

Augmented Reality Displays for Communicating Uncertainty Information in Automated Driving

Alexander Kunze
Loughborough Design School
Loughborough University
Loughborough, UK
A.Kunze@lboro.ac.uk

Stephen J. Summerskill
Loughborough Design School
Loughborough University
Loughborough, UK
S.J.Summerskill2@lboro.ac.uk

Russell Marshall
Loughborough Design School
Loughborough University
Loughborough, UK
R.Marshall@lboro.ac.uk

Ashleigh J. Filtness
Loughborough Design School
Loughborough University
Loughborough, UK
A.J.Filtness@lboro.ac.uk

ABSTRACT

Safe manual driving performance following takeovers in conditionally automated driving systems is impeded by a lack in situation awareness, partly due to an inappropriate trust in the system's capabilities. Previous work has indicated that the communication of system uncertainties can aid the trust calibration process. However, it has yet to be investigated how the information is best conveyed to the human operator. The study outlined in this publication presents an interface layout to visualise function-specific uncertainty information in an augmented reality display and explores the suitability of 11 visual variables. 46 participants completed a sorting task and indicated their preference for each of these variables. The results demonstrate that particularly colour-based and animation-based variables, above all *hue*, convey a clear order in terms of urgency and are well-received by participants. The presented findings have implications for all augmented reality displays that are intended to show content varying in urgency.

Author Keywords

Augmented reality; automated driving; reliability; trust; uncertainty; visual variables.

CCS Concepts

•**Human-centered computing** → **Mixed / augmented reality; User interface design; HCI theory, concepts and models; Interaction techniques**; Visualization design and evaluation methods;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AutomotiveUI '18, September 23–25, 2018, Toronto, ON, Canada

© 2018 ACM. ISBN 978-1-4503-5946-7/18/09...\$15.00

DOI: <https://doi.org/10.1145/3239060.3239074>

INTRODUCTION

Vehicles equipped with automated driving systems (ADSs) promise advancements in road safety and efficiency while simultaneously invoking a significant shift in the driving experience of human operators [24]. In current passenger vehicles, the primary task of human operators is the complete execution of the dynamic driving task (DDT). In future vehicles, however, the significance of this task might shift and other, now secondary or tertiary tasks, may become more prominent and replace the driving task as foremost activity [25]. For the foreseeable future, human operators are expected to be fallback-ready and perform the DDT in cases of system failures or other critical situations within a reasonable amount of time [8, 42]. Previous research has indicated that it is beneficial to communicate the uncertainties of a system to prepare users for such takeovers despite the engagement in non-driving related tasks (NDRTs) [1, 16]. The existing proposals, however, do not explore various display formats but solely rely on visual information presented in the instrument cluster. This requires users to constantly shift their attention between road, instrument cluster, and NDRT, leading to an increased probability of missed critical events. Further, optimal ways of presenting information about the system's uncertainty regarding specific functions, e. g. lateral control, have yet to be explored. Addressing these shortcomings, this publication presents an initial evaluation of variables for an augmented reality (AR) based visual uncertainty display.

RELATED WORK

Automating systems under the expectation that a fallback-ready user can resume the previously automated task in case of system failures entails several human factors challenges, foremostly a lack in situation awareness at the time of a takeover [7]. To support users with the acquisition and maintenance of situation awareness, their attention management can be influenced through trust calibration [17, 39]. Lee and See [27] identified three essential aspects for achieving appropriate trust: *calibration*, *resolution*, and *specificity*. *Calibration*

refers to the agreement between the user’s trust in automation and the capabilities of the automation. *Resolution* is the degree to which the user’s trust judgement differentiates different capability levels. *Specificity* refers to the degree of differentiation between different components or aspects of the trustee and includes functional and temporal specificity. The former relates to the differentiation between different functions while the latter describes the sensitivity to changes. A combination of calibrated trust as well as a high resolution and specificity can potentially alleviate both disuse and misuse of automated systems. Systems can support the user with the acquisition of appropriate trust by providing information along the dimensions *abstraction* and *detail*. *Abstraction* refers to information regarding the performance, process, and purpose of the automation [26], with the former being the most crucial [18]. *Detail* describes the entity that is to be trusted, ranging from the overall system to its single functions and modes. System-wide trust theory suggests that users merge their trust across multiple aids, independent of their individual reliability levels [23, 11]. This is problematic as the user’s distrust of an unreliable function might affect other highly reliable functions and might lead to the unreasonable disuse of these. No publications are known to the authors of this publication that investigate the impact of explicitly communicating the reliability levels of different vehicle system components or functions on operator behaviour and if communicating reliability can ameliorate the contagious effects, justifying the need for developing a function-specific display to explore this.

