GOT INFO? EXAMINING THE CONSEQUENCES
OF INACCURATE INFORMATION SYSTEMS

Ing-Marie Jonsson[1], Helen Harris[1],
Clifford Nass[2], Leila Takayama[2]
[1]Toyota InfoTechnology Center
Palo Alto, California, USA
E-mail: {ingmarie, helenh}@us.toyota-itc.com

[2]Stanford University, Dept. of Communication
Stanford, California, USA
E-mail: {nass, takayama}@stanford.edu

Summary: It is a desirable goal to balance information given to the user with the potential adverse effects on cognitive processing and perception of information systems. In this experiment, we investigated the minimum level of information accuracy necessary in an in-car information system to elicit positive behavioral and attitudinal responses from the driver. There were 60 participants, and each drove in a simulator for 25 minutes; driving performance data was automatically collected, and drivers later completed questionnaires for attitudinal data. Participants were divided into three groups of drivers: a group driving with a 100% accurate system, another driving with a 70% accurate system, and one group driving without an in-car system. There was a definite positive effect on driving performance with the in-car system, and results show that decreasing the accuracy of the system decreases both the driving performance and the trust of the in-car system. Data also indicates that female drivers have a higher tolerance of inaccuracies in an in-car system; design implications are discussed.

INTRODUCTION

Automobile manufacturers are now making in-car systems and features available in all vehicles, not just the luxury models. Interacting with these systems affects the driver's on-road attention and cognitive load. Screen-based interaction requires the driver's focus to move from the road to the screen (Lunenfeld, 1989; Srinivasan, 1997). Mental tasks requiring operations with images produce more pronounced and different alterations in the visual search behavior than those corresponding to verbal tasks (Recarte et al., 2000). Attention theory suggests that speech-based interaction would be less detrimental to driving than interactions with a visual display (Wickens, 1984). There are however, results that demonstrate that even simple conversation can disrupt attentive scanning and representation of a traffic scene (McCarley et al., 2001).

Numerous studies on the impact of using mobile phones while driving have been done over the past years, but the distraction and safety consequences of other types of speech-based interactions while driving are still not fully understood (Strayer et al., 2003). Recent studies on speech-based interactions with in-car computers confirm some previous results on speech interaction over mobile phones, with an increase in reaction time and the perception of distraction (Lee et al., 2001).

For the experiment described in this paper, the in-car system provides verbal information that pertains to road and traffic conditions. In this case, the credibility of the content becomes a key factor in determining the effect on the driver (Tseng et al., 1999). In human-human interaction, reliability and trust are central to relationships and form a basis for communication and actions. These properties are also recognized as important for human-computer communication (Brave & Nass, 2002), and we would predict that varying the accuracy of an in-car information and warning system would likely affect the driving task performance and the driver's attitude about the trustworthiness and expertise of the system (Kantowitz et al., 1996).

IN-CAR INFORMATION AND HAZARD WARNING SYSTEM

This experiment was set up to investigate the impact of information and warning prompts on driving performance and to determine the threshold of error at which a driver exhibits improved driving performance and both likes and trusts the system.

The information prompts, ranging from informing the driver of the current speed limit, to warning the driver about potential road hazards, were recorded and later played back by the in-car system. The prompts were all worded to inform and warn about features and hazards related to the driving environment (Jonsson et al., 2004). To identify the threshold of errors that would be acceptable, there were two (three with the "Silent" system without information prompts) variations of the in-car system. The first variation was "100 Percent Accurate" ("100") in which each of the 33 alert-prompts was appropriate to immediately upcoming events in the driving simulator. The second variation of the in-car system introduced deliberately flawed information. In this case, the same recorded prompts as the "100" system were used, however a small set of them were moved around to create errors of omission and commission. The result was a "70 Percent Accurate" ("70") system where 30% of the prompts were either inaccurate or inappropriate. The third variation of the in-car system was "Silent," and did not give any prompts to the driver while driving.

METHOD

The experiment was a 3 ("100," "70," "Silent" Condition) x 2 (participant gender), balanced and mixed between-participant design. All participants were informed that the study would last an hour and that the experiment was set up as two driving sessions interspersed with three online questionnaires. The entire study was conducted in one room where both the driving simulator and the computer for the web-based questionnaires were placed.

Participants

A total of 60 participants, 30 male and 30 female, were recruited to participate in the study. Participants were between the ages of 18 and 25, had a valid driver's license, and a majority had computer game experience. All gave informed consent and were debriefed after the experiment.

