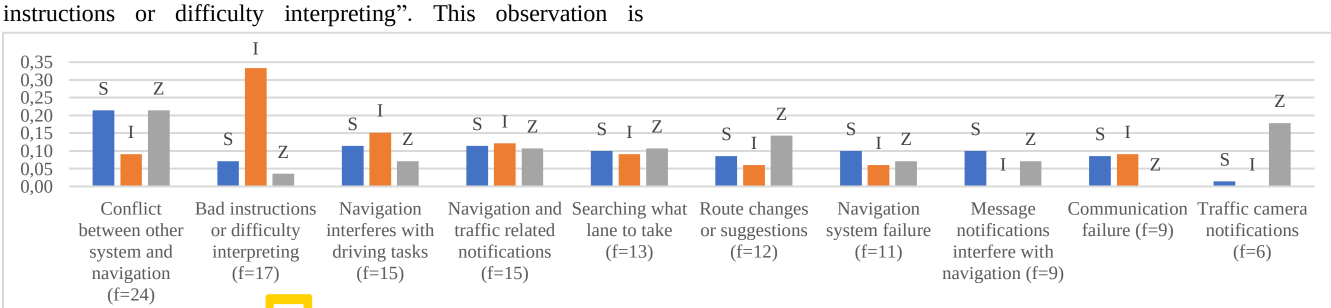

Figure PA: Relative frequency per code (x-axis), per group (S, I and Z) for reported navigation-related distractions (nav_distractions), sorted descending from left to right by total number of occurrences of the code (f).

No significant differences were found when comparing the means of the Likert-scale questions (Appendix HB) among the groups (S to I, I to Z, Z to S) using Mann-Whitney tests. However, three comparisons approached significance: 'distraction_awareness' between S and I (p=0.0504), 'behavior_wrong_turns' between Z and S (p=0.0635), and 'interruptions' between S and I (p=0.0694). Although these comparisons were not statistically significant, they were of interest due to their proximity to significance and their relevance to the study's focus on groups I and S. Consequently, the comparisons 'distraction_awareness' between S and I (p=0.0504) and 'interruptions' between S and I (p=0.0694) were selected for further investigation in the subsequent experiment.

5 EXPERT INTERVIEW

The second part of the requirements elicitation involved conducting and analysing the results of an expert interview, described below.

5.1 Goal of expert interview

To provide an expert's point of view and context to the data attained during the survey and literature review, but also to add new information, an interview was organised. One specific goal was to use the data from the interview to improve the accuracy of the theoretical model formulated during the requirements elicitation phase, for instance by enabling more accurate prioritization of driving performance indicators.

5.2 Interview design

The interview was conducted using a semi-structured format, guided by a list of topics shared with the interviewee two weeks prior. The following topic list was used: (1) distraction by navigation systems, (2) relations between navigation systems (specifically: interface and its surrounding context) and distraction, (3) effects on driving performance, (4) indicators of driving performance, (5) measuring driving performance, (6) how policy could be helpful. The topic list was designed to effectively dictate which topics would be discussed, but at the same time provide room for free interpretation by the participant, not to introduce unnecessary limits to the answers.

5.3 Participant profile

The interviewee is a scientist linked to the Institute for Road Safety Research (SWOV), experienced in analysing naturalistic driving data, more specifically with regards to distraction in traffic. ~~The participant mentioned an example where they were involved in~~ a large scale European study that had 120 cars fitted with cameras and sensors allowing them to collect approximately 100.000 hours' worth of driving data. Here, the participant focused on mobile phone use, while his colleague focused on navigation system use. The participant mentioned research that they would be starting, into Intelligent Speed Assistance (ISA), which is a part of many navigation systems. The hypothesis, as the participant described it would be that drivers drive faster on average with ISA, as they now have better overview on the current speed limit and are more confident getting closer to this limit.

