UNIVERSITY OF GOTHENBURG

Impact of IVIS on driving performance and safety on the road

Zlati Yordanov
Asif Hussain

Bachelor of Software Engineering and Management Thesis

Report No. 2010:029

University of Gothenburg
Department of Applied Information Technology
Gothenburg, Sweden, May 2010

Impact of IVIS on driving performance and safety on the road

Zlati Yordanov
I.T. University of Gothenburg
zlati_yordanov@hotmail.com

Asif Hussain
I.T. University of Gothenburg
mirzaasif21@hotmail.com

Abstract

Automatic Speech Recognition is a state-of-the-art technology which is changing the way people interact with computers nowadays. Automotive industry profits from this technology while using it for controlling the in-vehicles devices, in a so called In-Vehicle Information Systems (IVIS). It is a problem when IVIS becomes a hindrance instead of assistance, so a system with the least distraction caused while using it, is to be preferred. The three most common interactive modalities of the user interface are: Manual, Auditory-Visual and Auditory only. With the help of a simulator and a prototype user interface we put the three interactive modalities to the test, measuring the distraction from the main task of the driving and performance of the human-machine interaction. A lab experiment conducted with the help of 15 drivers helped us assess the impact of those interactive modalities on driving performance. We found out that IVIS with Auditory only mode causes less distraction than Auditory-Visual mode. The worst mode was Manual which caused more stress, frustration and worse performance in all variables we used in the study.

Keywords: ASR (Automatic Speech Recognition), IVIS (In-Vehicle Information System), Interactive modalities, Auditory-Visual mode, Auditory only mode, Manual mode.

I. Introduction

In-Vehicle Information Systems have become a vital part of luxury cars nowadays. IVIS comprises of variety of diverse gadgets, for instance navigator, entertainment, communication devices, warning and emergency help systems. It allows a lot of freedom to the driver when interacting with in-vehicle devices. Automotive industry focuses on optimizing these IVIS systems. Nowadays voice controlled IVIS systems are out on the market, for example, Ford SYNC and JaguarVoice. They are easy to use, because they allow hands on the steering wheel and eyes on the road.

In-Vehicle Information Systems have different interactive modalities, Manual, Auditory only and Auditory-Visual. Safe driving is a first priority task for a driver while using IVIS. Since the complexity of IVIS, drivers can be distracted in various ways while trying to communicate with it. For example, some modalities demand physical effort while others cause higher cognitive load, leading to not being able to keep focus on the road (Graham & Carter, 2000). IVIS are designed to assist drivers and the more complex they get to communicate with the less helpful they are, therefore causing more irritation. In the upcoming paragraphs we are going to review what the literature says about these interactive modalities.

The studies have proven the effectiveness of speech based mode over manual control. Using manually-operated in-vehicle systems while on the move has a detrimental effect on driving performance. In the experiments, speech control resulted in improved lane keeping, fewer collisions and reduced target reaction times, compared to manual control (Graham & Carter, 2000). A survey study done by Nuance exposed that in almost all driving scenarios, users prefer the voice control over manual control (Nuance, 2009). Another research reveals that address entry with voice recognition is safer and shorter than with a keyboard and it is a recommended approach (Tsimhoni et.al, 2004). It is exposed that Manual mode is not safe and it leads to worse driving performance, which may have dangerous consequences.

Auditory only mode of IVIS is better to lower the distraction on concurrent tasks while driving. Speech based input has potential to reduce the risk of In-Vehicle Information System operation to a manageable level (McCallum et.al, 2004). The Auditory-only feedback is probably the best modality for a car environment, to avoid any visual distraction effects (Graham & Carter, 2000). It is suggested that speech recognition should only rely on speech dialog, instead of visual display (Vollrath et.al, 2008). These empirical findings uncovered that Auditory only mode causes less distraction on driving, hence safe and better driving performance.

In Auditory-Visual mode the interaction between driver and IVIS system is vocal and visual. It is indicated that visual presentation of multimodal IVIS can act as a redundancy or complementary modality to auditory presentation, which aids the resource relieving demand (Fu et.al, 2004). Providing visual as well as auditory feedback did adversely affect concurrent driving performance (Graham & Carter, 2000). A report on in-car distraction study disclosed that elimination of visual display from speech recognition causes less distraction (Vollrath et.al, 2008). In the light of these discoveries, we understand that Auditory-Visual mode is the source of additional distraction, which leads to lowered driving performance.