Benefits of Uncertainty Communication

Recent research has investigated the impact of communicating the overall uncertainties of automated systems in the automotive domain. Beller et al. [1] explored the communication of uncertainties by displaying a schematised uncertain face in the instrument cluster in critical situations. The results indicate that communicating uncertainties improves both driving safety following takeovers as well as operators’ monitoring behaviour. As such, minimum time-to-collision figures were significantly lower for participants that were presented with the uncertainty information compared to those who did not have this information available. Further, secondary task performance indicated that participants of the experimental group directed their attention more to the field relevant for driving (FRD) in critical situations than the control group, potentially explaining the difference in takeover performance. Supplementing this work, Helldin et al. [16] presented uncertainty information with a higher resolution in the form of seven bars in the instrument cluster, with each bar representing one level. Participants to which this information was available were able to take over the DDT faster and were also able to perform NDRTs for a longer time than the control group. Corresponding with these findings, the benefits of presenting uncertainty are reaffirmed by publications in other domains such as aviation and military [6, 10, 33, 47].

The previous work is limited in that the presented solutions require users to shift their focus towards the instrument cluster in order to gain knowledge about the system’s current uncertainty. While this has already benefited the practicability of NDRTs [16], solutions that present the information without

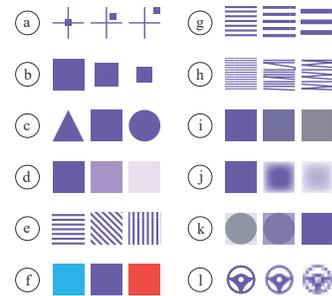


Figure 1: Extended set of visual variables, increasing uncertainty from left to right: (a) position, (b) size, (c) shape, (d) value, (e) orientation, (f) hue, (g) grain, (h) arrangement, (i) saturation, (j) crispness, (k) transparency, (l) resolution [31]

requiring the user to glance towards the instrument cluster might further improve this. Additionally, none of the existing solutions investigate the communication of uncertainties with a higher functional specificity.

Methods of Communicating Uncertainty

Besides showing the general benefits of communicating uncertainty, the best suited methods for conveying uncertainties need to be investigated. Noah et al. [36] explored the use of human-centric and system-centric reliability displays in the context of automated lane keeping and compared different display types, namely qualitative, quantitative, and representational displays. Their findings indicate a similar matching accuracy of the different display types and suggest that participants can more readily match system-centric displays with intended reliability levels. To facilitate the practicability of NDRTs in automated driving, Kunze et al. [24] proposed the unobtrusive communication of uncertainties using a combination of haptic seat feedback and peripheral lighting. In a military context, Neyedli, Hollands and Jamieson [35] investigated visual displays varying in display type (random mesh, pie) and proximity (separated, integrated) for showing target identity and corresponding system reliability information. Their findings indicate that integrating the reliability information leads to more appropriate reliance. Additionally, it was shown that the graphical display of uncertainty is equally effective as verbal or numerical communication [10, 47].

Graphically integrating uncertainty information with the affected data is one of the major challenges within Geographic Information Science (GIScience) [30, 38]. The use of abstract, manipulable signs, also referred to as *visual variables*, is a frequently employed strategy within GIScience. Initially proposed by Bertin [3], the basic set of visual variables has subsequently been extended, as of now including the following (see Figure 1) [12, 29, 34, 28]:

- (a) Position: alteration in the x, y, z location
- (b) Size: alteration in area, length, or repetition
- (c) Shape: changes in form
- (d) Value: alteration in the relative lightness or darkness
- (e) Orientation: alteration in the alignment
- (f) Hue: alteration in the spectrum colours

- (g) Grain: changes in granularity
- (h) Arrangement: changes in regularity
- (i) Saturation: alteration in the spectral peakedness
- (j) Crispness: alteration in the boundary sharpness
- (k) Transparency: changes in opacity
- (l) Resolution: alteration in the spatial precision
- (m) Movement: animated changes in position
- (n) Frequency: animated changes in animation speed

With the exception of *shape* and *arrangement*, all of the variables are at least marginally suitable for communicating a logical order [41]. MacEachren [31] investigated the intuitiveness of the presented visual variables for communicating uncertainty in a geographical context. The results indicate that particularly position (a) and crispness (j) provide a logical order for communicating uncertainty, followed by value (d), size (b), and transparency (k). The automotive domain can build on these findings and further investigate the use of these variables in a driving context, whereby the different connotation must be considered. While uncertainty visualisation in geographical maps refers to the reliability of the underlying data, the uncertainty information of an automated system refers to its inherent reliability and infers the likelihood of a takeover. As such, an increasing automation uncertainty implies a higher urgency for human operators to focus their attention on the current driving situation, whereas uncertain areas in maps may be visualised with the opposite intention.