Experimental Apparatus

STISIM *Drive* from Systems Technology, Inc., was used as the driving simulator. The simulator was running on a PC and back-lit-projected onto a big screen. Participants sat in a real car seat

while "driving" and controlled the driving simulator with a forced feedback steering wheel and two pedals, an accelerator and a brake. The simulator set-up was used for both an introductory course and the main driving course, and the same simulator configuration and driving courses were used by all participants. The introductory driving session was designed to familiarize participants with the controls and the feedback of the driving simulator. The main driving course took participants an average of twenty-five minutes to complete. There were three web-based online questionnaires. The first questionnaire asked for general information such as gender and age in addition to driving and computer game experience. The second was a blame attribution questionnaire, and the third asked about the quality of the car, the information system, and the effects of driving with the in-car system. For the "Silent" condition, the questions referring to the information system were omitted.

Measures

Driving Performance. Data for driving performance was generated and collected automatically by the driving simulator. The simulator was instrumented to save data on the driver's steering, accelerating, braking, accidents rate, speeding, and obedience to traffic rules. The data was primarily analyzed by indices built from these variables. **Accidents** include collisions and off-road accidents; **bad-driving** combines the number of speeding instances, light tickets, center-line crossings, collisions, and off-road accidents.

Credibility and Trust of the In-Car Information System. Attitudinal data was collected through standard and custom-made questionnaires. Berlo and McCroskey (Rubin, 1994) indices measure how credible the in-car information system is perceived to be. Contrasting adjectives were paired on opposite sides of a 7-point scale such that, for example, *reliable* and *unreliable* would appear at different ends. The Berlo **qualification** index combines the attributes of being *experienced, qualified, trained, skilled, informed,* and *candid*. The index was very reliable (alpha = .91). The McCroskey index of **authoritativeness** included *reliable, informed, qualified, intelligent, valuable,* and *expert* (alpha = .89). **Trustworthiness of character** was based on whether the system was *honest, pleasant, unselfish, friendly,* and *nice* (alpha = .60).

RESULTS

The credibility of the information system and driving performance were analyzed using a two-way ANOVA with the variant of the in-car information system and gender as the between-participants factors.

Driving Performance

Accidents. One effect of decreasing the accuracy of the information system was a noticeable decrease in driving performance. Drivers who interacted with the "100" system had fewer **accidents** ($F(2.36) = 4.84$, $p < .034$) than the "70" group. Male drivers also had more accidents than female drivers, however not significantly more ($F(2, 36) = 3.69$, $p < .06$, M = 2.84, SD= 1.6, and M = 3.80, SD = 1.2), based on a two-tailed t-test, $t(38) = 2.2$, $p < .04$. See Figure 1.

Furthermore, drivers in the "100" system had significantly fewer accidents than those without an information system ($F(2, 36) = 4.66$, $p < .038$). This suggests that alerting drivers of hazards in

the road helped them to avoid these hazards and in fact made them more aware of the driving environment. An important observation is that male drivers in the age group 18-25, a high-risk group, had significantly fewer accidents, indicating the information system was effective in improving their driving performance. When comparing **accidents** between drivers in the "70" system and drivers without an information system, there is an interaction, though not significant, between gender and variant of information system. The "70" system female drivers have fewer accidents, while male drivers in the "70" system have approximately the same number of accidents as the "Silent" drivers ($F(2, 36) = 3.36$, p $< .075$, M=3.3, SD = 1.2 and M = 4.2, SD = 1.2), based on a two-tailed t-test (t(18), 1.9, p<.072). See Figure 1.
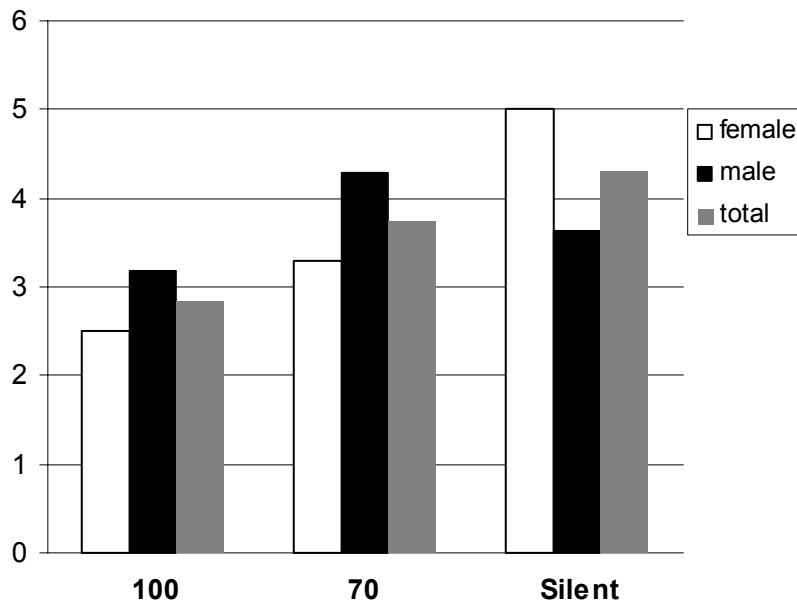