The aim of this research is to examine the distraction effects of the three IVIS interactive modalities on driving performance. For this purpose we are going to use a prototype interface and a simulator to test these modalities.

This paper begins with a brief introduction of the problem. It describes the background and related research on the topic. Furthermore, it explains the purpose of this research. The second chapter gives details of the methodology. It outlines how data was collected, its implications and which instruments were used for gathering it. Third section deals with results coming out of the data and discusses them. The documents end with conclusion.

II. Method

The investigation of the problem demands a systematic and scientific approach. It should be consistent with the area of research. The method should be well practiced and have empirical authentication with the problem domain. We came up with the idea of lab experiment by using a driving simulator, because this approach is suitable for this study. A literature review reports that 9 out of 15 experiments were performed on simulators (Baron & Green, 2006). Moreover, (Fu et.al, 2004) and (Graham & Carter, 2000) also used similar simulators. So the use of a simulator is evident from various studies. Further, this segment explains the sampling method, participants, apparatus and procedure.

Sampling our research data was inspired by a method described in the literature as the Lane Change Task (Vollrath et.al, 2008). Being developed by Deimler-Chrysler it is a very common standardized approach for assessing workload, stress or distraction while driving (Mattes, 2003).

II-A. Participants

Subjects between the age of 19 and 31 performed different tasks while driving .The subjects were selected by the criteria of having a driver's license and represented both students from IT university of Gothenburg and people who have non IT background. In total 15 people were tested of which 11 men and 4 women.

II-B. Apparatus

Figure 1: Prototype Interface

A prototype interface of an In-Vehicle Information System (IVIS) running on a PC was used as a voice controlled IVIS with the three interactive modalities (Figure 1). For the more complicated task of entering an address, a Tom Tom 910 GPS navigator with touch display was used.

Figure 2: Driving Simulator

The driving simulator is a Logitech G25 Racing Wheel game installed in a Saab 9.5 Kombi cockpit. It consists of three pedals, gear change joystick, a joystick steering wheel and a Hitachi CPX1 XGA projector with 1280x1024 resolution. Figure 2 shows the driving simulator. It collects data for the driving performance of the subjects like for example the mistakes he/she makes, time of the drive, speed, distance etc. The simulator track consists of a road going through both the countryside and small towns. The road is of standard width of 6.6 meters and it takes approximately 15 minutes to get to the end of it .It has uphills, downhills, turnings, trees and warning signs. To make it even harder for the drivers there are also cars going in both directions (car parameters width: 1.2m, length: 4.57m). Some of them drive over the speed limit, so that the subjects have to keep track of what is going on behind the car as well in order not to crash. The driving simulator was also keeping track of the steering input (how much the driver uses the steering wheel), the time and value of over speeding, time and value of crossing the edge excursion, time and value of crossing the centerline.

Moreover, a stopwatch was used to get the total time of "eyes off the road" while performing the different tasks with the prototype interface.

To furthermore enrich our data collection we used the NASA TLX: Task Load Index which is a task load assessment tool, allowing users to assess workload on human-machine interaction (Hart & Staveland, 1988).

II-C. Variables

Eyes off the road:

The first measurement we used was "eyes off the road" so that we can assess the length of the time interval of not concentrating on the driving, for the three modalities of the system. We assume that "eyes off the road" variable will be zero for the Auditory only mode, because other distractions except for peering at the system display are not taken into account. So the variables we will count are Auditory-Visual, Auditory only=0, and Manual.

Task Distraction:

Task distraction will be calculated using the NASA TLX: Task load index scales. We calculated the average values for each scale in this index.

Driving Mistakes:

Speed Exceedance: The maximum speed is 90 km/h which is also the limit on a high speed road in Sweden where the study is conducted. The simulator calculates the %Time and %Distance of which a driver breaking the law is over the maximum speed limit.

Centerline crossings: The simulator is counting the Time in seconds and Value in meters for each crossing of the centerline of the road.

Road edge excursions: The road edge is 2 meters wider than the lane and if a driver crosses it, it is recorded in the data as a mistake.

Out of lane driving: This is the variable giving %Time and %Distance of which the subject has driven the car out of the boundaries of the road lane which in this case is 3.3 meters. Lane parameters are shown in Figure 3, borrowed from (Fu et.al, 2004).