Display Formats, Visual Processing and Attention

It further needs to be addressed how the methods should be applied, namely using which display type. Depending upon the degree of abstraction, visual displays can be categorised as digital, analogue, representational, or contact analogue [5]. Higher levels of abstraction are thereby thought to increase the cognitive effort required for linking the presented information to the real world, favouring contact analogue displays to keep the mental workload low [5, 21]. While all of the mentioned display formats can be used in a head-up display (HUD) that projects information directly into the FRD, contact analogue HUDs (caHUDs) superimpose information in positional and temporal relativity to the environment, i. e. augmented reality (AR) information, while the other displays are conveying content detached from the surroundings. As a consequence, drivers interacting with digital, analogue, and representational displays on a HUD are not focusing their attention on the road, but rather at the HUD content, leading to higher distraction levels and worse task performance despite a subjectively higher rating compared to head-down displays (HDDs) [43]. This favours the use of caHUDs in the automotive domain, as evidenced by various applications, including traffic augmentation to increase trust in automated driving systems [49], systems to improve object detection during nighttime [40, 19], or providing visual information about the potential braking path of a vehicle [44].

To facilitate both the quality and quantity of displayed information in visual displays, knowledge about perception is essential [48]. Phenomena such as postattentive blindness, change blindness, postattentive amnesia and memory-guided search highlight that visual processing is critically influenced

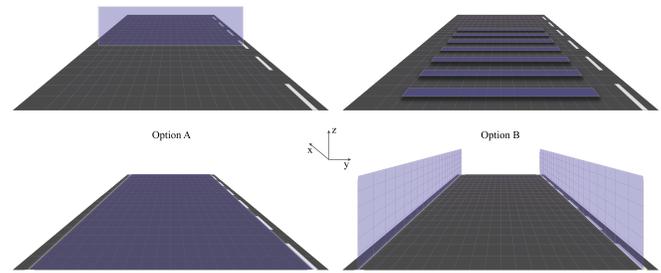


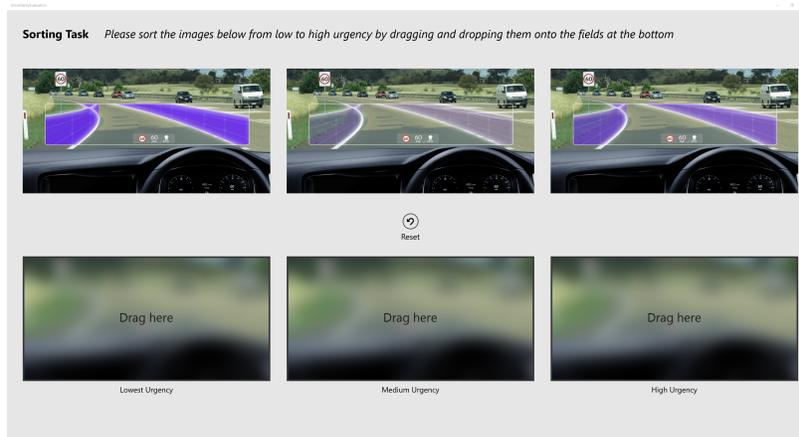
Figure 2: Basic layout options (purple) for presenting longitudinal (top) and lateral uncertainties (bottom) as overlays on AR displays

by the viewer's state of mind and attentional focus [13]. The term visual attention is thereby used to delineate the processes used to select areas for analysis. Repeated cycles of stationary eye periods, fixations, and rapid eye movements to new locations, saccades, enable humans to see detailed information of larger regions despite the limitation of only seeing detailed colour and shapes in a small part of the visual field at any given moment [13]. A limited set of visual features, however, can be detected with a single glance via low-level visual processes, often referred to as *preattentive* properties. Among these preattentive visual features are size, orientation, hue, density, luminance, and motion, showing some overlap with the previously presented visual variables [22, 45, 46, 14, 20]. Preattentive processing is thereby hindered by the joint presence of more than one visual feature, unless this is done in a redundant manner. Further, the preattentive properties of visual features are not equal but rather vary in their *popout* capabilities, with colour, size, orientation, contrast, and animation having the strongest effects [48].

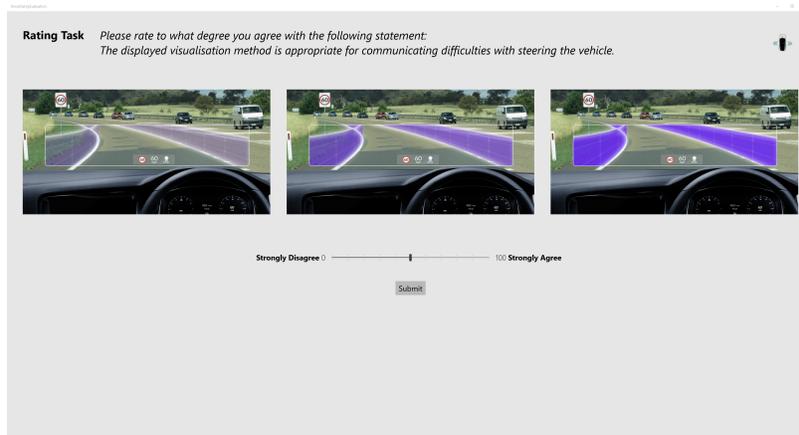
DEVELOPMENT OF AN AR UNCERTAINTY DISPLAY

The basic functions that an ADS has to execute can be summarised as the DDT, consisting of tactical functions such as the planning of manoeuvres as well as trajectories and operational functions, e. g. pre-cognitive, natural reactions. Similar to currently available automation systems (lane assist, adaptive cruise control (ACC)), the DDT can be split into a lateral and longitudinal component [42]. These discriminable functions can be used as a starting point for investigating the communication of function-specific uncertainties.

