INFLUENCE OF HMI ERGONOMY ON DRIVERS IN COOPERATIVE SYSTEMS AREA

Tomáš Jizba

Czech Technical University in Prague, Faculty of Transportation Sciences, Konviktská 20, Prague 1, Czech correspondence: jizbatom@fd.cvut.cz

ABSTRACT. Modern vehicles are getting smarter and utilize more and more the advantages of Advanced Driver Assistance Systems (ADAS). Deployment of upcoming technologies, such as cooperative systems (V2X), will most likely be the key step towards a significant reduction of accidents across the globe. Some of these systems extend driver’s field of vision, so the driver can be forewarned against a wide range of threats. Unlike technology, the human processing capacity has remained almost unchanged over centuries. Therefore, it is necessary to bear in mind that drivers have restricted capabilities to process multiple warnings. In this context, an important question arises: How can be V2X warnings integrated into the Human Machine Interface (HMI) of vehicles, and what warning policy is needed to ensure high usability, acceptance, efficiency and understanding of such a warning interface from the driver’s point of view.

Using a human centered design approach, two concepts of visual driver-vehicle interface for V2X warnings were developed and evaluated. One of those interfaces was based on a 1-stage warning policy and the second one utilized a dynamic 3-stage warning approach. The research revealed that drivers prefer to have dynamic warnings that correspond with the level of danger to which they are exposed, rather than only one warning displayed in a situation of imminent danger. However, with regards to overall usability score and workload measurements, results showed that the 3-stage system performed worse in comparison with the 1-stage system.

KEYWORDS: V2X, warning stages, GUI, ADAS, cooperative systems, HCD, usability, acceptability, UX.

1. INTRODUCTION

The Internet, as a powerful medium, influences our lives more than ever before and is starting to be present in distinct areas of human interest. One of these areas is transportation, where vehicles are no longer considered to be isolated units, but they are seen as a part of communication chain. In this sense, one can see transportation as one of the biggest representative of the so-called Internet of Things (IoT), where particular mobile nodes (vehicles) are capable of regularly exchanging their status via a wireless network. Related technologies are to be integrated within new vehicles to improve safety, comfort, efficiency and overall experience by communicating with traffic flow and infrastructure. According to long-term predictions it is expected that 75% of the estimated 92 million cars shipped globally in 2020 will be fitted with internet-connection hardware (smart OBU) [1]. These connected vehicles bring brand new opportunities for suppliers, OEMs and developers, but it is always necessary to bear in mind that the vehicle and traffic environment are very specific in terms of safety requirements. Especially during HMI design development, safety issues play an important role and markedly influence the final look of intended systems or applications.
1.2. SAFETY VIA VEHICLE-TO-X TECHNOLOGY

In spite a decreasing trend in recent years, the number of fatalities still represents a significant economical loss for society as a whole. In terms of accidents, lack of driver awareness and attention has been identified as the leading cause of car accidents with an estimated percentage of 26% to 56%. Nearly 90% of severe accidents could be positively influenced through an Advanced Driver Assistance System (ADAS)\[4\]. These systems integrate safety features and technologies designed to avoid collisions and accidents by alerting the driver to a potentially dangerous situation, or directly taking over control of the vehicle in case of such a situation. In upcoming years, the ADAS should utilize the available wireless connectivity of V2X to enhance passenger safety via mutual exchange of cooperative messages.

Regarding the V2X, these can be divided into several categories. One of those categories distinguishes V2X systems based on so-called warning stages, i.e. based on levels of warning that they provide to drivers. In general, one-stage warning systems provide drivers with only so called Imminent Crash Warning (ICWs), whereas two-stage warning systems provides ICWs plus so called Cautionary Crash Warnings (CCWs). Finally, a multi-stage warning system is based on a continuous warning strategy which means that the driver is regularly informed about the state of given event/danger. The question arises, how the driver handles several information outputs from more-stage warning systems (i.e., two and more stage systems) which are forwarded to him/her in short time proximity. One can expect that in certain situations, drivers will be facing multiple warnings at the same time. Since the amount of a driver’s attentional resources is limited, it is necessary to find a suitable way how to present these gathered warnings in order not to overwhelm the driver. Nowadays, visual displays represent the prevalent interface when it comes to warning of the driver inside the vehicle. Hence, it is essential to investigate how drivers react upon cumulated visual stimuli (i.e., the number of different warnings and information presented on a display) in different traffic situations.