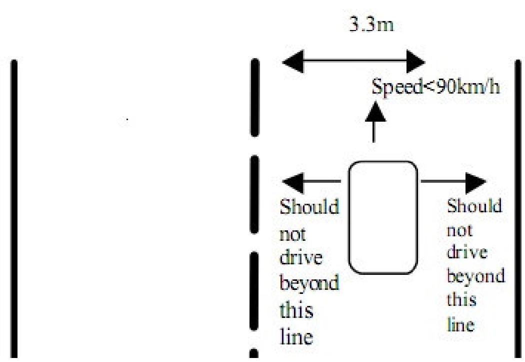

Figure 3: Lane parameters

Total Dangerous Position Index & Total Over Speed Index:

To get the Total Dangerous Position Index and Total Over Speed Index we firstly need to calculate the Arithmetic Mean (AM or Average) for each value of the drives. We use the formula;

$$AM(a_1, \ldots, a_n) = \frac{a_1 + \cdots + a_n}{n}.$$

If n numbers are given, the average of a_1 to a_n is the sum of $a_{1,\ldots,}a_n$ divided by n, where n is the count of the numbers (Miller, 2003).

When we have calculated the average values AM(values) we can calculate the variable D(total-dangerous-position) with our formula:

D(total-dangerous-position-index)**=**
[AM(Centerline crossing time) **x** AM(Centerline crossing MaxValue)**]** **+**
[AM(Road edge excursion time) **x** AM(Road edge excursion MaxValue)**]**;

Where AM(Centerline crossing time) & AM(Road edge excursion time) are in seconds and AM(Centerline crossing Max Value) & AM(Road edge excursion MaxValue) in meters.

Moreover, the centerline is the limit line which should not be crossed to the left of the road and the road edge is the border to the right (see Figure3). If a subject crosses one of the lines it is counted as a mistake, same criteria counts for driving over the speed limit.

D(total-dangerous-position-index) is an approximate value of the mistake severity MaxValue over time and is useful because when a driver goes out of the lane it is the time and MaxValue that shape the derivative of this mistake. As we have approximate maximum value of this derivative we can multiply it by the approximate time for which it occurred and get as a result the final value describing the shape of the derivative (rectangular in this case because we have the same approximate MaxValue for each second of the mistake and we can calculate its area). If the data shows that a mistake occurs more often than another mistake, but the second mistake has much bigger value we can compare them and be able to assess which modality of the prototype system is safer. So, to calculate the D(total-dangerous-position-index)we need to add the dangerous position of the car from both the centerline and road edge.

Another variable we are interested at is D(total-over speed-index).To calculate this we are using our formula:

D(total-overspeed-index)**=**
AM(Maximum Speed Exceedance time)**x** AM(Maximum Speed Exceedance Value).

Where AM(Maximum Speed Exceedance time) is in seconds and AM(Maximum Speed Exceedance Value) in km/h.

We calculate D(total-over speed-index) the same way as D(total-dangerous-position-index), only that the derivative is formed by the average time of the Maximum Speed exceedance multiplied by the average Maximum Speed exceedance value.

II-D. Procedure

Subjects had to make a test drive with the simulator so that they can get to know how to use it and avoid any extra distraction. They had a chance to see how the prototype and Tom Tom interfaces work, by having a 10 minutes instruction session beforehand. They also practiced the scenarios with the interfaces prior to test. Prototype interface scenarios included: choosing a contact from a list of contacts and dialing it, choosing a song (from CD/MP3 or radio submenus) and playing it. Further, Tom Tom was used for the scenario of entering an address to the system.

In this study we are using the "Wizard of Oz" (WOZ) method where a person pretends to be an ASR. The use of WOZ method is a common approach in many studies (Baron & Green, 2006). We prepared three interactive modalities of the system, Auditory-Visual, Auditory only and Manual;

Auditory-Visual: the display was the PC that we used to run the program, when the simulation was executed, the subjects had to give speech commands to the system and look at the display to get visual feedback and have a better overview of the menu navigation.

Auditory only: we used exactly the same scenario just that we took away the display and the subjects had to give only speech commands to the system while driving.

Manual: subjects had to navigate through the prototype interface menu to complete the given tasks while using the mouse of the computer for choosing a contact and playing music scenarios and Tom Tom navigator for entering address.