Following the findings presented in the previous sections, this paper explores the communication of uncertainties regarding the lateral and longitudinal vehicle motion using visual variables in an AR display. Building on research conducted within GIScience, the uncertainty information can be overlaid on the FRD. To enable overlays, areas relative to the vehicle movement must be defined, leading to three possible planes: x-y, y-z, and x-z. Based on metaphors (e. g. crash barriers for lateral uncertainty) and learned representations (e. g. ACC for longitudinal uncertainty), the options shown in Figure 2 were generated. As the uncertainties are intended to be displayed simultaneously, the different layout options must be compatible. A combination of lateral option B and longitudi-



(a)



(b)

Figure 3: Interface layout for sorting (a) and rating session (b)

nal option B was chosen as this leads to clear visualisations without interference, e. g. in the form of overlaps.

STUDY OBJECTIVES AND METHOD

The primary purpose of this study is to evaluate a large set of basic visual variables regarding their suitability for communicating lateral and longitudinal uncertainties. The suitability is thereby determined by investigating the intuitiveness of ordering the variables and subjective preference ratings. The results of this study enable the informed development of a prototype interface for subsequent behavioural studies in a driving simulator.

Participants

A total of 46 participants (16 female) with an average age of 30.24 years ($SD = 9.40$) participated in the experiment. All participants held a valid driving license for an average of 10.11 years ($SD = 9.14$) and averaged an annual mileage of 10,338.04 ($SD = 28,999.64$). Five participants indicated to have used vehicles equipped with ACC or lane keeping assistants. Participants received no monetary compensation for taking part in the study.

Uncertainty Visualisation Method

The visual variables were selected based on their anticipated ordinal characteristics [41] as well as their practicability in the chosen layout. As such, variables a to n with the exception of *arrangement*, *shape*, and *resolution* were included in the experiment. Following a 3-step scale employed by MacEachren [31] and using the previously presented visual layout options, 11 visual variables were used to convey low, medium, and high reliability for each vehicle function, resulting in a total of 66 variations. These variations were placed into a generic driving scene to generate context (see Figures 3, 4, and 5). The layout option selected for lateral uncertainties is characterised by a large overlay area and therefore affords the direct transfer of the variables without further modification. In contrast, more subtle variables such as changes in line *orientation* or *grain* were not visible when using the chosen layout for longitudinal uncertainties. Instead of lines, the *orientation* and *grain* of the bars themselves was changed (see Figure 4, *Orientation* and *Grain*). Further, the variable *position* was varied on different axes. Regarding lateral uncertainty, the position was varied to indicate a wider or narrower lane (changes in y-axis) while the bars for longitudinal uncertainty were varied in their height (z-

