This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

Evaluating mixed reality notifications to support excavator operator awareness

Markus Wallmyr[1, 2] Taufik Akbar Sitompul[1, 2] Tobias Holstein[1, 3] and Rikard Lindell[1]

[1] School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden
[2] Department of Product Management, CrossControl AB, Västerås, Sweden
[3] Hochschule Darmstadt, Darmstadt, Germany
markus.wallmyr@gmail.com,{taufik.akbar.sitompul, tobias.holstein, rikard.lindell}@mdh.se

Abstract. Operating heavy vehicles, for instance an excavator, requires a high level of attention to the operation done using the vehicle and awareness of the surroundings. Digital transformation in heavy vehicles aims to improve productivity and user experience, but it can also increase the operators mental load because of a higher demand of attention to instrumentation and controls, subsequently leading to reduced situation awareness. One way to mitigate this, is to display information within the operators' field of view, which enhances information detectability through quick glances, using mixed reality interfaces. This work explores two types of mixed reality visualizations and compares them to a traditional display setup in a simulated excavator environment. We have utilized eye-tracking glasses to study users' attention to the task, surrounding awareness, and interfaces, followed by a NASA-RTLX questionnaire to evaluate the users' reported mental workload. The results indicate benefits for the mixed reality approaches, with lower workload ratings together with an improved rate in detection of presented information.

Keywords: Mixed reality, Human-vehicle interaction, Situational awareness, Head-up display, Excavator, Heavy-vehicles.

1 Introduction

Many industrial sectors, such as agriculture, construction, material handling, forestry, mining, and transportation, rely on mobile machines to support their work, which generally referred to as heavy vehicles. Operating these vehicles can be a complex and demanding task. The operator may be required to drive the vehicle while performing a task that requires additional attention and control, for example lifting an object. Completing the task can also require a high level of precision and be performed in co-operation with other workers, where un-safe operation causes unnecessary risk to the health of nearby people. To assist the operator, modern heavy vehicles are increasingly being equipped with information systems and assistive systems, where the interaction is done via a visual user interface. A result of this transformation is that many heavy vehicles are now equipped with multiple displays inside the cabin, including displays from both the original manufacturer and third-party equipment suppliers. It is not uncommon for an operator to have more than five displays: one main instrument cluster, one secondary vehicle display, a navigation system display, one or more displays for add-on equipment, and finally, one or two camera monitors. With more and more information displayed visually to the operator, the cognitive load on the operator is becoming untenable leading to increased mental fatigue. The current state of affairs prompted Sanches et al. [1] to state that “the need for research that informs the design of effective, intuitive, and efficient displays is a pressing one”.

Moreover, operators often spend a large portion of their working days operating the machine and they are required to be both efficient and safe while carrying out their tasks. Reports of past accidents show that the human factor has a big impact on accidents. For example, a recent study on earth moving equipment show that 46% of the cases involved misjudgment of the hazardous situation by the operator [2]. In addition, an analysis of wheel loader accidents shows that incidents where mining and non-mining personnel were hit, struck or run over by the wheel loader are the most common wheel loader-related accident, which represent 41% of the total fatalities [3]. Assisting operators’ awareness could thus help improve the safety and wellbeing of operators, surrounding workers, and pedestrians.

While the intent of an increase in-vehicle information systems is to aid the operators performing their tasks and allow them to be more informed about the operation, one possible drawback of these additions is that the operators will spend more time focusing on displays and digital information, resulting in reduced attention to the surrounding environment and potentially leading to an increase of hazardous situations [4]. Alternatively, the operators might miss vital information that is displayed, either due to a high mental workload or that the information is presented outside of their main field of view [5]. These drawbacks could result in a lack of situational awareness [6], when the operator is missing critical elements needed for a full comprehension of what is happening and anticipation of what will happen. A lack of situation awareness has been listed as one of the bigger causes of accidents in safety-critical situations [7]. Studies on how to support operators’ awareness show that safety-enhancing systems, for example collision avoidance systems, can assist operators while carrying out their tasks [8, 9]. However, having to look away in order to obtain the information can result in an increased risk of incidents. A notable example of this situation is the Llanbadarn Automatic Barrier incident report, where a train passed a crossing with the bars raised. The immediate cause of this incident was that the operator of the train was occupied with the machine interface and therefore failed to check the crossing indicator [4].

One way to improve the operator’s situation awareness is to study various approaches in which the information can be communicated more naturally integrated into their normal workflow. For example using mixed reality interaction technologies such as head-up displays where information can be presented within the operator’s line-of-sight through the windscreen [10, 11] . This might help the operator to effectively detect and obtain the presented information, thus potentially reducing the risk

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

of missing any vital information. Another approach is to use other senses, but these senses are also already in use in this context. For audio, there are existing audio alerts, the communication between co-workers, and noise pollution from machines and the surrounding area. Some operators that we interviewed also describe how they use hearing already, for example through the different sounds that an excavator's bucket makes when touching various materials. This audible feedback can reveal if the operator is digging gravel, hitting a rock, or colliding with a buried pipe. The tactile channel is also affected by existing vibrations from the engine and hydraulics of a heavy vehicle and through the movement in the terrain during groundworks. Moreover, there are prior studies [12–14] that indicate a strong user preference for the visual channel, compared to other senses, which can then be used to improve performance further.

In this paper, we evaluate how visual cues deployed through mixed reality displays might affect information processing and situational awareness of heavy vehicles' operators. More specifically, by comparing the performance of test subjects using a traditional head down display against two types of mixed reality interfaces when displaying navigational information, as well as warning messages for obstacles in an excavator, which is a common type of heavy vehicle. The excavator is a versatile machine that can be used for many purposes, such as digging, material handling, demolition and groundworks. Operating an excavator puts a significant mental workload on the operator [15]. While performing the tasks at hand, the operator controls the boom and the attachment along all 3 axes, often with high accuracy requirements. Additionally, the operators also have to pay attention to their immediate surroundings, in order to prevent accidents involving people or objects.