5.4 Interview set-up

The idea to interview an participant from SWOV came up while reviewing work that it published **[literature]**. A request for contact and for the interview was made by email which eventually resulted in an appointment. The interview took place at the SWOV office in The Hague and lasted approximately 75 minutes. The interview was recorded by use of the author's smartphone, which was tested for this purpose, prior. The author started at the first item in the topic list and asked in-depth questions. Divergence of topic was allowed, but the author did steer the discussion back to the topic list at a certain point until eventually all topics on the list were discussed. Halfway during the interview, a few print-outs of preliminary data from the survey were discussed, which was unplanned, but deemed valuable as it would help guide the conversation.

5.5 Analysis methods

Before analysis, the interview audio recording was transcribed using ChatGPT whisper API [**L0AB**], with help of the OpenAI package for Python v0.27.0 [**L0AA**]. The transcription was then ~~relistened,~~ double checked, edited and coded manually by the researcher. The coding method used in the study involved a systematic process of analysing the transcript to identify and categorize relevant facts and statements related to specific topics. The author initially read the interview to gain a general understanding of the content. Then, during a subsequent reading, facts and statements pertinent to the topics of driving distraction, driving performance, policy or other related areas of interest were extracted. The extracted information was organized in a table, where each entry was annotated with the corresponding topic category. This coding process allowed for the identification and categorization of key information within the interview data, enabling the researcher to focus on specific themes or topics of importance for further analysis. The researcher utilized the coded data as a foundation to construct the summary (section 5.6), which aimed to provide a concise overview of the key findings and insights from the interview. This systematic approach ensured that the summary was grounded in the data and aligned with the specific themes of interest.

5.6 Findings

Visual distraction was identified as the most significant and dangerous form of distraction while driving. The interviewee mentioned that the probability of a car accident increases substantially after just two seconds of visual inattention. To measure and analyse these distractions and their effects, the Wiener Fahrprobe was proposed as a suitable method. Despite its reliance on subjective experiences, this approach offers the advantage of flexibility in capturing and describing unforeseen situations, and less resource intensive for a small scale naturalistic driving study.

Several key measures have been identified for evaluating driving performance, including SDLP (standard deviation from lane position), steering jerk, breaking delay, abrupt breaking, time to collision, time headway, post encroachment time, and speed control. The researcher noted that to get reliable data on these measures, large quantities of data are required, although slightly less so on SDLP. Task breakdown in driving assistant applications are more effective in safely allocating drivers' attention compared to tasks that impose a time constraint. Attention and cognitive workload are important considerations in driving, and the concept of the "bathtub curve" has been introduced to illustrate the relationship between workload and attention. It emphasizes the need to maintain an optimal level of task difficulty and task load to ensure the driver can direct sufficient attention towards the road.

Navigation systems play a helpful role in driving performance by alleviating search behaviour and allowing drivers to focus more on driving tasks as opposed to navigating. The participant stated there is likely not a strong relation between navigation and the workload being too high or too low. As driving assistance systems assume increasingly prominent roles, their impact on driver attention becomes a subject of investigation. This investigation includes understanding the "bathtub curve" and its implications for workload and attention. Future developments may involve navigation systems guiding drivers' roles on specific sections of the road, when driving assistants are switched on or off.

6 NATURALISTIC EXPERIMENT

6.1 Goal of experiment

The naturalistic experiment has the purpose of validating the theory that was formulated during the requirements elicitation phase of this study.

6.2 Experiment design

Describe which variables; overall conceptual setup; link back to research questions.

6.3 Sampling

6.4 Experiment set-up

More the practical aspects: tools and such.

As an especially important ethical consideration to this study to minimize risks to the participants, researcher and other traffic, the researcher has explicitly requested compliance with traffic laws at all times during driving sessions, and to always put safety first when making decisions while driving.

6.5 Analysis methods

6.6 Findings

7 DISCUSSION

The findings of this study provide insight into relationships between distraction while driving and different types of navigation systems. In this section context is provided to the findings by discussing their implications, as well as stronger and weaker points to their validity and reliability.

The analysis method for the coded open answers to “nav_distractions”, “nav_behavior”, “nav_bad_instructions”, and “nav interruptions” compared internal relative frequencies within groups, across groups. A drawback to this approach is that the internal relative frequencies may have provided a distorted view on the data, as differences across groups may have been inflated or otherwise made less visible, due to internal dominance of one, or a few other code(s) inside a group. Another potentially more valid approach could have been to assign participants a dichotomous variable containing whether the code applied or not. However the first approach was chosen as to make optimal use of the small sample, potentially trading off a certain level of validity.