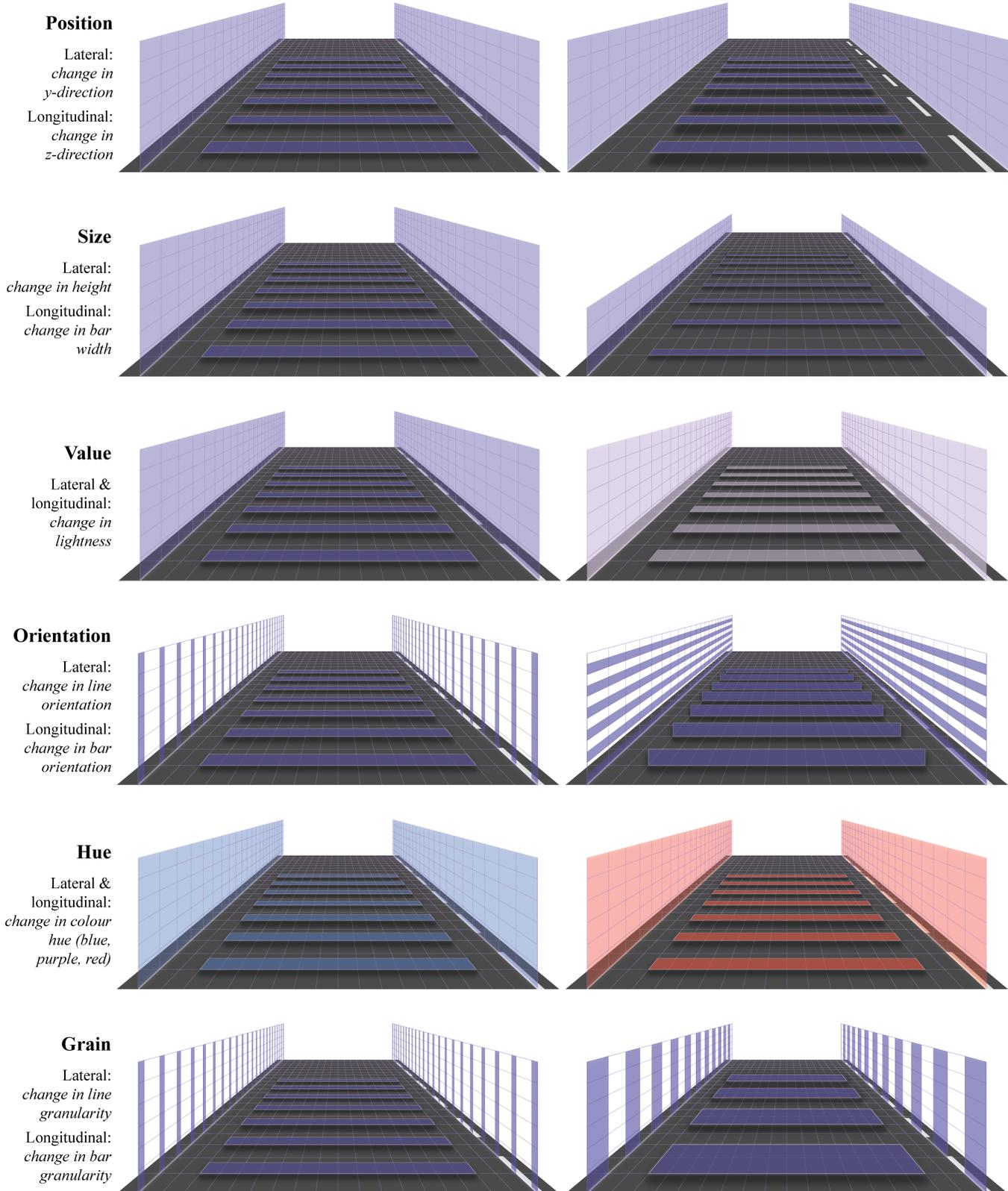
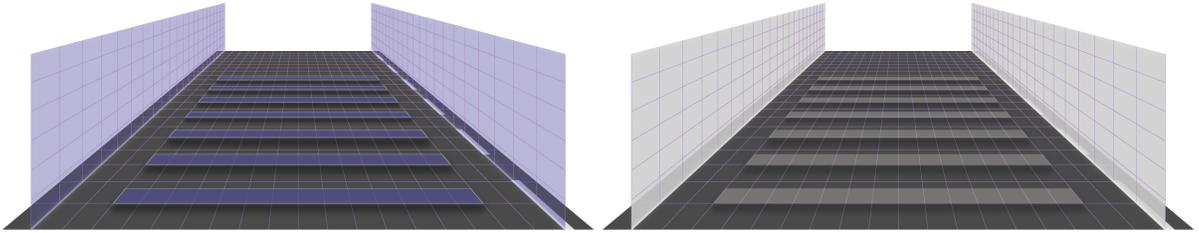


Figure 4: Implementation of visual variables in the chosen layout options, ordered according to condition 1 with increasing urgency from left to right for all static variables; intermediate steps were omitted to improve visibility in print (see Figure 1)