The paper will first cover related work, followed by an overview and explanation of the visual information presented to the operator. It will then present how the user evaluation in a simulated environment was performed, followed by its results, including both quantitative and qualitative findings. After a summarizing discussion on the results we wrap up this paper with limitations of the study and outlines for future work, as well as the conclusion.

2 Related work

Research on the use of mixed reality displays to aid users with their task completion stems as long back as the 1940s, where head-up displays were used to aid pilots of military aircraft and has since been explored in many domains [16, 17]. Mixed reality displays have also shown potential in the automotive vehicle sector for their possibilities to increase safety and aid user experience [18–20], to reduce workload and increase driver comfort [21]. For example, a head-up display speedometer produces improved performance and shortened response time to hazardous situations [22] as well as better speed control [11].

While the vast body of work around mixed reality displays for vehicles is related to on-road vehicles, there is also a potential use for heavy vehicles [23]. One related application area, that has been both researched in academia and used in industry, is

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

the use of virtual and augmented reality for operator training and simulation. These studies suggest that VR and AR-based training can improve the quality of training for becoming operators [24, 25].

One example of evaluation in heavy vehicles, is in the forest harvester, where Englund et al. [26] investigated the result of moving the graphical interface from a traditional display to a head-up display. They found that the operators were positive about seeing system information closer to the area of operation. This work was however limited to a productivity perspective, where production data that was normally presented on a head-down display was moved to a head-up display.

Using mixed reality displays has also shown potential to improve the ergonomic situation, both in terms of comfort and fatigue [27]. This is an area of note, as one of the major reasons for work-related injuries in heavy vehicles, is awkward postures of the operators [28]. Furthermore, Akyeampong et al. [15] evaluated two different interface designs to improve usability and ergonomics while controlling an excavator. Their design was based on an augmented interaction style and their setup used a Virtual Reality (VR) simulation using VR headsets for visualization. They concluded that “the major strength in both HMIs designs was the heads-up display concept” .

Taking prior research into account, there is still work needed to be done to understand how mixed reality displays and augmented interaction can be used in heavy vehicles, both in terms of the implementation of the technology and interaction design. For example, on the use of mixed reality displays to improve excavator operators comfort and safe operation, by presenting warnings and navigational cues closer to the operators line of sight.

3 Conceptualization

Heavy vehicles generally have a cabin with large windscreens to enable the operator to view the surroundings around the machine (see examples in Figure 1). This is important for both the general operation, as well as to be able to see obstacles and hazards around the machine. Moreover, displays in the cabin are normally placed quite

Fig. 1. To the left, a cabin with large windscreens for good visibility. To the right, an example of a cabin with four displays covering the windscreen, including one display which is not visible

low, as to not obstruct the operator’s field of view. This low placement puts the displays outside the operators' area of attention with the apparent risk of the operators missing key information. The size of the display is often limited due to space or cost, thus restricting the amount of information that can be comfortably displayed and reducing the size of presented information. Modern vehicle displays commonly range from 3 to 15-inch displays, with some niche examples of both smaller and larger displays. With more information being presented, additional screens might be needed to fit the new data. Due to space limitations, these displays might be placed more in-line-of-sight. However, this placement causes additional drawbacks, since the displays are now hindering the operator's ability to monitor the surrounding area. We asked ourselves, what if we could use the windscreen as the display to keep the operator notified in terms of navigation and the presence of potential obstacles, without obscuring visibility of the surrounding area.

3.1 Using the simulator to create situations for ideations

The approach of this experiment was to perform the evaluation in a simulated scenario to understand how the information that is presented within line of sight of the operator would be used. Prior observations in the simulator have shown difficulties in obstacle detection and difficulties to know the position of the bucket in relation to other objects. We thus sought to support the operator to be aware of obstacles while progressing through the track, as well as to support the navigation of the excavator and its boom. For further insights, we drove the excavator through the simulation to detect challenging areas, as well as recording them through screenshots. Examples of

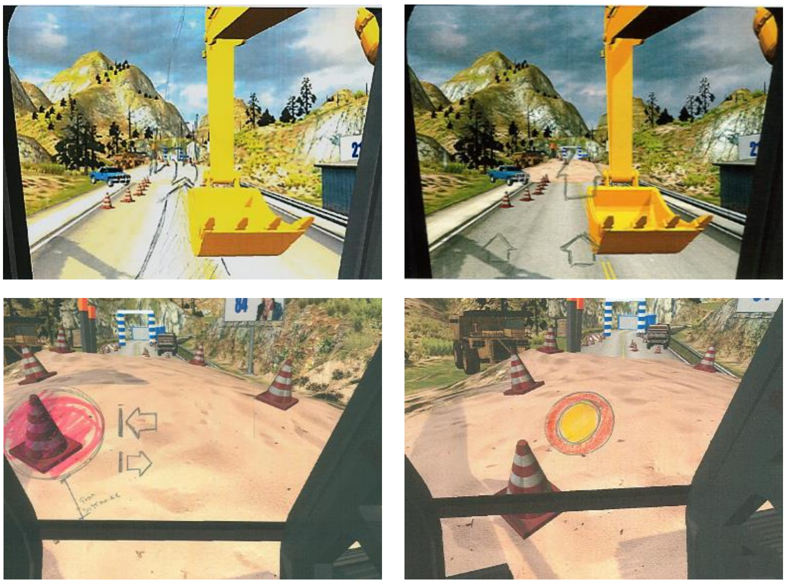

Fig. 2. Example of pencil sketches of navigational support as well as warning notification for the operators

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

the challenges are the course presented included tight paths where the user will be close to obstacles, as well as human avatars positioned where the machine structure hindered visibility. We then used these screenshots as the base for sketching. This enabled us to ideate based on the environment in which the users were operating.

From these screenshots, several designs were made and discussed internally. Fig. 2 shows some examples of the sketches that we made, where the upper figures show images of how to provide graphical aid to help the operator to navigate. The lower images show examples pictographic icons for obstacle notifications.

