Lead me the right way?! The impact of position accuracy of augmented reality navigation arrows in a contact analogue head-up display on driving performance, workload, and usability

Lisa Pfannmüller\textsuperscript{a}, Matthias Walter\textsuperscript{a}, Klaus Bengler\textsuperscript{a}

\textsuperscript{a}Technische Universität München, Institute of Ergonomics, Boltzmannstr. 15, 85747 Garching, GERMANY

The contact analogue head-up display (cHUD) is a promising advancement of the conventional head-up display technology. Information can be presented in an augmented reality way, directly superimposed on the driving environment. In order to achieve increased driving safety and comfort, the fit between virtual information and the real world is critical. A discrepancy, e.g. due to imprecise sensor data, must not lead to confusion or distraction. In a static driving simulator study, the position accuracy of a contact analogue navigation arrow was manipulated (three levels) and the effect on navigation errors, lane-keeping performance, subjective workload and usability was investigated with and without a secondary audio-verbal n-back task \((n=2)\), as well as in easy and difficult navigation situations. Position accuracy had a significant impact on navigation errors. In particular, deviations up to 6 m from the ideal position led to significantly more navigation errors as compared to the ideal position. This effect was independent of the complexity of the navigation situation and of whether a secondary task had to be solved or not. Decreased position accuracy also significantly reduced usability ratings. These findings have important implications for the technical development of future automotive cHUD systems and related design concepts for navigation systems.

**Keywords:** contact analogue, head-up display, driving simulator, augmented reality, usability

1. **Introduction**

Modern automobiles are already equipped with a variety of driver-assistance systems (DAS) in order to relieve the driver and raise the quality of today’s mobility with respect to maximised safety and comfort (Bengler et al., 2014). Although fully-automated driving is still a vision of the future, the degree of assistance and automation in modern cars is increasing in order to fulfill the desire for safe, efficient and comfortable individual mobility. In light of this development, the driver’s main role has become a supervisory one: he/she has to monitor the systems’ operations and their appropriateness with respect to the traffic situation and may need to intervene in case of criticality. New display technologies are required to support the modern driver in his/her new role (Bengler, Götte, Pfannmüller, & Zaindl, 2015).

While traditional head-down displays (HDD) draw the driver’s attention away from the street, head-up displays (HUD) enable the presentation of information in the driver’s primary field of vision via a virtual image, which is mirrored in the windshield. This allows the driver to monitor information, e.g. of DAS, with minimum accommodation, adaptation and eye movement effort, hereby decreasing eyes-off-the-road times and speeding up reaction times to unexpected traffic events (Gengenbach, 1997; Gish & Staplin, 1995; Horrey, Wickens, & Alexander, 2003; Kiefer, 1998, 1999).

While the use of conformal symbology (Naish, 1964) in a HUD may already further enhance the above-mentioned advantages of this technology (Alexander, Wickens, & Hardy, 2005; Lauber, Bray, Harrison, Hemingway, & Scott, 1982) and reduce the likelihood of cognitive capture (Caird, Horrey, & Edwards, 2001; Tufano, 1997), a further advancement is the contact analogue head-up display (cHUD). Information in the cHUD can be presented in an augmented reality (AR) manner (Azuma, 1997), i.e. superimposed on the real driving environment at the very location where the information is needed (see Figure 1). This combination of real and virtual information is supposed to reduce the driver’s mental interpretation and the transformation effort necessary to transfer information provided by the system to the driving environment. It may also enhance situation and system awareness and understanding.

AR cueing, as a means of highlighting critical objects, regions or events in the real world, has already been shown to help direct the driver’s attention to hazards (Ho & Spence, 2005; Rusch et al., 2013; Schall et al., 2013), to improve target detection (Yeh & Wickens, 2001) and reduce collision involvement (Kramer,
Cassavaugh, Horrey, Becic, & Mayhugh, 2007; Lee, McGehee, Brown, & Reyes, 2002). However, AR information in the driver’s primary field of view risks interfering with the driver’s perception of, and reaction to, critical objects or events in the outside world because of masking, crowding or divided attention (Schall, Rusch, Lee, Vecera, & Rizzo, 2010).

