Ubiquitous Co-Driver System and Its Effects on the Situation Awareness of the Driver

Hannu Karvonen, Tuomo Kujala, and Pertti Saariluoma

Abstract—The aim of this paper is to explore the effects of ubiquitous computing in cars on the situation awareness and expectations of the driver. In a driving simulation environment with participants using a co-driver system, we investigated how people took and recovered from misinformation provided by the system. The system presented safety-critical information about the upcoming curves on the road, but in the experiment part of the messages contained false information. The effects of this information on participants’ behavior were investigated. On the grounds of the experiment, we discuss two approaches for investigating drivers’ situational awareness, which are based on either mental workload or mental contents.

I. INTRODUCTION

UBIQUITOUS computing (ubicomp) envisions a new world in which computing is seamlessly integrated into our everyday life [1], [2]. In cars, ubicomp is gradually becoming a natural part of our driving tasks with advanced context-aware computing systems such as navigation, collision avoidance, and adaptive cruise control (ACC) systems. The ultimate purpose of these context-aware driver assistance systems is to increase the safety, efficiency, and enjoyment of driving [3]. However, this intelligent technology does not necessarily have only positive effects on driving. It has been noticed that if the psychological effects of these systems are not sufficiently considered in the design, they can also have a detrimental effect on driving safety [4].

According to Burnett and Porter [5], the increase of in-car ubicomp systems means that there will be several new functions and information systems available for the driver. These functions and systems also bring along various controls and displays which can distract the driver and potentially disturb the main task of driving. Therefore, there is an urgent need for usable, integrated designs for these systems. Future in-car ubicomp system design has to take this need seriously. Thus, it is necessary to call more attention to the possibilities opened by user psychological research, i.e., the analysis of psychological processes involved when humans interact with this kind of technology (see e.g., [6]).

The traditional human-computer interaction (HCI) does not apply as such in analysing in-car ubicomp systems. Costagliola et al. [7] have defined three main factors why traditional HCI design techniques are inappropriate for ubicomp systems in the automotive domain. These factors include the limited attention which is available for interaction with a system, the displays in cars which can show only a reduced amount of information, and the fact that ubicomp systems cannot rely on an input device such as a mouse, making the current implementation of the point-and-click paradigm inadequate. They also note that there is a need for establishing new techniques and approaches which carefully take into account the specific issues of in-vehicle information systems (IVIS). Concerning the same topic, Rakotomairainy and Tay [8] state that future in-car systems must be aware of 1) factors that characterize the current situation (outside and inside a car), 2) the driver’s current physical and physiological conditions, 3) the driver’s intention or need (e.g., direction), and 4) the most appropriate and non-intrusive ways to interact with the driver.

From our point of view, the key problem is how to apply psychological knowledge of users to analyze and understand the demands of tasks. This is typically a user psychological problem, because it should be clearly defined and explained in psychological terms what kind of difficulties and benefits users meet when emerging technological applications are adopted [6], [9], [10], [11]. With in-car ubicomp systems it is essential to pay attention to the way we could best explicate the problems of mental contents. Traditionally, workload or mental capacity has formed the most important basis for analyzing human system interaction [12], [13], [14]. However, it does not easily capture the most essential aspect of human system interaction in the case of navigation and other similar in-car ubicomp systems, because these systems are designed to provide information contents for the users. This is why it is essential to focus on different aspects of mental contents, i.e., information contents in mental representations [11].

Mental contents as a general basis for interaction research are a vital conceptualization, because the concept makes it possible to unify a number of earlier findings under one single approach. Expectations [15], [16], [17] and schema
dependency [16], [17], [18] (for historical curiosity see [19]) provide two examples of phenomena which can eventually be seen as arising from incorrect or inadequate mental contents.

Expectations and the mental contents creating them are important e.g. in analyzing possible cases of misinformation provided by a system. Errors in databases, input information, or limited contacts to satellites can make systems provide misinformation to drivers. There have been cases in which a driver’s blind trust in a navigation system has led him to cause an accident when the system has given incorrect information. Therefore it is important to know how people react in counter-expected situations with these systems.