3.2 Operator support and virtual scenario

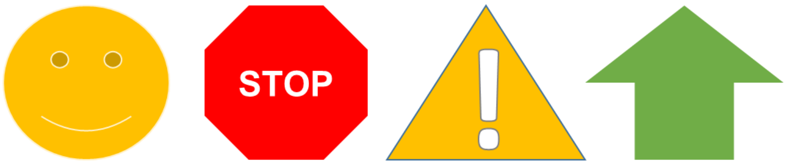

Fig. 3. The notification symbols used to notify the operator in the simulated environment. From left: human warning (shown with stop or warning), stop, warning, navigation direction.

A visual warning system was designed based on the sketches. The aim was to aid the user to perform the assigned task in the track as well as to support the user in detecting hazards. This system consisted of simple graphical figures shown on a display. The selected design of the symbols aimed to be naturally recognizable by the user, by using familiar polygonal shapes resembling traffic signs, see Fig. 3. Whether this assumption was successful in this context is however not verified, as the aim of the study was to evaluate the different placement of information rather than the information design as such.

When an obstacle was close to the excavator, a warning triangle was shown together with a numeric value indicating the distance to the obstacle. When the excavator was in imminent danger to collide with an object, a hexagon-shaped stop sign was shown in place of the warning icon. When an avatar, representing a virtual human, was close to the excavator, an additional circular warning sign was shown besides the other warnings. The user was free to look around and operate the machine within the open 3D environment. For direction guidance there was an arrow, pointing from the excavator's bucket towards the next pillar in the scenario, similar to the arrow icon commonly used in automotive and mobile navigation systems.

Fig. 4 shows a schematic illustration of the scenario used. The purpose was to represent a varied set of operator challenges, as will be further described in the qualitative results. The path was made with inspiration from field observations, for example where an operator had to move through a narrow passage, as well as to be guided in a precision bucket operation task. We selected to not involve a digging task, as this would have resulted in a static scenario with machine rotations. The main task in the scenario was to demolish three pillars, comprised of stacked cubes placed on top of each other and located to the side of the machines proscribed path. The first pillar was

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

placed in the beginning of the track and the two last pillars at the end of the track. The pillars were of different heights and their top two cubes were colored orange. The user was supposed to use the excavator's bucket to tip the pillars by knocking the orange cubes placed on their top. Getting to the pillar, the user had to navigate along a track with some narrow passages restricted by cones, a pile of sand with a curved path, and other vehicles in the area. Also, there were two human avatars along the track, one standing next to the path and another one constantly walking back and forth over the path.

The purpose of the creation of this scenario was to provide the user with several observable challenges and to force the user's attention into several different fields of focus. To complete the task successfully, the user needed to remain aware of the surrounding area and monitor several directions, such as the path forward, the ground surrounding the machine and the rotation of the excavator track, the bucket, as well as looking up when approaching the pillars.

Fig. 4. Left: A bird view of the virtual scenario.

4 Evaluation

To evaluate the use of mixed reality displays for presenting information and the use of symbolic notifications during operation, a study was performed with fifteen users. In this study the users completed a scenario with the help of the visual support system, using two types of mixed reality displays, as well as a reference in the form of a head-down display that represents the traditional monitor used in today's excavator cabins. Each of them displayed the same kind of information, which are the directional arrow and the obstacle detection presented above.

One of the mixed reality displays was a head-up display, in the form of a physical representation of a head-up display in the cabin's windscreen. The other mixed reality display was a projected display inside the virtual world, placed further away in front of the excavator. The head-down display was also placed on the operators' side, similar to the actual placement in existing cabins. The mixed reality displays are subsequently more in line of sight, while the head-down display requires the user to look away from the primary working area to look at what is presented on the display.

4.1 Simulator overview

The simulation was performed in a CAVE-like room. When compared to full-scale industrial vehicle simulators, it offers a cost-effective solution that can be easily reconfigured for prototyping. It has also been evaluated in favor of a traditional monitor-based simulation because the user can look around and is immersed within the virtual world [29]. Compared to VR-based simulations, it also enables the mixing of real and virtual content to test various forms of user interfaces.

The user sits on a chair, whose left armrest is equipped with a head-down display and a keyboard to control the excavator's movement. The excavator's boom and bucket were controlled using a traditional joystick interface, placed on the user’s right armrest. In front of the user there was a prototype head-up display made by using a transparent film, acting as a combiner, attached to a metal frame (0.6 x 0.9 m). The image created was full color and was generated from a projector, placed directly under the display. The virtual world was projected by a head-worn laser projector onto the reflective wall material, giving a wide field of view of 83.6° x 47.5°. This creates an undistorted image, centered in front of where the user was looking.

Fig. 5. To the left, a schematic picture of the CAVE simulation room and its primary components. To the right, a photo that shows all three display solutions at once (the simulation only showed one at a time)

The mix of different sources of image generation and traditional displays deployed in this particular setup works in this arrangement due to the specific combination of projectors and reflective materials. The image on the head-up display was designed to be clearly visible and achieved due to the bright projector, rated at 1200 lumens. The head-worn projector, with a value of just 15 lumens, emits a light level low enough not to create interference with the head-up display, even when faced directly. But its image was still clearly visible on the highly reflective material used for the CAVE walls.

The virtual environment and all visualizations are modeled and run using Unity [30], where a mobile phone attached to the head-worn projector hosts the virtual world. The head-down display computer hosts the visualization for the head-up display and the head-down display. Fig. 5 shows all the displays together, including the virtual content projected on the CAVE wall.

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

4.2 Interacting in the simulation

The excavator simulator allows the user to drive a virtual excavator and operate its boom and bucket. The body of the excavator is rotated through the joystick's horizontal axis. The lower excavator's boom is moved up and down through the joystick's vertical axis, while the upper boom is extended and retracted by twisting or turning the joystick around the Z axis. Finally, two buttons on the top of the joystick control the bucket movement of the excavator. The user can drive the excavator by pressing the keys on the numerical keyboard. One can argue that a keyboard is not a natural interaction device in an excavator, but the use of a keyboard made it natural and easy to understand the navigation of the excavator, especially for those who have experienced computer gaming. Performing the evaluation with professional drivers would probably have benefited from a joystick also for navigation of the excavator.