Furthermore, although augmented reality applications have already been implemented in many areas, automotive cHUD technology still faces several challenges that have prevented its series application so far and may have a critical influence on its functionality, potential benefits and risks, and on acceptance.

Despite generally higher requirements concerning optical quality and display area, the position accuracy of the virtual elements is critical and may affect the perceived quality of augmentation. Whether the virtual information appears to be located correctly with respect to the outside environment is strongly affected by the quality and precision of the utilised sensor data, but also by the vehicle’s movements. Conventional GPS sensors are only accurate up to approximately 3 m, street maps and predictive road data – the basis for modern navigation systems – only reach maximum accuracies of 5 – 10 m (without DGPS). In a cHUD, inaccuracies like these may result in a deviation of the virtual information in the cHUD from its ideal position. Moreover, due to the optical characteristics of a cHUD, pitch and yaw movements of the car can lead to displacements of the virtual image in horizontal and vertical directions by several metres (Schneid, 2009). Despite this, the benefits of the cHUD might still outweigh.

Up to now, it has not been investigated whether, or to what extent, a deviation of contact analogue navigation information from the ideal location in the real environment has an impact on driving behaviour, perceived workload or usability. A discrepancy between virtual information in the cHUD and the real driving environment bears the risk of driver distraction: Visual attention may be attracted to the conflicting virtual information, demanding increased cognitive effort to interpret the discrepancy. In the worst case, the proposed benefits of the cHUD might be reversed, leading to increased workload and decrements in driving performance and usability. A potentially negative effect of the deviation of contact analogue information from the intended position could, however, occur in particularly complex, ambiguous traffic situations, which have been shown to increase workload (Hancock, Wulf, Thom, & Fassnacht, 1990; Harms, 1991; Noy, 1990), and/or during a secondary task. Even non-visual tasks, such as conversations, have been shown to interfere with driving (Strayer, Drews, & Johnston, 2003; Strayer & Johnston, 2001) and increase cognitive activity and workload (Conti, Dlugosch, Schwartz, & Bengler, 2013; Dlugosch, Conti, & Bengler, 2013).

In the current driving simulator study, we therefore investigated the impact of three different levels of position accuracy of a contact analogue navigation arrow on driving performance, perceived workload and usability. In addition, two degrees of situation complexity were implemented for the driving task, and an audio-verbal secondary task was used to induce additional cognitive workload to participants.

Figure 1. Visualisation of a contact analogue head–up display (cHUD) and a conventional HUD.
2. Method

2.1 Participants
Thirty-two participants took part in this study. All participants had held a valid driver’s licence for at least three years and reported having a minimum annual mileage of 10,000 km. All participants reported having normal or corrected-to-normal vision. One participant had to be excluded due to simulator sickness and another one due to problems with data recording. The resulting 30 participants (eight women, 22 men, \( M_{\text{age}} \) 26 years, range: 21 - 44 years) were included in the data analysis. Sixteen participants reported having experience with head-up displays.

2.2 Position accuracy, tasks and dependent measures
Three levels of position accuracy were defined for the contact analogue navigation arrow: (a) ideal position, (b) a deviation of \( \leq 3 \) m, and (c) a deviation of \( \leq 6 \) m in an \( x-/y\)-direction from the ideal position (see Figure 2 for an illustration). The exact positions within each level were determined randomly for each trial. The arrow was first visible 60 m before the manoeuvre point. In addition, the next navigation manoeuvre was indicated by a voice announcement.

![Figure 2. Left panel: Visualisation of the levels of position accuracy; upper-right panel: example of the ideal position of the navigation arrow; lower-right panel: example of the highest level of inaccuracy (\( \leq 6 \) m).](image1)

![Figure 3. Example of an easy (left) and a complex, difficult situation (right).](image2)
The route was composed of 10 different navigation situations; five easy, unambiguous situations and five complex, ambiguous situations. Examples of an easy and a difficult situation are shown in Figure 3. Intersections with only one possibility to turn left/right, i.e. to follow the navigation instructions, were categorised as easy. Complex intersections with several possibilities to turn left/right were categorised as difficult due to the fact that the navigation instructions could be potentially misleading here.