Errors made as a consequence of incorrect expectations are not similar to errors caused by cognitive underload or overload. Saariluoma [20] demonstrated that for example chess players may make serious errors when they have incorrect presuppositions and expectations concerning the situation, although it cannot be assumed that their processing capacity would be overloaded. From the user psychological point of view, expectations are thus different from workload as a type of explanatory factor of behavior [11].

To study these issues, we built a context-aware co-driver system which was tested in a driving simulator. Our previous work [21] has indicated that a similar kind of a driver tutoring system could be helpful in assisting in driving in a more economical and safe style, especially on unfamiliar roads. However, with the co-driver system we wanted to investigate situations in which the system gives safety-critical guidance messages about the upcoming curves. The effects of false messages on driver behavior could be examined with this system: a part of its guidance messages gave wrong information on purpose. The co-driver system resembles navigation systems, which have been investigated quite thoroughly under IVIS research. Nevertheless, the effects of possible false guidance messages on driving safety, which was our interest in the experiment, have not been previously under thorough investigation.

II. METHOD

A. Experimental Design

Two trials were driven on two tracks in the experiment: one without any secondary systems, and another with the co-driver system. The co-driver system gave traffic sign and textual information about the upcoming curves on a Head-Up Display (HUD) (see Fig. 1). In addition, a female voice was used to give the messages verbally. The information about the curves included the direction and the sharpness of the curve (a traffic sign), the distance to the curve (text), and the recommended speed for the curve (traffic sign). This information was always given on a predefined location on the road, depending on the distance to the upcoming curve.

In the trial with the co-driver system, the drivers first received correct messages from the system. After driving for a while, the system occasionally started to give wrong guidance messages. The hypothesis was that when these false messages were given, the drivers would become confused and make several driving errors. The number of driving errors in the results would therefore be significantly higher with the false messages than with the correct ones.

![Fig. 1. The drive view with the co-driver system’s HUD message.](image)

B. Participants

The 24 participants were the same as in our previous experiment [21] and consequently had experience of the driving simulator. They were recruited through public email lists of the University of Jyväskylä and divided into two groups on the grounds of their driving experience: novice (driving experience 0-5000 km) and experienced drivers (driving experience 25 000-230 000 km). The participants included 12 women and 12 men with ages between 19 and 39 (M = 24.08, SD = 4.48). Women and men were equally divided into the driving skill level groups. As a reward for taking part in the experiment, the participants were given a movie ticket.

C. Environment and tools

The driving simulation environment was built in the User Psychology Laboratory of the University of Jyväskylä by using a tailored version of an open-source car simulation program called Racer (www.racer.nl). The use of a driving simulator instead of a real car was especially justified with the co-driver system, because in this way possibly hazardous situations in a real environment due to false guidance messages would not occur. Further justification for this type of experimental design can be found in the research by Lee et al. [22]. They successfully investigated the possibility of using simple PC-based equipment for identifying problematic or unsafe older drivers. Their research indicates that it is possible to measure the participants’ cognitive and perceptual abilities effectively with low-cost PC-based simulators. Driving simulators can sometimes be blamed for poor simulation of driving with a typical passenger car. However, what is the prototype of a car that should be used in research? There are thousands of different vehicle models
that differ from each other in numerous ways. The use of a simulator can also be justified by the possibility of using high-risk traffic conditions or environments and the good access to traffic, vehicle, and subject data [23]. Results acquired in simulators have been found fairly consistent with those obtained in real vehicles according to the research of Santos et al. [24] and Engström et al. [25].

In our simulator, the high resolution drive view was projected through a data projector. The HUD was simulated by projecting the HUD information with another data projector over the drive view. The HUD information was placed on the bottom of the screen on purpose, because in this way it would not block objects from the drive view. The experimenter controlled and scheduled the HUD information from another computer. The drive view also included an analogous speedometer. As the main interface for driving, a steering wheel, pedals and an automatic transmission were used.

The road used in the trial without the co-driver system simulated a Norwegian countryside landscape. The road used in the trial with the co-driver system was a similar road, but it had more sharp curves. The roads were driven for about 4 kilometers in each trial.