4.3 Additional evaluation apparatus

The user was recorded both using a side video camera and an eye-tracker. The digital video camera was placed facing the user within the simulator to obtain verbal comments as well as reactions while performing the scenario. The eye-tracking equipment was based on a pair of eye-tracking glasses from Pupil labs [31, 32], giving a recording of where the user is looking while performing the evaluation.

The equipment was calibrated for each user, using the software on-screen display calibration, before starting each test run. The accuracy of the gaze tracking is 0.6 degrees and the gaze precision is 0.08 degrees. The quality was adequate for detecting what part of the environment, or display, that the user is paying attention to. A more detailed study on exact areas of attention would require a more exact calibration.

Each user was also asked to fill in a NASA Raw Task Load Index (NASA-RTLX) form after completing the evaluation [33].

4.4 Users

The evaluations were conducted with fifteen users based on the guideline that already 10-12 users would suffice to provide a significant result [34]. Twelve of the participants were male and three were female. The eye-tracking result from three users had to be left out, due to the inability to capture their eye gazes. For the remaining users, two users had prior experiences of operating a real excavator, seven users had prior experiences of industrial vehicles or industrial simulators, and six users had prior experiences of a virtual reality headset or head-worn projection systems. The age distribution was defined in intervals, with no one under 25, eight users reported an age between 26-35 years, and four users being 35 or older.

4.5 Method

Before the tests started, users were informed about the purpose of the study and introduced to the equipment, the controls, as well as the task to perform in the scenario.

The users were evaluated one at a time. Each user was first equipped with the head-worn projection system as well as the eye-tracking equipment. The eye-tracking equipment was calibrated using the on-screen based calibration in the software. The recordings was started after the user was informed about the purpose of the recordings, how we would use them, and that they had given informed consent for participating in the experiment. Each user was also informed that he or she could discontinue the study at any time.

The user was then asked to complete the given scenario. When the scenario was completed, it was restarted, and the user was asked to do it with the next type of displays. Thus, each user performed the scenario for each of the display types. The order was randomized to avoid familiarity bias. Completing all three runs in the scenario took around 15 minutes, with some minor variation between each of the participants. Each run was generally completed faster than the user's prior run.

After completion of the test runs, the users were asked to fill a questionnaire to document their experiences with the prototype. The questionnaire was divided into two sections. The first part contains general questions about age and earlier experiences. This was followed by a NASA-RTLX questionnaire [33], which is used to rate the workload differences between the type of displays [35].

After finishing the evaluation, two approaches were applied to the results: collecting quantitative results and seeking for qualitative findings. The questionnaires were compiled into a spread sheet and the eye-tracking videos were manually examined where all gazes at the different displays were recorded, including information in which section of the track they occurred. From this data box plots were generated, as presented in the result's section. ANOVA evaluations were also performed, from where the significant results are presented. When visible in the recording, the occasions whether the presented information was overlooked, were also listed. Additionally, other observations detected during the eye-tracking review, i.e. the first-person perspective of the user evaluations, were also noted and compiled together. Finally, the verbal comments by the users were transcribed into text.

5 Results

The following section presents the quantitatively measured results followed by more qualitative findings. As a baseline, the two mixed reality displays are compared to the head-down display and we will later look at the differences between each version of the mixed reality displays.

5.1 Quantitative results evaluation

The NASA Raw Task Load Index (NASA-RTLX) evaluation was used to get a measurement of the workload related to the different display types [36]. NASA-RTLX is a version of NASA-TLX that is simpler to fill in by the participants, while still providing similar quality of result [35, 37]. The pen-and-paper version of the questionnaire was used [33] and the questions were filled in for each set-up after all 3 test runs were

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

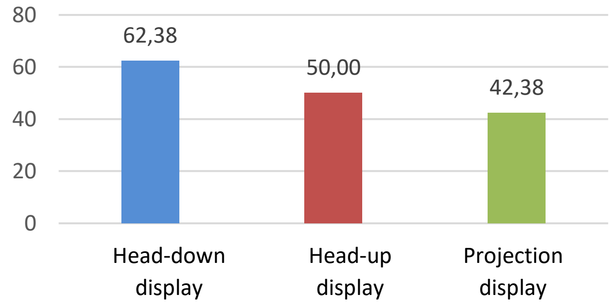

Fig. 6. Overall NASA-RTLX workload ratings for each display type.

completed. The NASA-RTLX consists of six workload categories: mental demand, physical demand, temporal demand, performance, effort, and frustration level. Each score is rated subjectively by the user on a 20-step scale (-10 to +10). The overall score for each type of displays was obtained by dividing the sum of the workload scores by seven, presented in Fig. 6.

The results show that the users rated the head-down display, placed at the left side of the armrest, with a higher workload rating than both mixed reality displays. An ANOVA test between the head-down display and each of the mixed reality displays show a sigificant difference between the head-down display and the projection display ($F(1,24) = 6{,}545$, $p{=}0{,}017$).

Looking at the sub-scale categories, mental demand, physical demand, effort, and frustration show the largest differences, which are presented as average values with standard deviations in parentheses in Table 1, as well as box plots in Fig. 7. It can be seen that the mixed reality displays showed better results in all categories when compared directly with the head-down display, with a statisticaly significant

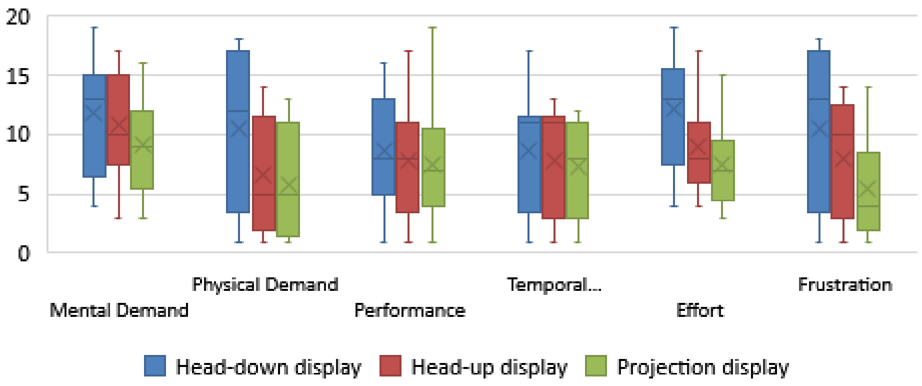

Fig. 7. NASA-RTLX subscale result for each type of displays represented with box plots. X represents the mean and the median is marked with a line separating the 2nd and 3rd quartile. The whiskers mark the largest or smallest data item not being an outlier.