1.3. GOALS AND OBJECTIVES OF THIS WORK

Within this work, two types of warning systems were tested, one on the basis of one-stage warning policy (see Figure 1) and the second one on the basis of multi-stage warning policy comprising three warning stages (see Figure 2). Human factors literature or several conducted studies (e.g. 8, 9) suggest that multi-stage warning systems enhance a driver’s performance over a 1-stage system. However, there is currently a lack of available data regarding visual distraction, workload, and driver comfort in order to favor multi-stage warning systems over 1-stage systems. Therefore, the main goal of this work was to assess the impact of these two visual warning interfaces on a defined user group concerning acceptability, usability and workload.

2. IMPLEMENTATION AND TESTING

The proposed systems were tested within an artificial environment, which should substitute the real traffic environment. The test in the real environment may be quite dangerous with respect to the intended purpose of the system. Therefore, a simulator study was chosen for the final evaluation of the systems and their mutual comparison.

2.1. SIMULATOR DESCRIPTION

The experiment was conducted on the CTU’s high-fidelity driving simulator (see Figure 3). The participants were seated in a Skoda Octavia II mock-up...
that was placed on a static base, and surrounded by 3 wide-screens on which the road and traffic environment was projected. The mock-up was fully equipped with a roof, mirrors, a passenger seat and original controls, such as pedals, a steering wheel, dashboard (speedometer), and automatic transmission. Sounds of the engine and blowing wind were simulated as well by using an Akai sound system with 7 loudspeakers situated in and around the mock-up. The experiment supervisor was seated closely behind the mock-up. There he had access to a control computer (MAC, Intel i5 2.7 GHz processor, 8GB, Intel Iris graphic card) which was connected to an in-vehicle display (7’’ iOS tablet) by means of a Lightning / USB cable. He could observe the projected environment, participant’s behavior, and additionally communicate with the participant during the sessions. Three PCs (Intel i7 processor, NVIDIA Quadro K4200 graphic card), Matlab, ZModeler and Rhino software were used for simulation purposes. The PCs generated real-time images with a refresh frequency of 60 Hz, and sent them to the three full HD (1920x1080) projectors. The projectors projected the images on a screen in front of the mock-up and on the side screens that were inclined to the front screen at 90°.

![Figure 3. Škoda Octavia II mock-up.](image)

As a centralized source of visual and auditory warnings a nomadic device in the form of a 7’’ tablet with a high-resolution display (2048 x 1536 px) was used. The tablet was situated above the middle console (see Figure 4), in a way that its position fulfilled recommendations from ESoP and JAMA, and simultaneously its presence did not obstruct driver’s field of vision. Moreover, research [12] suggests this position to be optimal in terms of reaction times, duration of lane departures, perceived workload and eye-movement latencies to presented stimuli.

### 2.2. Scenario Elaboration

For the impact assessment, a highway scenario containing several potentially dangerous events was created. The impacts were studied during conditions of normal visibility on the highway which consisted of 2x2 traffic lanes (see Figure 5) with approximately 10% of the traffic volume in the right lane and it contained passenger vehicles. The lanes were mainly straight with a few gentle curves. Each participant encountered two traffic jams during experimental trials. These traffic jams were approximately 3000 meters apart from each other and their regulation was based on the vehicles speed. For example, vehicles had to slow down when approaching the jam (i.e., 50 km/h at the tail), drive in a stop-and-go mode in the jam (i.e., 0-50 km/h) and accelerate when leaving the jam (i.e., towards 100 km/h). At the beginning of each traffic jam emerged temporary heavy braking vehicles, which created another dangerous event for participants. Besides traffic jams, the scenario contained one place with a shallow puddle (7x7 meters) simulating a slippery road across both lanes (see Figure 5), and one place with an immobilized vehicle on the edge of the right lane.