D. Procedure

At the beginning of the experiment, the participants were told about the basic circumstances related to the experiment, such as that the trials were video-taped and the participants will be presented anonymously in the results of the experiment. After this, the participants were instructed only to drive straight ahead and told that the goal is to keep the vehicle inside its own lane as accurately as possible. They were also instructed that there is an 80 km/h speed limit on the road.

The first trial was driven without any extra systems. In the second trial the participants drove on another road of a similar level of difficulty and this time assisted by the co-driver system. Before the second trial, the participants were told that the car is equipped with a Head-Up Display in which different messages will be presented during driving. They were told that the messages are given by a co-driver system which gives information about the upcoming curves: whether the upcoming curve is gentle or sharp, the curve’s direction, the number of meters to the curve and the recommended speed for the curve. The participants were instructed to act on the co-driver system’s messages in a way that seems best for them.

In the trial with the co-driver system, the system first gave six correct guidance messages about the upcoming curves so that the participants would get used to the system. In the following 12 curves, six of the guidance messages were false and six were correct. The curves with false guidance messages were picked randomly. These curves were also changed with every other participant so that the properties of certain curves would not have an effect on the results. The information of the upcoming curves in the false messages was incorrect for the direction, sharpness, and the recommended speed for the curve.

E. Data Analysis

Several methods were used to gather data from the experiment. Qualitative data was captured by conducting semi-structured interviews after the trials and by instructing the participants to think aloud freely during the trials. The data from the interviews as well as the real time comments of the participants while driving were analyzed. Quantitative data of the trials was analyzed by counting the number of driving errors occurred during the trials. A driving error was defined to occur when the bonnet of the vehicle covered a white line marking the lane on the road on either side of the driving lane.

Other quantitative data was gathered with modified NASA Task Load Index (NASA-TLX) [26] forms after both trials. The form was modified to a 1-10 scale and the meter for physical demand was removed because of its irrelevancy in the experiment. Two additional meters were used with the form given after the trial with the proposed system. Due to these modifications, any weighting processes were not used in the method. The modified forms therefore included the following scales: mental demand, temporal demand, effort, performance, frustration, usefulness of the guidance messages, and obtrusiveness of the guidance messages.

In the analysis phase, the differences between genders and between novice and experienced drivers were also compared. Suitable parameters, such as means and standard deviations, were counted from the quantitative data. These results were also run through statistical significance tests (nonparametric Wilcoxon Signed-Rank Test and Mann-Whitney U-Test).

III. Results

A. The Interviews

The interviews focused on the participants’ subjective experiences with the proposed system. The participants gave generally more negative than positive comments about the co-driver system. Apparently, this was due to the system’s false guidance messages. When asked “Did you trust the system?”, almost every participant said that at first they did, but after noticing the false messages their trust was lost.

The use of verbal instructions in addition to the HUD messages was generally appreciated. This was mainly because the participants could focus more on the driving since they did not have to look at the textual messages to understand the message. Most of the participants, however, mentioned that the HUD messages were also useful because if they forgot the content of the verbal message, they could check the instruction from the HUD before the curve.

The false messages were reported to affect the driving considerably. Especially with the first false messages, the participants said that their driving was distracted in the
curves. They also said that after noticing that the system is giving them false messages, they did not care about the messages anymore. After this, the participants said they lowered their speed.

B. The Thinking-Aloud Method

Several comments regarding the distance approximations could be captured while driving with the co-driver system: most of the participants felt that the distance to the curve given by the system was difficult to estimate. Other comments related to the verbal messages, which were commonly experienced as positive. The HUD symbols were also found very useful and illustrative.

The comments at the beginning of the trial were more positive than at the end of the trial. One notable phenomenon could be seen when the system gave the first false messages: many of the participants wondered if they had misunderstood or heard the message incorrectly. However, they stopped blaming themselves after a few false messages when they realized that the system was responsible. At this point, some of the participants still tried to find some sort of logic in the functioning of the system: they for example wondered if the system was giving instructions for the curve after the upcoming one. After they realized that the system cannot be trusted at all, the participants said they stopped paying attention to the instructions.