Each subject had to drive three times following the three scenarios but using a different modality. While they were driving we also placed a person with a stopwatch to the side of the driver so that he can get the "eyes off the road time" excluding the time when the driver looks at the scenario paper.

After each scenario a previously printed copy of the NASA TLX: Task Load Index was handed to the subjects so that they can answer the different questions connected to the mental load, physical load etc. during the drive and grade it on a scale from 1 to 20 . Each of them had the chance to take a second look at the answers from the previous drive so the grades for the specific drive could be comparable. This was necessary because subjects could not remember the grades they gave for their previous drives.

III. Results and Discussion

Evaluation of eyes off the road

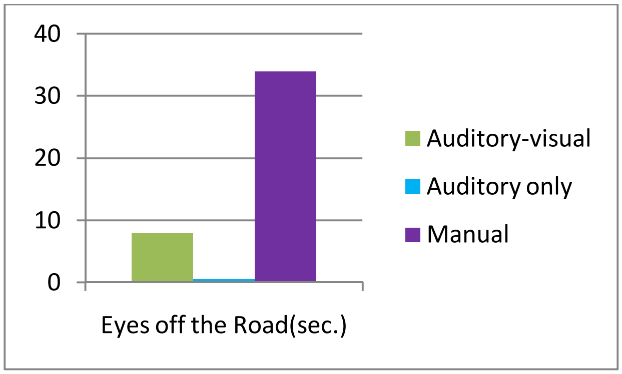

Figure 4: Eyes off the road difference.

Eyes off the road measurement results state, Auditory only = 0, since there is no display to distract. The average values for the subjects driving with Auditory-Visual mode is =7.9 seconds, and with Manual mode is =33.85 seconds. So we can clearly see that the Manual mode distracts the driver's eyes the most. The Auditory only mode is preferred over Auditory-Visual and Manual mode. Figure 4 shows the difference between the three modes.

Evaluation of Task Distraction

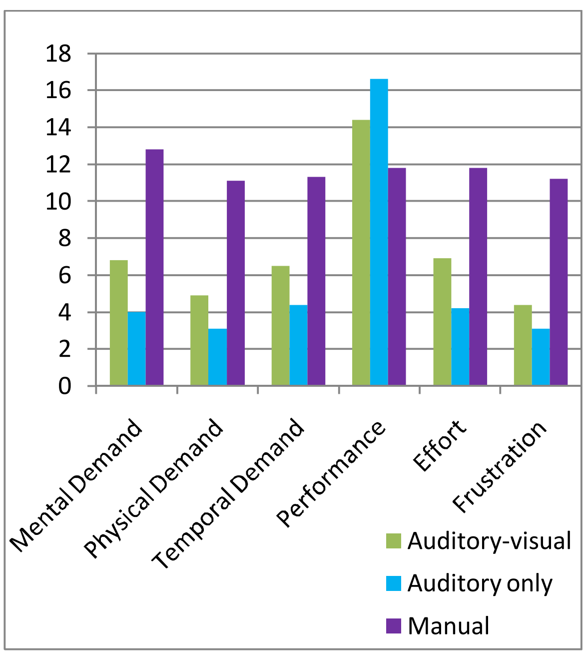

Figure 5: NASA TLX

Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration
6.8	4.9	6.5	14.4	6.9	4.4
4	3.1	4.4	16.6	4.2	3.1
12.8	11.1	11.3	11.8	11.8	11.2

Auditory- **Auditory** **Manual**

Table 1: NASA TLX Results

NASA TLX results are: best scores had Auditory only, followed by Auditory-Visual mode and worst results showed the Manual mode according to figure 5 and table 1. So we state that Auditory only mode has indisputable advantage in task load over Auditory-Visual mode and Manual mode. A more important fact we observe here is that Auditory-Visual mode causes more mental demand than Auditory only and literature stated the opposite according to (Graham & Carter, 2000).