![Figure 4. The tablet device within the cockpit.](image)

![Figure 5. Highway simulation with dangerous events.](image)

With respect to the ego vehicle, these dangerous events changed their urgencies. Since some of them only occurred temporarily, the triggered warnings differed in the 1-stage warning system compared to the 3-stage warning system. The reason was the warning policy which differed in both systems. Whilst the 3-stage warning system could display all three levels of warnings, the 1-stage warning system displayed only the most imminent warning in the given time, i.e.
A total of 18 participants took part in the driving simulator experiment. The average age of the participants was 32.5 years (SD = 10.05 years). Table 2 shows some personal characteristics of the participants.

The main experiment consisted of one training session and three experimental sessions, with a total duration of approximately 1 h. In the first experimental session, participants were informed that they will drive through a highway test track without any assistance system (i.e., without any secondary task). The participants were asked to drive in the right lane and maintain driving speed 100 kph unless they would be forced by the environment to change this speed. They were also told to brake only if necessary. After driving for about 10 minutes through the highway test track and responding to 8 dangerous events, the simulation ended. At the end of the first session, participants were asked to evaluate their experienced effort (RSME scale [14]) while driving on the test track.

Before the second and third experimental sessions, the participants were given pre-measurement Van der Laan questionnaires [15] in order to assess primary acceptability of both systems (1-stage and 3-stage). After the completion, an eye-tracker calibration followed using a head-mounted eye-tracking device SMI iView X HED. Once the calibration process was done properly, one could obtain an accurate estimate of the gaze position (gaze pointer) within the scene, together with particular durations. In this work, five glance characteristics were measured, namely: Mean Glance Duration, Total Number of Glances, Total Glance Time at the in-vehicle device, Maximum Glance Duration, and Standard Deviation of Glance Durations (i.e., variability of glance durations).

Regarding the driving instructions (i.e., speed, etc.), these remained the same as in the first session. Dual-task sessions were randomly reversed for each participant, so that some participants started with the 1-stage warning system and then proceeded with 3-stage system or vice versa. After each ride there again followed the questionnaire with the subjective workload assessment (RSME). When the participants finished both experiment sessions, the experiment supervisor gave them post-measurement Van der Laan questionnaires concerning systems’ after-acceptability. After filling out these questionnaires, the participants were additionally interviewed regarding usability (SUS) by using 5-point Likert scales. The whole experiment was finished by a free discussion with the participant.

### Table 1. TTC threshold for particular warning stages.

<table>
<thead>
<tr>
<th>Warning stage</th>
<th>Visual cue</th>
<th>Timing (TTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>🔄</td>
<td>&gt; 13s</td>
</tr>
<tr>
<td>2nd</td>
<td>🔄⚠️</td>
<td>3s &lt; TTC ≤ 13s</td>
</tr>
<tr>
<td>3rd</td>
<td>⚠️</td>
<td>≤ 3s</td>
</tr>
</tbody>
</table>

### Table 2. Personal characteristics of participants (n = 18).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>12 (66.7%)</td>
</tr>
<tr>
<td>Female</td>
<td>6 (33.3%)</td>
</tr>
<tr>
<td>Age (18-30 years)</td>
<td>8 (44.4%)</td>
</tr>
<tr>
<td>Age (31+)</td>
<td>10 (55.6%)</td>
</tr>
<tr>
<td>Driving license (≤ 6 years)</td>
<td>9 (50%)</td>
</tr>
<tr>
<td>Driving license (≥ 7 years)</td>
<td>9 (50%)</td>
</tr>
<tr>
<td>Smartphone while driving</td>
<td>15 (83.3%)</td>
</tr>
</tbody>
</table>

Only the high-level warning with the highest urgency.

Every new warning or new state of the warning was accompanied with an auditory cue in the form of a sound beep. This was supposed to direct driver’s attention towards the in-vehicle display, since the visual warning itself is not omnidirectional. The timing of warnings was based on the Time-To-Collision (TTC) parameter that determined individual warning stages (see Table 1).