Some points about potential weaknesses that I must process into narrative:

1. While coding I must admit that codes may not always be mutually exclusive. One case can be mentioned where codes are likely not mutually exclusive: for question 'bad_instructions', the two codes: 'bad timing or bad data' and 'confusing or conflicting'. Conflicting can be seen as a form of bad data, and also looking at the data it may be noticed that these overlap. As a result, confounding of variables may happen here.

2. There was only one interview. We cannot speak of saturation of results, or multiple experts confirming each other or debating each other. We have only one expert view. This missing of expert perspective is mitigated partly by triangulation between expert views, survey results and literature review.

3. The survey sample is quite small.

4. The survey sample is a 'convenience sample'.

5. Points 3 and 4 both mean I can by no means generalize to the Dutch population of car navigation system users.

8 CONCLUSION AND FUTURE WORK

The study aimed to examine the distinctions between smartphone-based car navigation systems and car-integrated navigation systems concerning their potential impact on driving performance, with a particular focus on the mediating role of distraction.

Additionally, this research sought to analyse the implications of these differences for the design of car navigation systems and the formulation of regulations pertaining to their future utilization.

9 REFERENCES

To compile.

[**L0AA**] https://pypi.org/project/openai Retrieved 2023-07-10

[**L0AB**] https://openai.com/blog/introducing-chatgpt-and-whisper-apis Retrieved 2023-07-10

APPENDIX HL: SURVEY CHANGE LOG

A change log was kept to log any changes made to the survey, displayed in table LM.

Time	*Submits*	*Event or change*
29/04 22:00	0	Finalized the survey.
30/04 11:00	4	Corrections spelling errors.
30/04 15:00	7	Added extra question at the end: 'would you like to participate in a follow up study? Leave your email address'.
01/05 12:45	18	Simplified phrasing of questions 8, 9 and 10 (includes the two matrix questions), after feedback that the questions are complicated.
08/06 15:55	80	The survey was made unavailable for more responses.

Table LM: Changes to the survey are displayed together with the time of the change and the number of submits at the time of the change.

APPENDIX HB: MEANS LIKERT-SCALE SURVEY QUESTIONS

	$\bar{x}(S)$	$\bar{x}(I)$	$\bar{x}(Z)$
distraction_manual	2,04	1,75	2,00
distraction_awareness	2,61	2,06	2,46
distraction_shift_focus	2,11	1,81	2,23
distraction_mental_load	1,79	1,56	1,85
distraction_glance_frequency	2,21	2,00	2,15
distraction_glance_duration	2,29	2,25	2,23
behavior_more_speed	1,81	2,00	1,45
behavior_less_speed	2,69	2,57	2,18
behavior_speed_control	2,04	2,50	2,00
behavior_lane_position	2,58	2,36	2,09
behavior_reaction_time	2,69	2,36	2,18
behavior_wrong_turns	2,81	2,50	2,00
behavior_operating_errors	1,88	1,71	1,55
bad_instructions	2,04	2,14	2,09
interruptions	2,08	1,57	1,64

Table UP: Means of answers to Likert-scale questions. Cells marked grey indicate mean comparisons for which a small likelihood exists that the difference is due to coincidence.