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

difference between the head-down display and the projection display specifically for physical demand (F(1,24) = 4,747, p=0,039), effort (F(1,24) = 8,653, p=0,007), and frustration (F(1,24) = 5,061, p=0,034). The largest difference in the workload categories was for the physical demand, where the head-down display scored twice as high as both mixed reality displays. This result is probably due to the user being required to move his/her head and re-focus the eyes every time they wanted to look at the head-down display. This is also something that has been commented by some of the users, where they stated a lower motivation to look at the head-down display.

The subjective workload scored highest for the category of mental demand and effort for all type of displays. This supports the notion that running the excavator was a mentally demanding task, but it was not causing frustration to the same degree. Additionally, the projected display scored slightly better than the head-up display with lower mental demand, effort, and level of frustration. Even though we did not get any comments regarding this from the users, it could be an indicator that the effort of moving the focal depth has an impact on the general effort.

In conclusion, the NASA-RTLX ratings indicate that mixed reality displays could be favourable for both information assimiliation and for the operator's physical wellbeing, but only the projection display showed a statistically significant difference. A larger and more diverse sample set would be beneficial for firm conclusions, as is also related to in the future work section.

Table 1. NASA-RTLX's average scores per category with standard deviation in parentheses.

Subscale	**Head-down display**	**Head-up display**	**Projected display**
EFFORT	12,2 (4, 6)	9,00 (3,8)	7,5 (3,5)
FRUSTRATION	10,5 (6,8)	8,0 (5,0)	5,5 (4,4)
MENTAL DEMAND	12,6 (4,4)	10,9 (4,4)	9,1 (4,1)
PERFORMANCE	8,6 (4,5)	7,8 (4,8)	7,4 (4,9)
PHYSICAL DEMAND	9,4 (6,7)	6,5 (4,9)	5,8 (4,6)
TEMPORAL DEMAND	8,7 (5,4)	7,9 (4,7)	7,2 (4,2)

Eye-tracking data and side recorded video results.

Each recording was analyzed to observe when the user looked at the presented information on the display. This was done by manually logging the glances for each of the displays used. While each user looked at the different displays to a varying degree, all

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

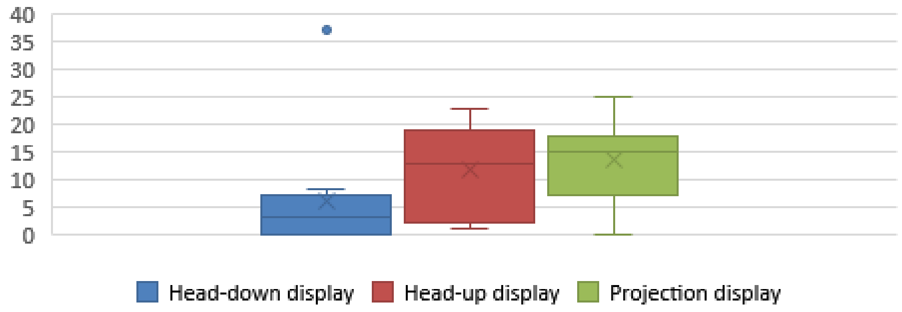

Fig. 8. The amount of display glances for each evaluated display type. The points outside the whiskers are considered as outliers.

users had a higher count of observed display glances for both of the mixed reality displays, in relation to the head-down display. On average, the users looked at the head-down display six times per run, the projected display 13 times per run, and the head-up display 12 times per run, see Fig. 8. The number of glances points to the assistive information being observed and used more when the information was closer to the users' line of sight. Taking away the outlier data from the sample set we get an ANOVA set of results between the head-down display and the head-up display at ($F(1,18) = 8{,}880$, $p=0{,}008$) and the projection display at ($F(1,18) = 16{,}419$, $p=0{,}001$).

Studying the different sections of the track shows a different usage for each display type, depending on the type of situation, seen in Figure 10 and Table 2. The first part, when the user tips the first pile, highlights an example of how the users generally used the supportive system; when unsure of what to do and when in need of additional information to be able to proceed. Here one user had issues in the beginning and thus used the display a lot more than the other users.

The next measured sequence was the avatar standing partly hidden next to the excavator path. The average gaze results show a higher detection rate of its warning symbol and its presence for the mixed reality displays, but not a significant difference according to the ANOVA analysis. The next measure shows the rest of the gazes detected while passing through the alley, where the ANOVA analysis of the data shows a significant deviation value for the projection display ($F(1,20) = 4{,}377$, $p=0{,}049$).

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

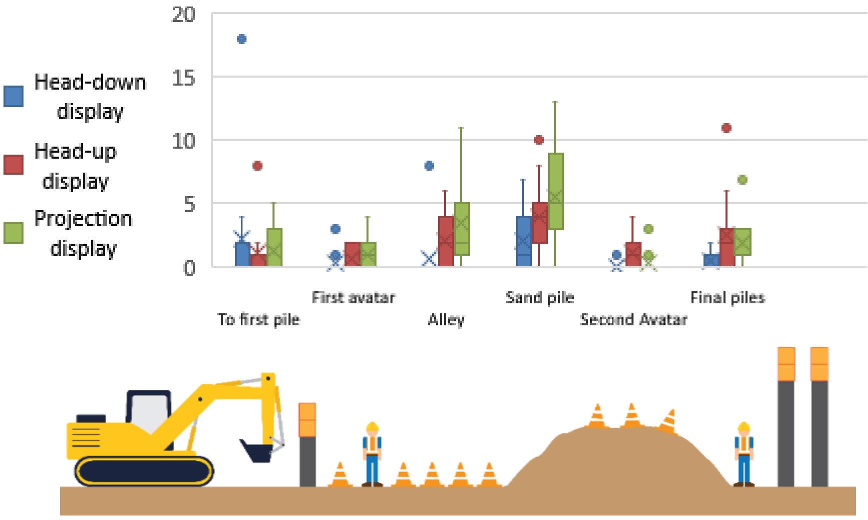

Fig. 9. Summed display glances for all users, divided in different parts of the track.