### 2.3. Participants an Procedure

A total of 18 participants took part in the driving simulator experiment. The main experiment consisted of one training session and three experimental sessions, with a total duration of approximately 1 h. In the first experimental session, participants were informed that they will drive through a highway test track without any assistance system (i.e., without any secondary task). The participants were asked to drive in the right lane and maintain driving speed 100 kph unless they would be forced by the environment to change this speed. They were also told to brake only if necessary. After driving for about 10 minutes through the highway test track and responding to 8 dangerous events, the participants arrived at the tail of the second congestion where the simulation ended. At the end of the first session, participants were asked to evaluate their experienced effort (RSME scale [14]) while driving on the test track.

### 3. Results Elaboration

#### 3.1. Workload Assessment

##### 3.1.1. Glance Behavior

Out of 18 tested participants in this experiment, 16 provided adequate eye-movement data. In the Table 3 one can find mean and standard deviation for each glance characteristic in relation to the warning system (1-stage/3-stage). Furthermore, a paired-samples t-test was chosen to test significant differences between both systems. Therefore, the table also contains t values, degrees of freedom, p values, and effect sizes for each pair (if significant difference).

There were four statistically different pairs of glance characteristics. It was no surprise that there were statistically significant differences in the Total Glance Time (t(15) = -12.974, p < .001) and the Number of Glances (t(15) = -15.179, p < .001) between both systems. The drivers in the 3-stage trial were confronted with a lot more warnings compared to the 1-stage trial, which resulted in a higher division of attention (see the Total Glance Time in the Table 3). Besides, warnings in the 3-stage system occupied the display much longer, which tempted drivers to look at the in-vehicle display more often. In any case, the drivers spent a significantly higher portion of time observing the in-vehicle display with the active 3-stage warning system (19.551 ± 5.722 s) compared to the 1-stage.
warning system (3.499 ± 1.041 s). Apparently, the higher Total Glance Time was conditioned by a larger Number of Glances on the display (20.56 ± 4.72) in comparison with the 1-stage system (3.75 ± 0.93).

A higher number of visual stimuli on the display probably caused a significantly higher Maximum Glance Duration (1.413 ± 0.165 s) while driving with the 3-stage system compared to the 1-stage one (1.103 ± 0.227 s). The 3-stage warning interface resulted in a more complex and more time demanding visual search task compared to the 1-stage warning system. Especially in the situation, where the drivers were confronted with multiple warnings at the same time, their ability to scan the 3-stage warning interface decreased, resulting in higher glance durations. Another reason behind the higher Maximum Glance Duration was the complexity of testing scenario. The low-level congestion warnings were triggered on straight sections of the highway with perfect visibility, so some drivers spent a longer time observing the interface without jeopardizing their safety.

The last statistically significant difference was found for Standard Deviation of Glance durations (t(15) = -5.467, p < .001). As one can see in the Table 3, the 3-stage warning system led to a higher dispersion (variability) of glance durations (.276 ± 0.035 s) compared to the 1-stage warning system (.149 ± 0.082 s). While using the 3-stage system, drivers tended to alternate their glance behaviour. This was done by alternating saccadic movements with a longer length of fixation and saccadic movements with a shorter length of fixation. This glance behaviour and glance characteristic, respectively, may most likely imply a higher workload of drivers using the 3-stage system.

### 3.1.2. Subjective Measurement – RSME Scale

In terms of subjective workload, RSME scores (Figure 6) revealed that drivers perceived the highest effort (47.67 ± 22.02) while driving with the 3-stage system. Less effort drivers perceived in case of the 1-stage system (35.33 ± 15.87), and they felt the least burdened when they were driving without any system (28.83 ± 13.99). Paired samples t-test showed statistically significant difference in means for all three combinations (see Table 4).

![Figure 6. RSME Scores.](image)

These findings correlate with the glance behaviour of drivers (3.1.1–3.1.1), so the drivers’ attentional resources were more burdened while driving with the 3-stage warning system than with the 1-stage system. In case of the 1-stage system, the perceived effort can be qualified as a little, whereas some effort or rather a lot of effort was perceived during the interaction with the 3-stage system.