Evaluation of Speed exceedance & Centerline crossing

Auditory-Visual mode scored 2.13 times speed exceedance in average, 1.8 times centerline crossings and zero road edge excursions and off-road accidents. Auditory only mode had 1.33 times speed exceedance, 0.66 times centerline crossings and zero road edge excursions and off-road accidents. Manual mode had 2.43 times speed exceedance, 3.5 times centerline crossings 0.2 times road edge excursions and 0.066 times off-road accidents per person in average. In Figure 6 we see that subjects who used the Manual mode committed the most mistakes in all the four measurement variables. Auditory-Visual mode caused almost twice as many mistakes in comparison to Auditory only mode in centerline crossing and speed exceedance and zero in the more serious mistake variables Road edge excursion and Off-road accidents so, we can state that Auditory only mode is safest in keeping the car in the borders of the lane and speed limitation.

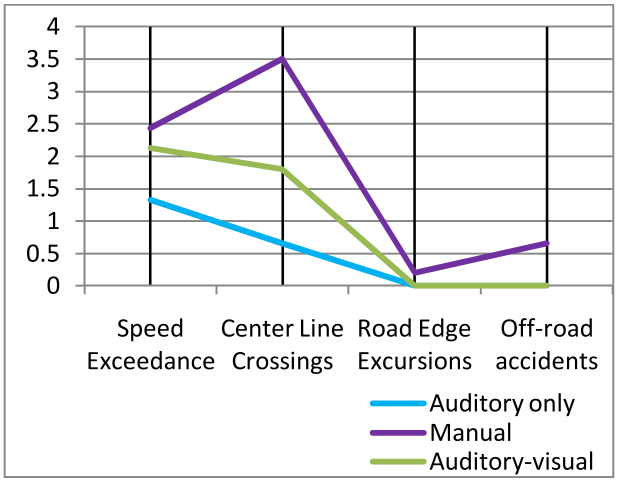

Figure 6: Speed exceedance and center line crossings.

Evaluation of Over Speed and out of Lane driving

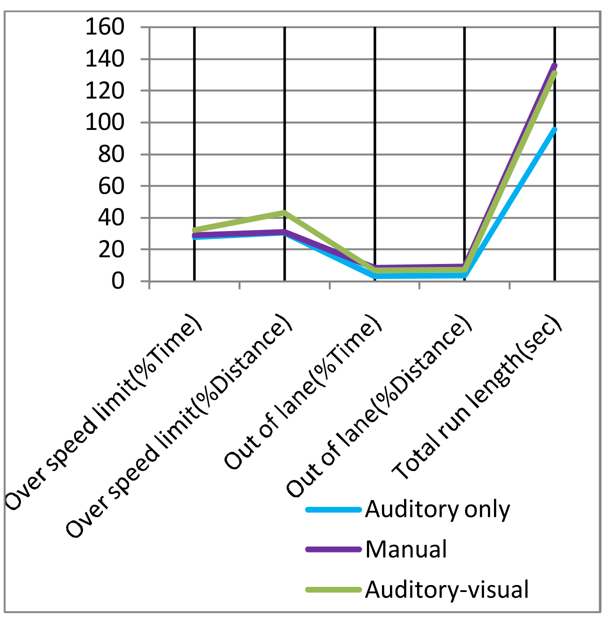

Figure 7: Over speed and Out of lane.

Figure 7 shows Over speed, Out of lane and Total run length. Measurement of the %Time and %Distance of Over Speed limit using Auditory-Visual mode scored %Time=32.2, %Distance=43, Auditory only mode scored %Time=27.7, %Distance=30.7 and Manual mode had %Time=29, %Distance =31%. We do not see big differences with the values except for subjects using Auditory-Visual mode were over the speed limit more than while using Manual mode and least with Auditory only mode.

Evaluation of %Time and %Distance of Out of lane driving shows that Auditory-Visual mode %Time=6.8, %Distance =7.2, Auditory only mode scores %Time=3.4, %Distance=3.53 and Manual mode %Time=8.6, %Distance=9.2%. Once again, subjects who used Auditory only mode were the most precise in driving in-lane most of the time and distance, followed by Auditory-Visual mode and Manual mode.

In the time of the total run length drivers using Auditory only mode finished the track fastest with 95.6 seconds in average, followed by Auditory-Visual mode scoring 131.06 and Manual mode with 135.9.

Auditory-Visual mode scored D(total-dangerous-position-index)=5.7 and D(total-overspeed-index)=1682.2, while Auditory only mode had D(total-dangerous-position-index)=0.5 and D(total-overspeed-index)=1221 and last but not least Manual mode scored D(total-dangerous-position-index)=9.3 and D(total-overspeed-index)=1668.6. One more time according to the results Auditory only mode scored the lowest "danger" index and we consider it as the safest, because subjects had the least mistake score.