The navigation required during the task became more complex when approaching the sand pile and the number of glances at the displays was observed to increase to the highest level at any stage of the track, regardless of display types, as users sought increased support and guidance. Still, the data gathered for the use of the mixed reality displays continued to be higher than the use of the head-down display, with a significant difference in a higher level of use for the projection display ($F(1,20) = 5{,}880$, $p=0{,}025$). Reaching the second avatar, the users were not reacting to the notification being shown in the same level as with the first avatar, with almost no readings on the head-down display and the projected display. Respectively, the head-up display was detected by half of the participants ($F(1,20) = 7{,}912$, $p=0{,}011$). When the users were trying to tip the last two pillars to complete the assigned task, they had to look upwards towards the top of the pile.

Table 2. Average glances for the different display types

	First pile	**First avatar**	**Alley**	**Sand Pile**	**Second avatar**	**Final piles**
Head-down display	2,3 (5,4)	0,4 (0,9)	0,7 (2,4)	2,1 (2,7)	0,1 (0,3)	0,5 (0,8)
Head-up Display	1,2 (2,4)	0,7 (0,9)	2,1 (2,2)	4,0 (3)	1,2 (1,3)	2,6 (3,3)
Projection Display	1,3 (1,8)	1,0 (1,3)	3,5 (3,6)	5,5 (3,8)	0,4 (0,9)	1,9 (2)

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

Here both versions of the mixed reality displays show more use over the head down display, with ANOVA analysis for the head-up display being ($F(1,20) = 4{,}246$, $p=0{,}052$) and ($F(1,20) = 4{,}482$, $p=0{,}047$) for the projection display. This high discrepancy is probably due to the information being closer to the area of operation and the head-down display being quite far outside the users' field of view forcing the users to significantly alter their head position to view the head-down display.

5.2 Qualitative observations

To complement the quantitative data, especially the gaze measurements, we also searched for qualitative findings [38]. This was done by watching the eye-tracking recordings and observing information that could inform us about the user's behavior. The spoken comments expressed by the users while running the excavator were also transcribed. This section will present a summary of these findings for the different sections of the track. Generally, the users just gazed at the displayed content for short periods of time and then back into the scenery, but their use of the displays varied across the track.

First part, tipping the first pillar

In the initial phase, the user mainly got acquainted with the current display type. The users could also navigate to the first pillar quite easily, without any obstacles in their path. As noted in the data above, a user drove past the first pillar and was able to see this in the navigation display.

Past the first avatar and through the alley

After tipping the first pillar, many users glanced at the arrow and from there started to drive in the designated direction. They then drove to the narrowing alley of cones. In the beginning of this alley, they passed the first avatar, which was on the right side of the machine and as such in a blind spot behind the excavator's arm.

At the same time, the natural focus of the user was straight ahead, where he/she was going to go to a path between an alley of cones to get towards the next objective. As such, the user was mostly looking at the ground or straight ahead. It seemed that the head-up display and the projected display made it easier for the user to give the display a glance to see whether any warnings were present. Still, some users did not react to the avatar warning symbols, nor the warning of close obstacles. The users that reacted often rechecked the warnings, for example to see if the human warning was still active.

The sand pile.

Following the alley section of the track, the users approached the sand pile. Here the user was supposed to go up and down the path, a passage that also required a bit of machine rotation. This was a more difficult passage to navigate and the users

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

switched from being mostly reactive to the displayed content, to also actively searching for more visual support. The use between the head-up display and the projected display also differed at this stage. The head-up display was mainly used to check the direction of the arrow, while the projected display got the most glances for the warning and stop symbols. One possible reason for this is that the projected display was presented slightly lower and thus appeared to be on the ground, where the user was looking when driving the excavator.

The second avatar.
Coming to the end of the sand pile, the user had to pass another avatar, this time walking back and forth over the track, to reduce the effect of the user knowing where the avatar was and thus neglecting it. However, even though the avatar was moving, there are almost no readings on the head-down display and the projected display, as the attention was already drawn more upwards, towards the final pillars. The head-up display is practically the only display where the human warning is detected. Possibly due to the head-up display being placed a bit more straight ahead in the line of sight for the user, and thus more visible.

The last set of box piles.
The user was now approaching the end of the path. At this stage, the user was about to move into the designated place and also working with the bucket of the excavator at an extended height. At this stage, the usage seems to rely more on an individual operation style chosen by the user. Some users used the display support and did it for several display types, while other users simply focused at tipping the piles, without using the displays for visual support.

Operating the vehicle and performing the task was a demanding assignment. Not only did the user have to drive the excavator on the simulation track, the user also had to operate its boom to perform the tasks. This required a certain level of control and some users expressed this by deeply exhaling after finishing the course. Some users also commented that “it's not easy. Very tiring also!”. Still, everyone seemed to enjoy the experience, with users giving comments like “The experience was great. I was engaged in the test simulation as it is my role to drive the excavator”.

6 Discussion

The user observation and the quantitative data together show that the user benefited from the mixed reality displays while operating the vehicle, both in terms of lower effort and increased glances to the presented content. This result shows a coherence between heavy vehicles and related domains such as the automotive domain, where the use of head up displays shows better detection performance and faster reaction time [22, 39, 40]. It also aligns with other studies showing information more in line of sight can increase both the safety and experience of the vehicle operator[19], as well as lower mental load in the task operation [15].