### 3.2. Usability Results

The participants were given instructions to mark a response to every statement within the SUS, and were asked not to spend too much time on any one statement. After the comparison of both means of SUS scores, one can conclude that the 1-stage system was perceived significantly more usable than the 3-stage system (see Table 5).

Regardless of the lower score of the 3-stage system (76.11 ± 7.96) compared to the 1-stage system (84.31 ± 5.54), both systems can be declared as usable. If the scores would be transferred to the adjective scale (see y axis in the Figure 7), the usability of the
<table>
<thead>
<tr>
<th></th>
<th>Baseline $\mu(\sigma)$</th>
<th>1-stage $\mu(\sigma)$</th>
<th>3-stage $\mu(\sigma)$</th>
<th>t</th>
<th>df</th>
<th>p(two tailed)</th>
<th>Effect size (Glass $\delta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 1-Stage System vs. 3-Stage System</td>
<td>–</td>
<td>35.33 (15.87)</td>
<td>47.67 (22.05)</td>
<td>-2.345</td>
<td>17</td>
<td>.031</td>
<td>0.67 (Medium effect)</td>
</tr>
<tr>
<td>Pair 2 Baseline vs. 1-Stage System</td>
<td>28.83 (13.99)</td>
<td>35.33 (15.87)</td>
<td>–</td>
<td>-3.472</td>
<td>17</td>
<td>.003</td>
<td>0.46 (Medium effect)</td>
</tr>
<tr>
<td>Pair 3 Baseline vs. 3-Stage System</td>
<td>28.83 (13.99)</td>
<td>–</td>
<td>47.67 (22.05)</td>
<td>-3.195</td>
<td>17</td>
<td>.005</td>
<td>1.35 (Large effect)</td>
</tr>
</tbody>
</table>

Table 4. Paired-Samples t-test – RSME.

<table>
<thead>
<tr>
<th></th>
<th>1-stage $\mu(\sigma)$</th>
<th>3-stage $\mu(\sigma)$</th>
<th>t</th>
<th>df</th>
<th>p(two tailed)</th>
<th>Effect size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUS Scores</td>
<td>84.31 (5.54)</td>
<td>76.11 (7.96)</td>
<td>3.615</td>
<td>17</td>
<td>.002</td>
<td>2.10 (Large effect)</td>
</tr>
</tbody>
</table>

Table 5. Paired-Samples t-test – SUS.

Figure 7. SUS Scores.

1-stage system would be evaluated almost as an excellent, whereas the usability of the 3-stage system would be marked as good. The reason behind such a significant difference in the usability scores is most likely a much simpler interface of the 1-stage system with substantially reduced functionality (e.g., only one warning per screen, no transitions of warnings, no position cue, etc.).

3.3. Acceptability – Van der Laan Scores

In Table 6 the results from pre-trial and post trial Van der Laan scales are shown. The table contains mean usefulness and satisfaction scores for both systems together with their standard deviations.

As it is apparent, both systems were rated equally in terms of pre-trial usefulness, and both suffered a certain decline in this score after the trials. In case of the 1-stage warning system, this decline was found to be statistically significant with $t(17) = 2.486$, $p = .024$ and $d = .59$ (medium effect). In terms of the satisfaction score, the participants showed negative trend after the trial with the 3-stage warning system, but with no statistical significance ($t(17) = 1.256$, $p = .226$). Even though the participants had higher expectations before using the systems, the overall attitude towards both systems seems to be positive (see Figure 8).

Figure 8. Post Van der Laan scores.

Apparently, the participants felt slightly more satisfied while driving with the 1-stage warning system. On the other hand, the 3-stage warning system was rated as more useful. Nevertheless, these differences were not found to be statistically significant.