Figure 1. Average Accidents per Driving Session

Bad Driving. Evaluating **bad driving**, the "70" drivers performed much worse than the drivers with the "100" system, ($F(2, 36) = 5.92$, p $< .020$), and most importantly, drivers in the "100" system performed significantly better than drivers with no system, (F(2, 36) = 4.76, p <.036). As expected, there was a significant difference in **bad driving** between male drivers and female drivers, with male drivers performing worse ($F(2, 54) = 5.38$, p <.024, with male drivers M = 35.87, SD= 16.04 and female drivers M = 27.47, SD= 12.19 based on a two-tailed t-test t(58)=2.28, p<.02). These results indicate that female drivers more readily accept inaccurate information given by an in-car information system. See Figure 2.

Credibility and Trust of the In-Car Information System

As anticipated, the decrease in accuracy was perceived by the participants and resulted in a lower rating of the system's **authoritativeness** ($F(2, 36) = 15.77$, p $< .001$), **trustworthiness of character** ($F(2, 36) = 7.91$, p <.007), and **qualification** ($F(2, 36) = 16.05$, p <.001). When the in-car information system makes obvious errors or omissions, the user will take notice and consider the "70" system to be flawed. The consequences of this is that the driver does not find the system trustworthy, and that the driver finds the system less qualified to make informed
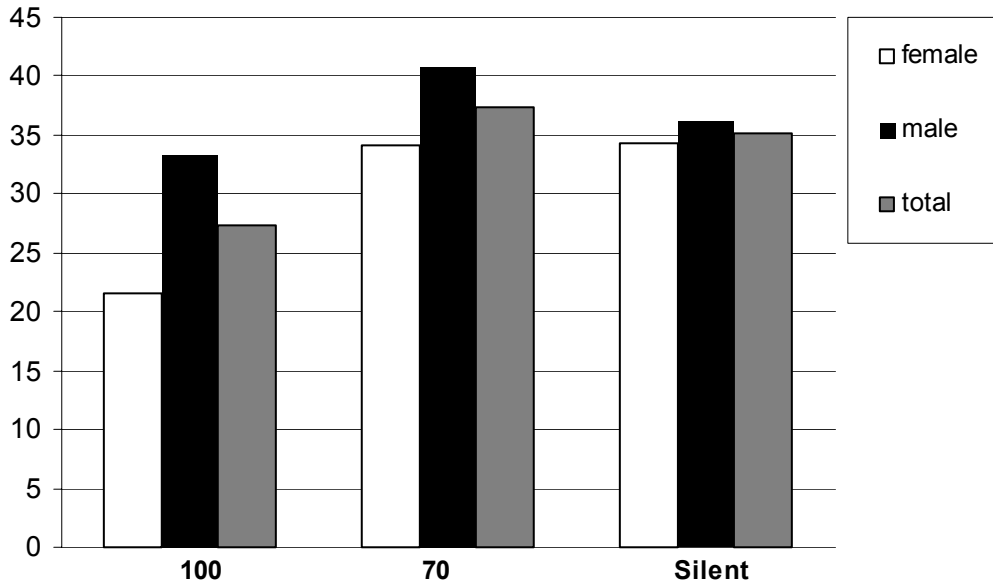
statements about the roadway. See Figure 3.

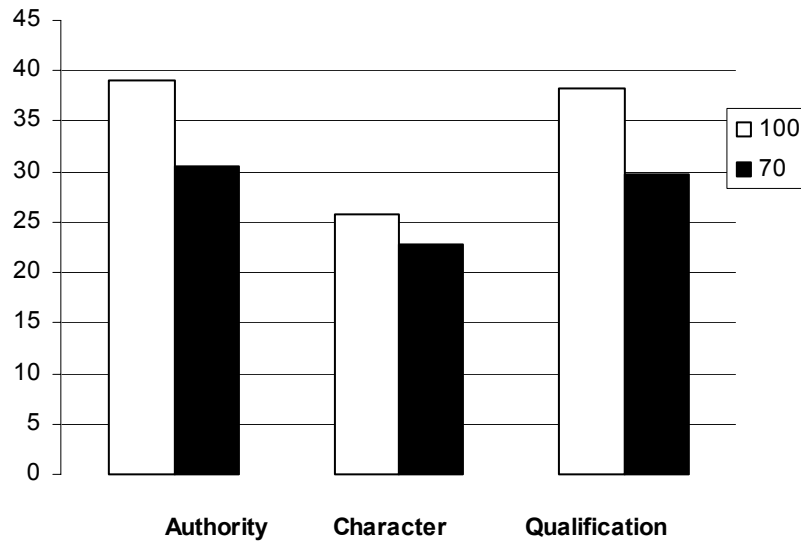**Figure 2. Bad Driving Index (Accidents, Speeding and Swerving)**

.