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

Moreover, moving the information into the line of sight did not result in it being always detected. The general pattern of use was instead that the information intake was dependent on the need for support as well as the current workload. For example, the first part of the track was relatively simple to complete and during that part the users were more reactive to items popping up on both types of the mixed reality displays, while the head-down display messages were only detected by a few users. When they approached the more complicated part of the rack, i.e. the sand pile, they had to observe the surroundings around the machine more. At this stage, they also used the displays differently. Some users consciously monitored the display as they progressed. Other users checked it more in the process of their eyes scanning the surroundings and checking the display as their eyes naturally passed it. In this situation, they were more prone to miss key warnings. However, having the information presented within the line of sight aided for both taking glances while passing the display, as well as detecting and reacting to information.

The head-down display got some criticism for being out of sight for the users and that they were less motivated to use it. The users also reported higher physical demands (see Fig. 7), which align with simulated evaluations of operators' posture in excavators, that shows higher discomfort when having too look down rather looking at a head-up display [27]. One user also expressed that the projected display was preferred over the head-up display, for example, "for me the first one (the projected display) was more, what do we say, user friendly". This was followed by the comment that "I could recognize the things, objects near to me and how I should go. But in the next one (the head-down display) I was just confused. I couldn't recognize, I couldn't concentrate where should I look." One hypothesis for this is that the information is presented in between modalities in the head-up display, for example, not directly connected to the virtual environment and not on a specific surface. This might appear more confusing for the users than the projected display, where the information is more incorporated in the visual field. Some users also noted that the projected display looked like it was projected on the ground in front of the machine, a type of visual approach that has been investigated in the automotive domain for augmented navigation [20, 41].

Prior studies on the use of mixed reality show higher performance and reduced effort for processing the presented information [15, 42]. However, what is interesting is that our test subjects used the task-assistive information (the arrow that guides the boom) to a rather low degree. From early evaluations, we had noticed that the users had difficulties finding the correct depth between the bucket and the object. The arrow was thus designed to show the direction between the excavator's bucket and the top of the pillars. Even though we informed the users about this functionality, none of the users used the arrow for detailed bucket operation. Instead, the users were focusing on the task at hand and not paying attention to other things that were happening. This would be interesting to further evaluate, for example, if there is a difference between task information that is really needed by operator to perform the work and information that is more assistive.

In summary, we conclude that the results of the data analysis of this experiment point towards a recommendation of moving critical information into the operators'

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

field of view. This information does not seem to create additional effort to the user, instead, our test subjects preferred the head-up displays and it improved information intake under complicated situations. Given the benefits observed of the projection display used in this experiment, we would also recommend the use of displays that can provide an extended focal depth, as it shows more significant results in reduced effort and physical demand on the user. We would however recommend further evaluations to involve a larger sample size to confirm these indications.

7 Limitations and further work

This study has used a rapid prototyping environment and a simplistic user interface, which is meant to evaluate the potential for mixed reality displays in heavy vehicles. Further evaluation of its potential benefits would gain from a higher fidelity simulation and a larger number of users, with different diversity and practical real-life experience of operating excavators.

The symbolic language itself can also be refined and evaluated, including additional types of information. The use of more human senses, for example, audio cues together with visual presentation, could further enhance detection and intake of vital information.

8 Conclusion

Advancements in digital transformation for industrial vehicles enable improved productivity, comfort, and safety. A well-designed interaction is vital to fully exploit this, and to avoid information loss and ergonomic strain. Furthermore, the design needs to avoid unsafe situations resulting from operators' decreased situation awareness because of their attention may be drawn from their task, or missing supportive information from the vehicle system. We have investigated and proposed the use of transparent interfaces based on mixed reality as an approach to facilitate the display of key information to the operators. To improve detectability, our designs display information close to the operators' focus area of attention. We have evaluated two mixed reality visualization approaches: a head-up display and a display projected in the virtual environment, and compared these designs with a head-down display. The designs presented an arrow, used for navigational guidance, and warnings symbols, used to inform the user about obstacles near the machine. The users was both more reactive and more seeking out information on the mixed reality displays. The eye-tracking measurement shows that the users used the mixed reality displays to a significantly higher level, on average twice as often than the head-down display. Moreover, the NASA-RTLX results indicate that the users had lower workload using the mixed reality interfaces, where the mixed realty displays scored better in each category.

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

9 Acknowledgements

This research has received funding from CrossControl AB, the Swedish Knowledge Foundation (KK-stiftelsen) through the ITS-EASY program, and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 764951.

10 References

1. Sanchez, J., Duncan, J.R.: Operator-Automation Interaction in Agricultural Vehicles. Ergonomics in Design: The Quarterly of Human Factors Applications. 17, 14–19 (2009).
2. Kazan, E.E.: Analysis Of Fatal And Nonfatal Accidents Involving Earthmoving Equipment Operators And On-Foot Workers, https://digitalcommons.wayne.edu/oa_dissertations/731/, (2013).
3. Kecojevic, V., Radomsky, M.: The causes and control of loader- and truck-related fatalities in surface mining operations. International Journal of Injury Control and Safety Promotion. 11, 239–251 (2004).
4. Department for Transport: Incident at Llanbadarn Automatic Barrier Crossing (Locally Monitored), near Aberystwyth, 19 June 2011. (2012).
5. Wallmyr, M.: Seeing Through the Eyes of Heavy Vehicle Operators. In: Bernhaupt R., Dalvi G., Joshi A., K. Balkrishan D., O'Neill J., W.M. (ed.) Human-Computer Interaction - INTERACT 2017. pp. 534–543. Springer (2017).
6. Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society. 37, 32–64 (1995).
7. Patrick, J., Morgan, P.L.: Approaches to understanding, analysing and developing situation awareness. Theoretical Issues in Ergonomics Science. 11, 41–57 (2010).
8. Ruff, T.M., Holden, T.P.: Preventing collisions involving surface mining equipment: A GPS-based approach. Journal of Safety Research. 34, 175–181 (2003).
9. Jo, B.W., Lee, Y.S., Kim, J.H., Kim, D.K., Choi, P.H.: Proximity warning and excavator control system for prevention of collision accidents. Sustainability. 9, (2017).
10. Lagnel, O., Engstr, J.: Better work environment with Head Up Display, JTI report - Agriculture and Industry, no 440. (2015).
11. Doshi, A., Cheng, S.Y., Trivedi, M.M.: A Novel Active Heads-Up Display for Driver Assistance. IEEE Transactions on Systems, Man, and Cybernetics, Part B. 39, 85–93 (2009).
12. Schwarz, F., Fastenmeier, W.: Augmented reality warnings in vehicles: Effects of modality and specificity on effectiveness. Accident Analysis and

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

Prevention. 101, 55–66 (2017).