C. Driving Errors

On the basis of the analyzed video material, it can be noticed that a driving error typically occurred because of excessive situational speed in curves. Only the last 12 curves were included when counting the number of driving errors, because in the first six curves all the guidance messages given were correct. Results of one participant had to be excluded in the analysis due to a technical fault during the trial in the recording equipment.

The difference between the mean number of driving errors committed between the trial without the co-driver system and with the co-driver system’s correct guidance messages was statistically significant (Z = -3.57, p < .001) (see Fig. 2). However, it must be noted that the effects of driving experience with the simulator were not controlled by counterbalancing the order of the trials because this difference was not the main interest in the experiment.

The mean number of driving errors with the co-driver system’s correct messages was only 0.91 (SD = 1.08). With the system’s false messages the mean number was 2.04 (SD = 1.58), which is close to the result of the participants’ trial without any systems. The difference in the mean number of driving errors between the trial without the system and the trial with the false messages was not found to be significant. However, the difference between the mean number of driving errors with the correct and false messages was statistically significant (Z = -3.23, p < .001). This means that driving errors occurred significantly more with the false guidance messages than with the correct ones. Qualitatively, these errors were also more severe and dangerous than the other driving errors made in the trials. When the first false messages were given, many of the participants even drove out of the other side of the road in the curves. The false messages also caused several panic brakeings when approaching the curve.

There were no significant differences when comparing the driving errors of novice and experienced drivers. Neither was the difference between the mean number of errors made by male and female participants significant. Thus, the level of experience or the gender of the driver does not seem to have an effect on driving with the co-driver system or on the reactions to the false messages.

D. The NASA-TLX

The mean for the reported mental demand was 6.88 (SD = 1.65) with the co-driver system, which means a 1.42 increase to the mean number of mental demand without the system. The difference between the means is statistically significant (Z = -3.26, p < .001). Also the means for temporal demand, effort, and frustration were reported to be higher and the mean for performance to be lower with the co-driver system than without the system. However, these differences were not statistically significant.

The usefulness of the co-driver system’s messages was evaluated in average as 3.23 (SD = 2.16). The difference between the co-driver system’s usefulness and obtrusiveness of the messages is also statistically significant (Z = -3.30, p < .001). This means that the messages were perceived as significantly more obtrusive than useful. According to the interviews, this was because of the false messages given in the trial. No statistical differences were found when comparing the answers of the novice and experienced drivers. The only statistically significant (U = 37.50, p < .05) difference between male and female drivers was that the
female drivers experienced their effort to be greater than the male drivers.

IV. DISCUSSION AND CONCLUSIONS

The purpose of our experiment was to examine one of the key psychological issues when prototyping ubicomp systems for automobiles: what happens when the system does not work according to the driver’s expectations?

The co-driver system can be considered an ubicomp system because it had the following qualities [27]: if the system worked correctly, its purpose was to disturb the driver as little as possible [2], it was context-aware (for location and timing), it used natural interaction methods (HUD and auditory information), it can be used in everyday activity, and its real implementation would utilize wireless networking (e.g., GPS information about the upcoming curves).

The results support the hypothesis made about the co-driver system before the experiment: the participants became confused with the first false guidance messages and several driving errors occurred in these situations. The difference between the mean numbers of driving errors as a whole was statistically significant when comparing the curves with correct and false messages. The driving errors made with the false messages were also more serious than the ones made with the correct ones or in the trial without the system, and they even included road departures.

The deterioration of performance with the false messages near to the level of participants’ first trials with the simulator may be at least partly explained by the malleable attentional resources theory [4]. The theory suggests that attentional capacity is capable of changing in size in response to changes in task demands [28]. The participants’ attentional effort, invested in the driving task and therefore in their situational awareness, may have decreased during the first six curves when they had the assistance of the system. However, this is just the other side of the coin. In this particular case, the other factor is the mental representation of the situation the participant formed according to the content of the HUD information. The participants’ mental models of the task did not include in this point the possibility of false information after the six successful curves. After the first false messages they still doubted their own comprehension instead of the system. In addition, mental demand in the co-driver system condition was in fact reported higher, not lower, than without the system. The mental workload got presumably higher especially after the curve with the first false message. The malleable attentional resources theory could explain the errors made in the curve with the first false message, but is insufficient for explaining the rest of the errors with the false messages. This means that we often have to pay more attention to the contents of mental representations than to the attentional load, when we consider the psychological explanations for errors [11].