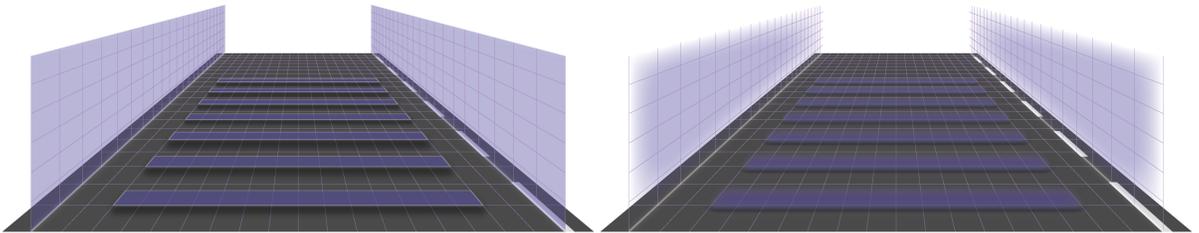
Saturation

Lateral & longitudinal:
change in spectral peakedness



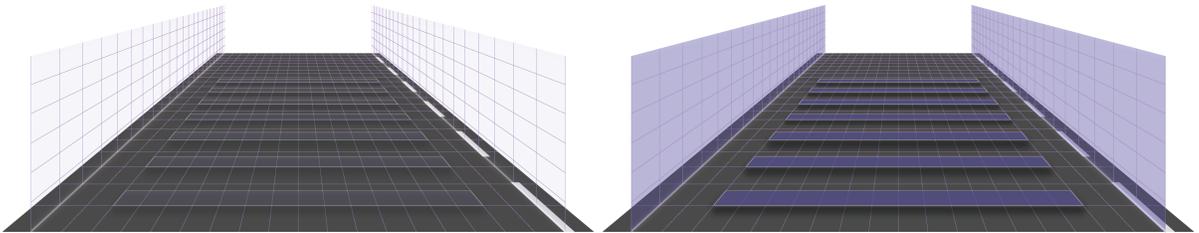
Crispness

Lateral & longitudinal:
change in boundary sharpness



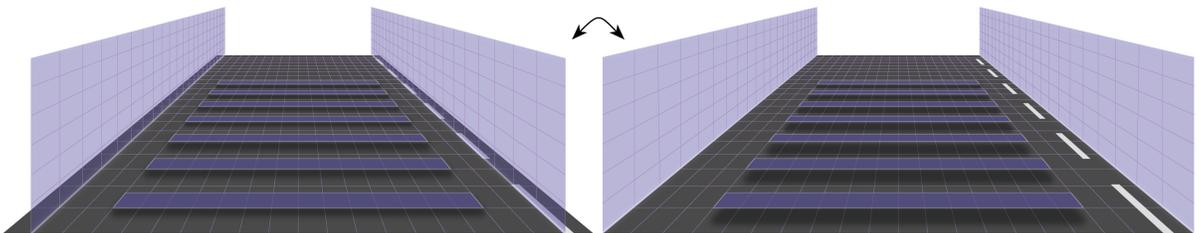
Transparency

Lateral & longitudinal:
change in opacity



Movement

Lateral & longitudinal:
animated position



Frequency

Lateral & longitudinal:
animated opacity

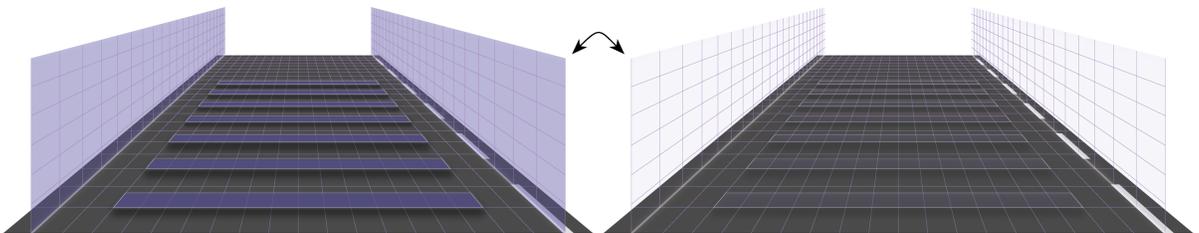


Figure 5: (Continued) Implementation of visual variables in the chosen layout options, ordered according to condition 1 with increasing urgency from left to right for all static variables; intermediate steps were omitted to improve visibility in print (see Figure 1)

Visual variables	Lateral				Longitudinal			
	Cond.	Mean	Mode	SD	Cond.	Mean	Mode	SD
Position	1	0.5652	1.00	0.4197	1	0.8261	1.00	0.3157
Size	2	0.7536	1.00	0.4170	2	0.9130	1.00	0.2506
Value	2	0.8841	1.00	0.3115	2	0.9420	1.00	0.2168
Orientation	1	0.8116	1.00	0.2812	1	0.8841	1.00	0.2380
Hue	1	0.9565	1.00	0.1148	1	0.9855	1.00	0.0695
Grain	1	0.7536	1.00	0.4047	1	0.5217	1.00	0.4906
Saturation	2	0.9565	1.00	0.2085	2	0.9275	1.00	0.2453
Crispness	2	0.7681	1.00	0.4197	2	0.5942	1.00	0.4259
Transparency	1	0.9130	1.00	0.2881	1	0.8696	1.00	0.3444
Movement	1	0.9130	1.00	0.2506	1	0.7101	1.00	0.3379
Frequency	1	0.8986	1.00	0.2740	1	0.8986	1.00	0.2343

Table 1: Sorting scores for each variable and driving function for the higher scoring condition; *Cond.*: sorting condition with higher mean score, condition 1 refers to the order shown in Figures 4 and 5, condition 2 means that the order shown in the figures is reversed; *Mean*, *Mode*, *SD*: descriptive statistics for each variable in the indicated sorting condition

axis) to create the impression of barriers. Increasing urgency levels (condition 1) for animation-based variables, *movement* and *frequency*, were communicated with a higher speed or flashing rate, respectively. *Movement* was implemented by animating the position of the overlays analogous to the *position* variable. A pilot study with 10 participants was conducted to ensure that all levels for each variable can be confidently distinguished.

Design

A 2 (driving functions, between) \times 11 (visual variables, within) mixed design was used to investigate the suitability of the visual variables for each driving function. The order of the levels within both groups was randomised and balanced.