The secondary task consisted of an audio-verbal n-back task with n = 2 (Kirchner, 1958). Participants listened to a sequence of 10 numbers and had to repeat the penultimate number, i.e. the number two steps back.

Subjective workload was measured with the NASA rTLX (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1986). For each of the six subscales “Mental Demand”, “Physical Demand”, “Temporal Demand”, “Performance”, “Effort”, and “Frustration”, participants have to rate their perceived load on a scale from 0 to 10 in 0.5-steps. The overall workload index was obtained by averaging across the six subscales.

Usability was measured with the Systems Usability Scale (SUS) (Brooke, 1996). Participants had to rate each of the 10 items on a 5-point Likert scale with the endpoints fully disagree and fully agree, resulting in an overall score between 0 and 100.

Measures of driving or lane-keeping performance were navigation errors (defined as deviations of more than 6 m from the ideal line), the steering reversal rate (SRR, 5° gap size), which is defined as the proportion of the number of counted steering wheel reversals greater than the defined gap size, and the time needed (McLean & Hoffmann, 1975). The SRR was chosen because of its relatively low dependency on the road curvature compared to other lane-keeping measures (Knappe, Keinath, Bengler, & Meinecke, 2007). Data was analysed from approximately 200 m before the manoeuvre point until 200 m thereafter.

2.3 Procedure

The study took place in the static driving simulator at the Institute of Ergonomics, which consists of a fully-instrumented BMW 640i mock-up, surrounded by a 270° view on three front screens (resolution 1400 x 1050 pixel each) and three rear screens (resolution 1600 x 1200 pixel for the side mirrors and 1280 x 1024 pixel for the rear-view mirror). The SILAB simulation environment was used. The contact analog navigation arrow was directly integrated into the simulation environment (see Figure 2).

After participants had given written informed consent, filled out a demographic questionnaire and had familiarised themselves with the driving simulator and the secondary task, they drove the above described route of 10 navigation situations (five easy, five difficult/complex) six times, i.e. at every accuracy level, with and without the secondary task. The order of conditions was randomised, and the sequence of situations and the look of the scenery along the routes varied to minimise learning effects. A baseline of the secondary task was recorded either at the beginning, in the middle, or at the end of the experiment (randomised across participants). The secondary task started after the voice announcement of the forthcoming manoeuvre. Participants filled out the above-described questionnaires after each route, i.e. six times.

3. Results

To analyse the impact of position accuracy, secondary task, and situation complexity on driving behaviour, 3 x 2 x 2 (Position Accuracy [ideal, 3 m, 6 m] x Secondary Task [yes, no] x Situation Complexity [easy, difficult]) repeated measures ANOVAs were calculated for navigation errors and SRR separately. To analyse the impact of position accuracy and secondary task on perceived workload and usability, 3 x 2 (Position Accuracy x Secondary Task) repeated measures ANOVAs were calculated for the NASA rTLX overall score and the SUS score. A 3 x 2 (Position Accuracy x Situation Complexity) repeated measures ANOVA was calculated for secondary task performance (error rate).

Position accuracy had a significant effect on navigation errors, F(2, 58) = 4.67, p = .019 (Greenhouse-Geisser-corrected). Error rates increased with decreasing position accuracy (M_ideal = 0.15, SD_ideal = 0.23; M_3m = 0.17, SD_3m = 0.24; M_6m = 0.19, SD_6m = 0.27). Post-hoc t-tests (Bonferroni-corrected) showed that significantly more errors were made at the highest level of inaccuracy compared to the ideal position (p < .05, see Figure 4). Participants made significantly more navigation errors in easy than in difficult situations, F(1, 29) = 61.11, p < .001 (M_easy = 0.01, SD_easy = 0.08; M_difficult = 0.34, SD_difficult = 0.25). The n-back task had no significant effect on navigation errors, F(1, 29) = 0.83, p = .370.
Position accuracy did not significantly affect the SRR, $F(2, 58) = 0.07, p = .937$, but situation complexity, $F(1, 29) = 113.82, p < .001$, and n-back task, $F(1, 29) = 13.07, p = .001$. The SRR was higher in easy compared to complex situations ($M_{\text{easy}} = 0.62, SD_{\text{easy}} = 0.11$; $M_{\text{difficult}} = 0.53, SD_{\text{difficult}} = 0.11$) and higher when the n-back task had to be performed compared to when not ($M_{\text{without}} = 0.56, SD_{\text{without}} = 0.14; M_{\text{with}} = 0.59, SD_{\text{with}} = 0.12$).