APPENDIX HV: LIKERT SCALE QUESTION CORRELATION MATRICES

	distraction_manual	distraction_awareness	distraction_shift_focus	distraction_mental_load	distraction_glance_frequency	distraction_glance_duration	behavior_more_speed	behavior_less_speed	behavior_speed_control	behavior_lane_position	behavior_reaction_time	behavior_wrong_turns	behavior_operating_errors	bad_instructions	interruptions
distraction_manual	X	0,48	0,62	0,25	0,54	0,45	0,24	0,43	0,19	0,34	0,50	0,30	0,36	0,14	0,10
distraction_awareness	X	X	0,50	0,52	0,47	0,33	0,42	0,29	0,27	0,39	0,57	0,31	0,25	0,13	0,22
distraction_shift_focus	X	X	X	0,48	0,54	0,56	0,31	0,30	0,17	0,35	0,54	0,42	0,28	0,21	0,15
distraction_mental_load	X	X	X	X	0,27	0,40	0,31	0,17	0,27	0,28	0,40	0,25	0,44	0,28	0,35
distraction_glance_frequency	X	X	X	X	X	0,54	0,12	0,33	0,12	0,39	0,42	0,34	0,34	0,10	0,10
distraction_glance_duration	X	X	X	X	X	X	0,23	0,35	0,12	0,43	0,33	0,31	0,43	0,19	-0,02
behavior_more_speed	X	X	X	X	X	X	X	0,31	0,57	0,46	0,19	0,23	0,35	0,22	0,09
behavior_less_speed	X	X	X	X	X	X	X	X	0,45	0,42	0,47	0,27	0,32	0,32	0,01
behavior_speed_control	X	X	X	X	X	X	X	X	X	0,60	0,19	0,34	0,22	0,42	0,01
behavior_lane_position	X	X	X	X	X	X	X	X	X	X	0,40	0,33	0,27	0,18	0,28
behavior_reaction_time	X	X	X	X	X	X	X	X	X	X	X	0,37	0,21	0,25	0,18
behavior_wrong_turns	X	X	X	X	X	X	X	X	X	X	X	X	0,32	0,29	0,23
behavior_operating_errors	X	X	X	X	X	X	X	X	X	X	X	X	X	0,31	0,17
bad_instructions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0,14
interruptions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Tabel QE: Correlation matrix Likert-scale questions Full set (F).

	distraction_manual	distraction_awareness	distraction_shift_focus	distraction_mental_load	distraction_glance_frequency	distraction_glance_duration	behavior_more_speed	behavior_less_speed	behavior_speed_control	behavior_lane_position	behavior_reaction_time	behavior_wrong_turns	behavior_operating_errors	bad_instructions	interruptions
distraction_manual	X	0,45	0,56	0,18	0,52	0,39	0,32	0,28	0,16	0,29	0,49	0,37	0,25	-0,11	0,10
distraction_awareness	X	X	0,44	0,52	0,38	0,15	0,59	0,08	0,45	0,43	0,40	0,39	0,24	-0,03	0,23
distraction_shift_focus	X	X	X	0,41	0,50	0,56	0,44	0,19	0,24	0,45	0,41	0,46	0,21	-0,01	0,14
distraction_mental_load	X	X	X	X	0,13	0,16	0,42	0,25	0,35	0,34	0,44	0,03	0,39	0,29	0,43
distraction_glance_frequency	X	X	X	X	X	0,38	0,06	0,06	-0,14	0,21	0,37	0,31	0,30	-0,07	0,14
distraction_glance_duration	X	X	X	X	X	X	0,24	0,16	-0,19	0,24	0,14	0,26	0,39	-0,07	-0,26
behavior_more_speed	X	X	X	X	X	X	X	0,32	0,58	0,68	0,35	0,20	0,36	-0,10	0,08
behavior_less_speed	X	X	X	X	X	X	X	X	0,47	0,35	0,53	0,41	0,51	0,37	0,07
behavior_speed_control	X	X	X	X	X	X	X	X	X	0,66	0,37	0,35	0,16	0,31	0,04
behavior_lane_position	X	X	X	X	X	X	X	X	X	X	0,31	0,32	0,11	0,09	0,28
behavior_reaction_time	X	X	X	X	X	X	X	X	X	X	X	0,52	0,24	0,28	0,10
behavior_wrong_turns	X	X	X	X	X	X	X	X	X	X	X	X	0,16	0,13	-0,07
behavior_operating_errors	X	X	X	X	X	X	X	X	X	X	X	X	X	0,28	0,01
bad_instructions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-0,06
interruptions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table QR: Correlation matrix Likert-scale questions Smartphone (S).