13. Webb, A.K., Vincent, E.C., Patnaik, P., Schwartz, J.L.: A Systems Approach for Augmented Reality Design. In: AC 2016: Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience. pp. 382–389. Springer, Toronto (2016).
14. Vasilijevic, A., Miskovic, N., Vukic, Z.: Comparative assessment of human machine interfaces for ROV guidance with different levels of secondary visual workload. In: 2013 21st Mediterranean Conference on Control and Automation, MED 2013 - Conference Proceedings. pp. 1292–1297. IEEE, Chania (2013).
15. Akyeampong, J., Udoka, S., Caruso, G., Bordegoni, M.: Evaluation of hydraulic excavator Human-Machine Interface concepts using NASA TLX. International Journal of Industrial Ergonomics. 44, 374–382 (2014).
16. Freeman, M.H.: Head-up displays—A Review. Optics Technology. 1, 63–70 (1969).
17. Krevelen, R. Van, Poelman, R.: A survey of augmented reality technologies, applications and limitations. The International Journal of Virtual Reality. Volume 9, 1–20 (2010).
18. Kun, A.L., Tscheligi, M., Riener, A., van der Meulen, H.: ARV 2017: Workshop on Augmented Reality for Intelligent Vehicles. In: 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI ’17) Adjunct. pp. 47–51 (2017).
19. Gabbard, J.L., Fitch, G.M., Kim, H.: Behind the glass: Driver challenges and opportunities for AR automotive applications. Proceedings of the IEEE. 102, 124–136 (2014).
20. Kim, K., Wohn, K.: Effects on Productivity and Safety of Map and Augmented Reality. IEICE TRANSACTIONS on Information and Systems. E94–D, 1051–1061 (2011).
21. AblaBmeier, M., Poitschke, T., Bengler, F.W.K., Rigoll, G.: Eye Gaze Studies Comparing Head-Up and Head-Down Displays in Vehicles. In: 2007 IEEE International Conference on Multimedia and Expo. pp. 2250–2252. IEEE, Beijing (2007).
22. Sojourner, R.J., Antin, J.F.: The effects of a simulated head-up display speedometer on perceptual task performance. Human factors. 32, 329–39 (1990).
23. Wallmyr, M.: Reflections on augmented reality for Heavy machinery-practical usage and challenges. In: Adjunct workshop on Augmented Reality for Intelligent Vehicles. at Automotive User Interfaces and Interactive Vehicular Applications Adjunct - AutomotiveUI ’17. pp. 47–51. ACM Press, New York, New York, USA (2017).
24. Akyeampong, J., Udoka, S.J., Park, E.H., Carolina, N.: A Hydraulic Excavator Augmented Reality Simulator for Operator Training. Conference on Industrial Engineering and Operations Management. 1511–1518 (2012).
25. Wang, X., Dunston, P.S.: Design, strategies, and issues towards an Augmented Reality-based construction training platform. Electronic Journal

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

of Information Technology in Construction. 12, 363–380 (2007).
26. Englund, M., Lundström, H., Bruneberg, T., Löfgren, B.: Evaluation of Head-up display showing bucking information in final felling, SkogForsk Report, no 869. (2015).
27. Akyeampong, J., Nevins, L., Udoka, S., Carolina, N.: Using Digital Human Modeling to Enhance Work Visibility for Excavator. In: Proceedings of the 2013 Industrial and Systems Engineering Research Conference. pp. 1909–1919. ProQuest (2013).
28. Kittusamy, N.K., Buchholz, B.: Whole-body vibration and postural stress among operators of construction equipment: A literature review. Journal of Safety Research. 35, 255–261 (2004).
29. Wallmyr, M., Kade, D., Holstein, T.: 360 Degree Mixed Reality Environment to Evaluate Interaction Design for Industrial Vehicles Including Head-Up and Head-Down Displays. In: Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry. VAMR 2018. pp. 377–391 (2018).
30. Unity, http://www.unity.com.
31. pupil labs, https://pupil-labs.com/.
32. Kassner, M., Patera, W., Bulling, A.: Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. 1151–1160 (2014).
33. NASA TLX Paper and Pencil Version. NASA Ames Research Center (2016).
34. Macefield, R.: How To Specify the Participant Group Size for Usability Studies: A Practitioner ’s Guide. Journal of Usability Studies. 5, 34–45 (2009).
35. Hart, S.G.: Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50, 904–908 (2006).
36. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Advances in Psychology. 52, 139–183 (1988).
37. Byers, J.C.: Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In: Advances in industrial ergonomics and safety. pp. 481–485 (1989).
38. Schreier, M.: Qualitative Content Analysis. In: Flick, U. (ed.) The SAGE Handbook of Qualitative Data Analysis. pp. 170–183. SAGE Publications Ltd (2013).
39. Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., Kamiya, H.: Effects of display position of a visual in-vehicle task on simulated driving. Applied Ergonomics. 37, 187–199 (2006).
40. Liu, Y.C., Wen, M.H.: Comparison of head-up display (HUD) vs. head-down display (HDD): Driving performance of commercial vehicle operators in Taiwan. International Journal of Human Computer Studies. 61, 679–697 (2004).

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**

41. Ng-Thow-Hing, V., Bark, K., Beckwith, L., Tran, C., Bhandari, R., Sridhar, S.: User-centered perspectives for automotive augmented reality. In: 2013 IEEE International Symposium on Mixed and Augmented Reality. pp. 13–22. IEEE (2013).
42. Yang, J.: Visual Support System for Remote- Control Construction Machine Based on Autonomous Cameras, (2015).

This document is the accepted version of the conference paper published at Springer. The final publication is available at Springer via **https://doi.org/10.1007/978-3-030-29381-9_44**