IV. Conclusion

IVIS systems are used more and more in the vehicle industry nowadays. They are complex systems which are intended to help the driver so that he/she can easily command the devices in the car. Most interesting to this study are voice controlled IVIS systems, which bring less distraction to the drive because both hands are kept on the steering wheel and eyes on the road while giving commands with the help of speech. This is necessary as it is illegal to answer the phone in the car while driving in many countries but devices like hands free make it possible for us to make phone calls on the move. Controlling the entertainment and typing in GPS addresses tasks require a lot of physical and mental effort if done manually and previous studies have revealed the impeccable advantage of voice control over Manual control to communicate with the IVIS machines. The goal is as little distraction as possible and companies nowadays manufacture IVIS with Auditory-Visual mode mainly because of the visual feedback given to the driver, thus making it easier for him/her to understand the response of a command and not lose track of where in the menu he/she was navigating if something happened on the road. Other studies stated that it is unnecessary to have a display because even if it gives complementary feedback to speech it affects adversely the driving performance and leads to more resource demand.

Moreover, our research discovered that subjects state that Auditory-Visual mode cause more physical, mental, temporal demand, frustration, effort and worse performance on the road compared to Auditory only mode. We also conclude that Auditory-Visual mode causes more driving mistakes than Auditory only mode, but on the other hand the duration of being out of the safe driving zone was almost equal. For both driving out of the lane and over the maximum speed limit Auditory-Visual mode caused more mistakes and they were also more severe ones than Auditory only mode. The total driving time was also longer for subjects who used Auditory-Visual mode, so they needed more time to complete the tasks which were included. This study shows a complete picture of the different IVIS system modalities. Previous literature on the topic primarily concentrates on comparing IVIS with manual and voice input, and very few compare IVIS with Auditory-Visual mode and Auditory only mode. Instead we combined all three modalities so that the reader can have a better overview of the problem overall.

This study had also some limitations that we are going to outline so that researchers can continue working on it in the future. Firstly the prototype interface we worked on had to be connected to a touch display and placed close to the driver. The prototype system was supposed to run on it and thus not distract drivers as much as the bigger PC screen.

Secondly, the simulator track could be set up resembling a city environment so that subjects have to take turns very often and keep track of traffic lights and pedestrians. It would be interesting how this study could be applied to extremely stressful driving environments and see how it will affect performance and safety on the road.

References

(Baron & Green, 2006), Adriana Baron and Paul Green, *Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review, 2006.*

(Fu et.al, 2004), Zhang Fu, Lay Ling Pow, Fang Chen, *Evaluation of the difference between the driving behavior of a Speech based and a speech-visual based task of an in-car computer, 2004.*

(Graham & Carter, 2000), Robert Graham & Chris Carter, *Comparison of Speech input and Manual Control of In-Car Devices while on the Move, 2000.*

(Hart & Staveland, 1988), Hart, S.G., & Staveland, L.E. (1988). http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLXScale.pdf

(Mattes, 2003), Mattes S, The lane-change-task as a tool for driver distraction evaluation. In H. Strasser& K. Kluth & H. Rausch & H. Bubb.(Eds.), *Quality of work and products in enterprises of the future* (pp. 57-60). Stuttgart: Ergonomia.

(McCallum et.al, 2004), Marvin C. McCallum, John L. Campbell, Joel B. Richman, & James L. Brown, *Speech Recognition and In-vehicle Telematics Devices: Potential Reductions in Driver Distraction, 2004.*

(Miller, 2003) ,Steven J. Miller*, The Arithmetic and Geometric Mean Inequality, 2003.*

(Nuance, 2009), Nuance*, Automotive Voice UI Usability Study User Survey, Attitude, Experience, Motivation and Key Issues, Germany, 2009.*

(Tsimhoni et.al, 2004), Omer Tsimhoni, Daniel Smith, Paul Green*, Address Entry While Driving: Speech Recognition versus a Touch Screen Keyboard, 2004.*

(Vollrath et.al, 2008), Prof. Dr. Mark Vollrath, Dipl. Psych. Jannette Maciej, Dipl. Pysch. Ute Niederee*, In-Car Distraction Study Final Report, 2008*