	distraction_manual	distraction_awareness	distraction_shift_focus	distraction_mental_load	distraction_glance_frequency	distraction_glance_duration	behavior_more_speed	behavior_less_speed	behavior_speed_control	behavior_lane_position	behavior_reaction_time	behavior_wrong_turns	behavior_operating_errors	bad_instructions	interruptions
distraction_manual	X	0,04	0,40	0,04	0,31	0,14	0,34	0,46	0,45	0,31	0,36	-0,06	0,48	0,31	-0,28
distraction_awareness	X	X	0,21	0,54	0,31	0,25	0,36	0,33	0,27	0,20	0,42	0,00	0,09	-0,04	-0,03
distraction_shift_focus	X	X	X	0,52	0,58	0,67	0,34	0,64	0,47	0,51	0,60	0,28	0,16	0,21	-0,14
distraction_mental_load	X	X	X	X	0,50	0,74	0,00	0,16	0,07	0,13	0,53	0,50	0,26	-0,18	0,20
distraction_glance_frequency	X	X	X	X	X	0,77	0,24	0,70	0,40	0,30	0,54	0,40	0,00	0,13	0,00
distraction_glance_duration	X	X	X	X	X	X	0,22	0,44	0,37	0,22	0,60	0,50	0,17	0,18	0,15
behavior_more_speed	X	X	X	X	X	X	X	0,50	0,75	0,21	0,00	0,19	0,22	0,74	0,00
behavior_less_speed	X	X	X	X	X	X	X	X	0,70	0,34	0,40	0,28	0,05	0,39	-0,36
behavior_speed_control	X	X	X	X	X	X	X	X	X	0,42	0,21	0,47	0,12	0,83	0,18
behavior_lane_position	X	X	X	X	X	X	X	X	X	X	0,69	0,18	0,06	0,03	0,03
behavior_reaction_time	X	X	X	X	X	X	X	X	X	X	X	0,07	0,24	-0,24	-0,20
behavior_wrong_turns	X	X	X	X	X	X	X	X	X	X	X	X	0,12	0,42	0,53
behavior_operating_errors	X	X	X	X	X	X	X	X	X	X	X	X	X	0,19	0,06
bad_instructions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0,32
interruptions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Tabel QW: Correlation matrix Likert-scale questions Integrated (I).

	distraction_manual	distraction_awareness	distraction_shift_focus	distraction_mental_load	distraction_glance_frequency	distraction_glance_duration	behavior_more_speed	behavior_less_speed	behavior_speed_control	behavior_lane_position	behavior_reaction_time	behavior_wrong_turns	behavior_operating_errors	bad_instructions	interruptions
distraction_manual	X	0,67	0,84	0,43	0,65	0,69	-0,02	0,69	0,19	0,51	0,58	0,34	0,46	0,51	0,18
distraction_awareness	X	X	0,76	0,58	0,62	0,57	0,28	0,55	0,09	0,36	0,81	0,33	0,37	0,55	0,15
distraction_shift_focus	X	X	X	0,50	0,66	0,62	0,10	0,60	0,00	0,40	0,80	0,44	0,40	0,70	0,34
distraction_mental_load	X	X	X	X	0,37	0,64	0,71	0,12	0,33	0,50	0,34	0,61	0,76	0,73	0,64
distraction_glance_frequency	X	X	X	X	X	0,67	0,17	0,59	0,55	0,77	0,39	0,33	0,67	0,39	-0,14
distraction_glance_duration	X	X	X	X	X	X	0,24	0,57	0,49	0,86	0,41	0,30	0,74	0,60	0,29
behavior_more_speed	X	X	X	X	X	X	X	0,29	0,38	0,31	0,01	0,17	0,53	0,31	0,23
behavior_less_speed	X	X	X	X	X	X	X	X	0,36	0,52	0,41	-0,07	0,34	0,29	-0,13
behavior_speed_control	X	X	X	X	X	X	X	X	X	0,85	-0,09	0,37	0,73	0,32	0,15
behavior_lane_position	X	X	X	X	X	X	X	X	X	X	0,26	0,39	0,83	0,55	0,21
behavior_reaction_time	X	X	X	X	X	X	X	X	X	X	X	0,16	0,09	0,62	0,35
behavior_wrong_turns	X	X	X	X	X	X	X	X	X	X	X	X	0,67	0,68	0,68
behavior_operating_errors	X	X	X	X	X	X	X	X	X	X	X	X	X	0,58	0,39
bad_instructions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0,69
interruptions	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table QT: Correlation matrix Likert-scale questions Streaming (Z).