Figure 3. Credibility Ratings of the In-car Information System

CONCLUSION

The results from the experiment show that manipulating the accuracy level of an information system affects driving performance in the driving simulator. Statistics collected by insurance companies suggest that male drivers under the age of 25 drive more aggressively and have more

accidents than any other group of drivers. Our data shows that male drivers in the age group 18-25 had significantly more accidents and worse driving performance that female drivers, indicating that the driving simulator used for our experimentation did an adequate job of simulating the real driving experience; we are reasonably confident that other findings from this experiment are equally applicable to real-world driving.

The reduction in accuracy of information in the in-car hazard and warning system from 100% to 70% was enough to increase accidents and to degrade driving performance. However, it is extremely promising that a system with only 70% accuracy can help reduce the number of accidents compared to a car without an information system. We can conclude that the gender-neutral threshold for information in an in-car information system lies somewhere between 100% and 70%. Male drivers are aided by a 100% accurate system but not by a 70% accurate system. Female drivers accept a 70% accurate system and exhibit improved driving performance.

DISCUSSION

The focus for this experiment was to investigate the impact of an in-car information and warning system on driving performance and attitude towards the system. With the rapid introduction of information systems in cars, it is important for the car industry to further investigate and fine tune the acceptable level of information accuracy. Previous work has found 70% accuracy "acceptable and useful," but on a limited simulation of routing for the Advanced Traveler Information System (ATIS) (Kantowitz et al., 1996). As of now, we can say that people are aided by alerts, although it is not an across-the-board improvement. We have seen a clear benefit of in-car information and warning systems on driving performance, but it is also clear that a realistic system needs to be integrated with other types of information such as navigation, points of interest, and weather. The impact of driving performance by adding these information types should also be investigated.

With the advance of technology, detection of the driving environment improves, and more and more "safety features" are being added to cars. As a result, it will become more important to know how and when to send information to the driver. The bottom line is that the industry can build the safest and most technologically advanced systems in the world, but unless drivers actually find the systems to be trustworthy and pleasant, they will be turned off. This experiment was set up as one in a line of studies that aim to investigate driver acceptance and willingness to interact with and use in-car system targeted for both safe and pleasant driving.

ACKNOWLEDGMENTS

REFERENCES

Brave, S. & Nass, C. (2002). Emotion in human-computer interaction. In: J. Jacko & A. Sears (Eds.) *Handbook of human-computer interaction*. New York: Lawrence Erlbaum Associates: 251-271.

Jonsson et al. (2004). Don't Blame me I am only the Driver: Impact of Blame Attribution on Attitudes and Attention to Driving Task. *Volume two of Proceedings of CHI:* 1219-1222.

Kantowitz, B. H., Hanowski, R. J., & Kantowitz, S. C. (1996). Development of human factors guidelines for advanced traveler information system and commercial vehicle operations: The effects of inaccurate traffic information on driver acceptance of in-vehicle information systems. Washington, DC: Federal Highway Administration (FHWA-RD-96-145).

Lee, J., Caven, D., Haake, S., and Brown, T. (2001). Speech-based Interactions with In-Vehicle Computers: The Effect of Speech-based Email and Drivers' Attention to the Roadway. *Human Factors, 43*: 631-640.

Lunenfeld, H. (1989). Human factor considerations of motorist navigation and information systems. *Vehicle Navigation and Information Systems Conference Proceedings*: 35-42.

McCarley, J., Vais, M., Pringle, H., Kramer, A., Irwin, D., Strayer, D. (2001). Conversation disrupts visual scanning of Traffic scenes. *Proceedings of the Ninth International Conference on Vision in Vehicle*.

Recarte, M., and Nunes, L. (2000). Effects of Verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimentl Psychology: Applied*, *6*: 31-43.

Rubin, R., et al. (1994). *Communication Research Measures: A Sourcebook.* New York: Guilford Press.

Srinivasan, R., and Jovanis, P. (1997). Effect of in-vehicle route guidance systems on driver workload and choice of vehicle speed: Findings from a driving simulator experiment. *Ergonomics and Safety of Intelligent Driver Interfaces*: 97-114.

Strayer, D., Drews, F., and Johnston, W. (2003) Cell Phone Induced Failures of Visual Attention During Simulated Driving. *Journal of Experimental Psychology: Applied*, *9*(1): 23-32.

Tseng, S., & Fogg, B.J. (1999). Credibility and Computing Technology. *Communications of the ACM (5)*: 39-44.

Wickens, C., (1984). *Processing Resources and Attention, Varieties of Attention.* New York: Academic Press.