Position accuracy did not have a significant effect on overall workload measured by the NASA rTLX, $F(2, 58) = 1.21, p = .305$, but perceived workload significantly increased with the additional n-back task, $F(1, 29) = 86.39, p < .001$ ($M_{\text{without}} = 2.65, SD_{\text{without}} = 1.20; M_{\text{with}} = 5.03, SD_{\text{with}} = 1.53$).

Position accuracy had a significant effect on usability (SUS score), $F(2, 58) = 13.25, p < .001$. A decrease in accuracy led to a reduction in usability scores ($M_{0m} = 72.67, SD_{0m} = 8.28; M_{3m} = 70.27, SD_{3m} = 10.92; M_{6m} = 65.27, SD_{6m} = 9.06$). Post-hoc $t$-tests (Bonferroni-corrected) revealed significant differences in SUS scores between 0 m and 6 m, and between 3 m and 6 m (both $p < .001$, see Figure 4). The n-back task did not significantly affect usability ratings, $F(1, 29) = 0.81, p = .376$.

Performance in the n-back task was not significantly affected by position accuracy, $F(2, 58) = 0.80, p = .450$, but by situation complexity, $F(1, 29) = 5.97, p = .002$, interestingly with higher error rates in easy than in difficult situations ($M_{\text{easy}} = 0.15, SD_{\text{easy}} = 0.13; M_{\text{difficult}} = 0.12, SD_{\text{difficult}} = 0.12$).

4. Discussion and conclusion

The results suggest that position inaccuracies of up to 6 m in x- and y-direction have a significantly negative impact on navigation performance in a navigation task. This effect was observed in the increased number of navigation errors, but not in the SRR as a measure of lane-keeping performance. This may be attributed to the generally high variability in the driving data and therefore comparably low power of this driving performance metric. There have also been diverging findings concerning the change in SRR depending on gap size and primary and secondary task demands, especially regarding visual vs. audio-cognitive load (Engström, Johansson, & Östlund, 2005; Jamson & Merat, 2005; Macdonald & Hoffmann, 1980).

The significantly negative impact of greater inaccuracies on navigation errors was independent of the situation complexity and of whether a secondary task was performed while driving or not, i.e. no significant interaction of the accuracy level with one of the additional factors was observed. Nevertheless, the percentage of navigation errors in complex situations was, in general, significantly higher than in easy situations; in fact, almost no errors were made in easy situations, suggesting that the significant main effect of position accuracy on navigation errors mainly arose from difficult traffic situations. The lack of a significant interaction effect may again be a result of low power. Future research will follow up on this and will only concentrate on difficult situations.
The manipulation of position accuracy had no significant impact on perceived workload, as measured by the NASA rTLX overall score and by secondary task performance. It is possible that both measures were not sensitive enough to detect differences due to variations in position accuracy. These findings require further investigation and should be validated with another workload measure, such as the detection response task (DRT).

Finally, a significant impact of position accuracy on SUS usability scores was revealed. Especially the drop from 73 points at the ideal position, to 65 points at the level of highest inaccuracy (≤ 6 m deviation) represents a decline in usability from very acceptable to only marginally acceptable (Brooke, 1996). This finding undermines the importance of the accurate positioning of information in an automotive cHUD with respect to series application and customer demands. However, there will be follow-ups on the current findings in respect of the effect of position accuracy for other use cases or DAS than navigation, such as ACC (adaptive cruise control), and with respect to different display concepts for contact analogue information display.

The current findings highlight the importance of the accurate positioning of virtual information in an automotive cHUD. The results suggest a negative impact of position inaccuracies of AR information in a cHUD on driving performance and usability ratings, underlining the need for further technological advances with respect to series production. However, as some technological restrictions will always remain, research should focus on developing guidelines for robust, error-tolerant, safe and ergonomic concepts for information display in the cHUD within the technological boundaries.

References