Only after noticing the falsity of the messages did the drivers begin to doubt the system. Finally, at the point when they ceased to trust the system, they stopped following the guidance messages – at least consciously. The conclusion can be drawn that these kinds of safety-critical in-car systems should not give a single message wrong, because they are risky for the drivers as long as the drivers trust the system. In real life, rare errors by the system may be even more dangerous than frequent ones. Especially the importance of correct recommended speed information seemed to be critical. In a real situation, the combination of seemingly correct information about the other properties of a curve with wrong speed information could be very dangerous.

The results pinpoint the importance of technical robustness of in-car ubicomp systems so that they will not give false information in any possible situation. This can be assured by e.g., regular updating of databases and self-diagnostic systems. No information at all will be better than information without confirmation in situations described in our experiment. If there is even a minor possibility for false information, the driver should be made aware of this. The results indicate the importance of the system’s transparency and understandability, so that the driver will become aware of the limitations of the system. Further research should focus on the possible ways to make drivers aware of the possibility of wrong information at early stages of use. However, the resulting system could be in contradiction with the ubicomp ideal of system invisibility [2].

The co-driver system could be useful in a real automobile, assuming that it would not give false guidance messages. The system’s benefits would become obvious especially in hairpin bends of certain mountainous environments (e.g., Norway or Switzerland). Furthermore, the co-driver system could have a beneficial impact in unfamiliar and difficult driving conditions such as dark and foggy weather.

In the experiment, the mean number of driving errors decreased significantly between the trial without the system and the trial with the co-driver system giving correct information. The same kind of positive phenomenon can be observed with for example tutoring [21] or navigation systems. The co-driver system had elements like turn-by-turn guidance, the direction and distance to the curve, and verbal instructions which are also found in navigation systems. Therefore, the effects of false messages on driving with the co-driver system can also be basically applied to navigation systems: with the first false messages the driver does not necessarily doubt the system if the driver has built trust towards it. This can cause many kinds of hazardous situations, such as driving on a one-way street in the wrong direction. When the driver realizes that the system is giving wrong instructions, trust in the system will vanish, just like in our experiment. However, at this point the cautiousness of the driver may already come too late.

Theoretically, our main argument is that mental contents may provide us with an equally important conceptualization as mental capacity or workload in explaining human-system interactions.
interaction phenomena. These are not contradictory but complementary foundational concepts. We can use content-based psychological analysis in solving different types of problems. Capacity-based psychology is important when the system is too complex for the users, but it is not very good in analyzing information contents and related phenomena, such as expectations [29]. According to capacity-based thinking, we can fill the human working memory with any information, correct or incorrect, without causing problems as far as the capacity limits are not surpassed. This is why capacity is not a good basis for the scientific analysis of information contents in mental representations. It is impossible to differentiate between different information contents on the grounds of capacity. Consequently, it is necessary to develop a new content-based theoretical language to cope with the problems originating from mental contents [20], [30], [31].

V. FURTHER WORK

In the future, one of the major problems with these kinds of ubicomp systems will be: what happens if the system does not function correctly? How can drivers who are used to being aided by these systems adapt to situations without the system or to situations where the system is working improperly? Do they have the skills to drive without the systems?

By using the co-driver system as an example, we strived to demonstrate what are among the most important issues related to situation awareness to be considered when designing a new ubicomp system into an automobile environment. In addition to taking into account revised human-computer interaction design principles, we suggest that more attention should be devoted to investigating the interaction between the driver and the ubicomp system when something in the system goes surprisingly wrong. Our results suggest that focusing on the effects of mental under- or overload is not alone sufficient for understanding drivers’ situational awareness. Also the driver’s expectations for systems’ behavior should be put under careful scrutiny. It is evident that the role of mental contents will become increasingly more important when intelligence and knowledge types provided by ubiquitous systems essentially develop in the near future.

REFERENCES