APPENDIX HB: LIKERT SCALE QUESTION CORRELATION DIAGRAMS

For all correlation coefficients below, lower bound of the confidence interval is at least 0.30.

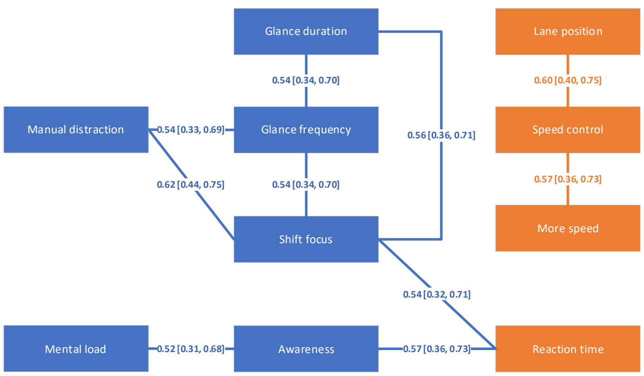

Figure QY: Cross-variable correlation coefficients and 95% level confidence intervals for Full set (F).

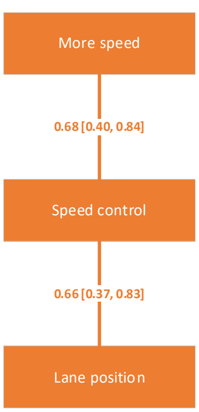

Figure QU: Cross-variable correlation coefficients and 95% level confidence intervals for group Smartphone (S).

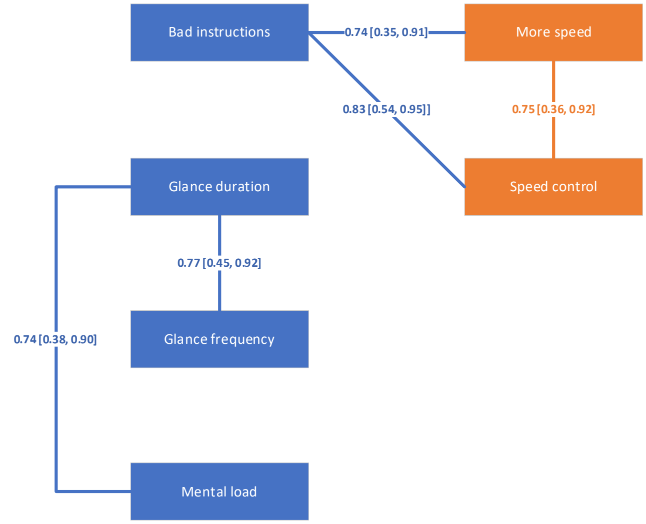

Figure QI: Cross-variable correlation coefficients and 95% level confidence intervals for group Integrated (I).

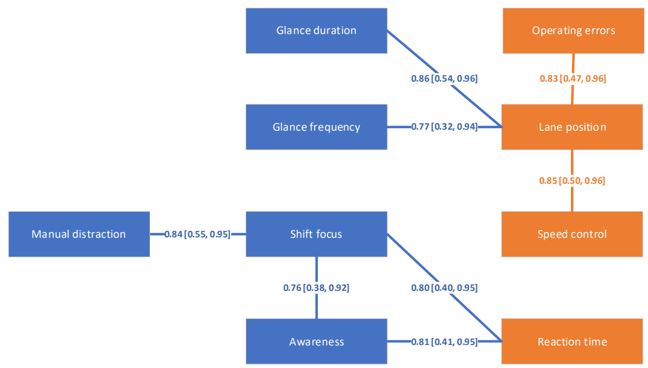

Figure QO: Cross-variable correlation coefficients and 95% level confidence intervals for group Streaming (Z).