Apparatus

The participants completed all tasks in a soundproof laboratory with controlled lighting conditions using a purpose-built Universal Windows Platform application on a 27" display with a resolution of 3840 \times 2160. Introductory information as well as all required instructions were embedded in the application, with experimenters and chaperones being present throughout all sessions for potential questions and comfort.

Procedure and Method

The experiment was split into two consecutive sessions, each lasting approximately 15 minutes per participant. Each session was introduced with a video summarising the purpose of the session and explaining the instructions for the upcoming task. The instructions ended with a sample question and participants were able to ask questions before starting with each task series. After each task, the participants were presented with an info screen that showed how many tasks were remaining and allowed a short mental break. Further, participants were required to click a button to proceed to the next screen in order to ensure that the mouse position was centred for each task.

The first session was intended to evaluate to what degree the different visual variables afford a consistent order. For this purpose, participants were presented with 11 consecutive forced

choice sorting tasks (one for each variable, see Figure 3a), with each task separated by the info screen. Participants had to sort three different levels of each variable according to the perceived urgency. The term *urgency* was chosen because pilot studies revealed that participants struggled with the concepts of system uncertainty and reliability, prompting a change to a more familiar term that infers a higher takeover probability. In addition to the sorting order, the time taken (*response time*) for each ordering task was recorded as an additional measure for intuitiveness [31], starting with the click of the proceed button on the info screen and ending with the drop of the last image on the task screen. The key objective of the sorting task was to investigate which variables afford a higher sorting consistency. A sorting measure for ordinal scales was developed to gain a single parameter that characterises the sorting consistency. The measure is based on the following fundamental conditions:

- Condition 1: Extreme A < Intermediate < Extreme B
- Condition 2: Extreme B < Intermediate < Extreme A

Exclusively visual variables with ordinal characteristics [41] have been selected, therefore a general order between two extremes can be assumed, with the intermediate always being between these two extremes. Each of these conditions have three quantifiable requirements. To achieve a score of 100 % for condition 1, for instance, Extreme A has to be placed left of the Intermediate and both Extreme A and the Intermediate have to be placed left of Extreme B (see Figure 3a). The fulfilment of zero requirements leads to a score of 0 % while the fulfilment of one or two requirements leads to 33.3 % and 66.7 % respectively. Swapping the Extremes A and B leads to the requirements of condition 2. Thus, the sum of the scores for each condition equates to 1.00. For each participant and visual variable, a score for conditions 1 and 2 was recorded, with the higher score indicating the potentially more logical sorting direction. The mean scores across participants of each group then indicate the sorting consistency and sorting direction. Thereby, condition 1 refers to the order indicated by previous publications, as shown in Figures 1, 4, and 5.