3.4. Final Decision

At the end of the whole experiment, the participants were asked to make a final decision, i.e. which of these systems they would like to have in their cars. The results (Figure 9) showed that 66.7% (12) of participants would rather use the 3-stage warning system, whereas the rest 33.3% (6) of them would prefer the 1-stage system.
4. Conclusion

Off-road glances (or glances to in-vehicle devices) are of particular concern for evaluating the potential for distraction from HMI systems. Glance characteristics, such as longer glance durations and/or a higher number of glances at an in-vehicle GUI may indicate its unintelligible design for the user resulting in higher investments of his/her attentional resources and in higher workload in general. Since the resources could otherwise be solely used for the primary task, their exploitation must be time effective. This work tried to evaluate two kinds of warning systems which were based on different warning strategies. Despite the more complex visual search task in the 3-stage warning system, results showed that there was no single glance to the display that was longer than 2 s threshold (see [16–18]) made by any of the 16 participants. Only one participant experienced glance duration longer than 1.6 s (1.68 s). Both systems, therefore, can be considered to be safe in terms of glance behaviour.

In spite of the higher perceived workload and the lower usability score, the drivers would probably equip their vehicles with the 3-stage warning system. However, further studies should be conducted to assess driving performance (e.g., braking reaction time, lane deviations, etc.) of drivers while interacting with this system. An optimal solution seems to be to switch between both systems depending on the current workload of the driver. In case of the higher workload, the 1-stage system would be active, whereas in less attention demanding conditions the 3-stage system would warn the driver continuously. This could be done via so called workload manager [19] that monitors the environment (e.g., lane parameters, weather, etc.) as well as the driver state (e.g., eye-blinks, heart rate, etc.), and makes decision about when information should be presented to the driver and in what form. Such an alternating policy may best address driver’s processing bottlenecks and may lead to an efficient human-system interaction without any impairment of road safety.

5. Recommendation for Future Research

It is clear that the aforementioned glance characteristics could differ in more complex scenarios with respect to the road type, road curvature, position of the in-vehicle device, environmental conditions (e.g., fog), surrounding stimuli (e.g., more stimuli in cities), and/or during additional task(s) at hand (i.e., during higher distraction). Within this study, the conditions were very convenient in terms of potential workload. Hence, additional experiments should be conducted to prove the designs under different conditions including the real world conditions.

One could also include a 2-stage warning system for comparison, i.e. a system with solely mid-level and high-level warnings. Many participants stated that the 1-stage system was useless in certain situations, because the dangerous event was visible in advance. They also pointed out that low-level warnings within the 3-stage system were showed too soon before the dangerous events. This may lead to so called "cry wolf" effect [20], which could harm the credibility of the whole 3-stage warning system. The 2-stage warning system could therefore represent a suitable compromise between these two systems. In addition, deployment of so-called Likelihood Alarm Display [21] could significantly enhance the credibility of multi-stage systems by providing likelihood of event presence.

The study also revealed that for half of participants the number of warnings was too high in certain situations, i.e. the screen was too cluttered. Therefore, an additional study should find out how many warnings should be presented per screen to increase situation awareness.

Finally, the warning modalities and their combinations, respectively, should be further examined. In particular, the low-level warnings could benefit from a haptic feedback, which could reduce the startle effect from sound cues and simultaneously supplement non-omnidirectional visual warnings.

<table>
<thead>
<tr>
<th>System</th>
<th>Usefulness</th>
<th>Satisfaction</th>
<th>Usefulness</th>
<th>Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Stage</td>
<td>1.356 (.448)</td>
<td>.653 (.412)</td>
<td>.956 (.525)</td>
<td>.653 (.543)</td>
</tr>
<tr>
<td>3-Stage</td>
<td>1.356 (.596)</td>
<td>.708 (.502)</td>
<td>1.133 (.672)</td>
<td>.458 (.729)</td>
</tr>
</tbody>
</table>

Table 6. Van der Laan – Pre-and Post Scores
ACKNOWLEDGEMENTS
This study was conducted as a part of other studies included in a double degree programme named Intelligent Transport Systems at the CTU in Prague and the FH Technikum Wien. The experimental part of the work was carried out at the Department of Vehicle Technology at the CTU, so I would like to thanks all the staff from this department who enabled me to use their simulator and equipment. Special thanks goes to Mr. Jan Válek, without whom it would not have been possible to carry out any experiment.

REFERENCES