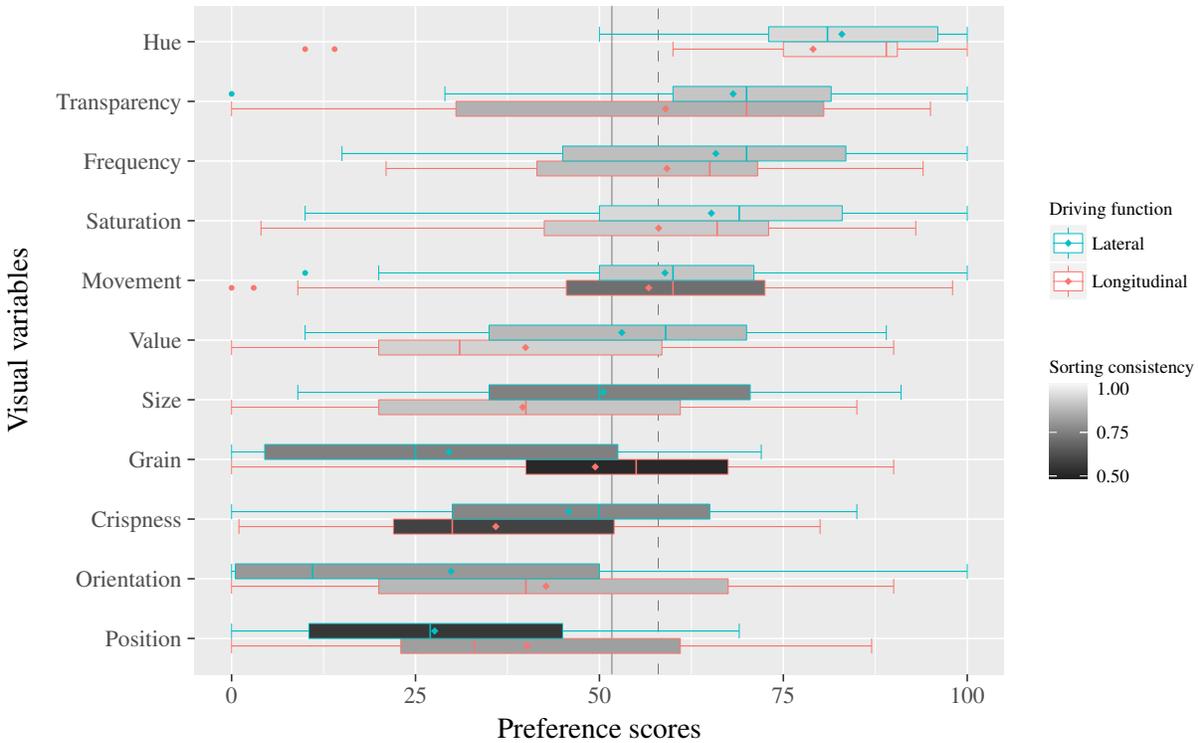


Figure 6: Preference scores for each visual variable relative to the driving function. The mean for each visual variable is indicated by a diamond within each bar. The dashed line indicates the median, the full line the mean of the overall preference scores.

The main purpose of the second session was to explore the suitability of the variables for communicating uncertainties. The variables may communicate a clear order, but could be unsuited for this particular application. Figure 3b shows the screen layout for the rating task, with the three levels of each variable in the top half of the screen, ordered as previously by the participant. The terms *lateral* and *longitudinal* were substituted with *steering the vehicle* and *braking and accelerating* to prevent misunderstandings. Further, a representational illustration was displayed in the upper right corner to provide additional clarification. Participants were asked to indicate their agreement with a statement shown above the images using a slider-scale ranging from *strongly disagree* (0) to *strongly agree* (100). A slider-scale was used instead of a Likert-scale to improve interactivity and allow for more granular responses. This was preferred over a pairwise comparison as that would have required $\frac{N \cdot (N-1)}{2} = 55$ questions instead of 11, meaning a higher workload for participants.

RESULTS

The responses of participants were collected and processed by the purpose-built application and subsequently analysed in R.

Sorting Scores

Table 1 summarises the results. The mode of all variables was either 1.00 or 0 depending on the sorting condition, indicating a dominant order with a clear intermediate. The variable *hue* achieved the highest sorting mean score in both the lateral

and longitudinal groups, with 0.9565 and 0.9855 respectively. Further, it was also characterised by the lowest standard deviation, suggesting a high sorting consistency among participants irrespective of the driving function. The variables *saturation*, *transparency*, *frequency*, and *value* were also characterised by high sorting mean scores (> 0.8). The sorting direction was consistent between groups, as indicated by the same maximum conditions for each variable. Compared with previous research, however, the direction of the variables *size*, *value*, *saturation*, and *crispness* was reversed.

Response Times

A 2×11 mixed-design ANOVA was conducted to assess the effects of visual variables and driving functions on response time. No main effect was returned either within or between driving functions, suggesting that participants took similarly long to complete the sorting task for each variable between and within groups. However, a significant interaction between driving function and variables was found [$F(1, 10) = 1.989, p = .0328$]. This interaction shows that between groups, the response times for the visual variables differed.

Preference Scores

Bartlett's test returned non-significant differences in the variances between and within groups for preference scores. A Shapiro-Wilk test and q-q plot indicated a non-normal distribution, prompting the use of nonparametric statistics. Thus, a

	Pos.	Size	Value	Orient.	Hue	Grain	Saturat.	Crispness	Transpar.	Movem.	Freq.
Position	●	-	-	-	< 0.001	-	-	-	-	-	-
Size	0.044	●	-	-	< 0.001	-	-	-	-	-	-
Value	0.005	-	●	-	< 0.001	-	-	-	0.009	-	-
Orientation	-	0.012	0.012	●	0.003	-	-	-	-	-	-
Hue	< 0.001	< 0.001	0.001	< 0.001	●	< 0.001	< 0.001	< 0.001	0.002	0.020	0.020
Grain	-	-	0.023	-	< 0.001	●	-	-	-	-	-
Saturation	0.005	0.031	0.023	0.023	-	0.016	●	0.014	-	-	-
Crispness	-	-	-	-	0.001	-	0.034	●	-	-	-
Transpar.	0.004	0.012	0.023	0.016	0.016	0.004	-	0.009	●	-	-
Movement	0.012	-	-	0.031	0.003	0.004	-	-	0.044	●	-
Frequency	0.004	-	-	0.005	0.016	0.004	-	-	-	-	●

Table 2: Significant p-values for fdr-corrected post-hoc pairwise sign tests on preference scores (above bullets: longitudinal; below: lateral)

rank-based test [9, 37] corresponding to a 2×11 mixed design ANOVA was performed to investigate the impact of visual variables and driving function on preference scores. There was a significant difference in the preference ratings among visual variables within group [$F(1, 6.65) = 20.02, p < .001$]. Figure 6 summarises the results. The visual variable *hue* returned the highest preference ratings for both the lateral [$M_{lat} = 82.96$] and the longitudinal group [$M_{long} = 79.04$], followed by *transparency* [$M_{lat} = 68.17, M_{long} = 59.00$], *frequency* [$M_{lat} = 65.83, M_{long} = 59.17$], and *saturation* [$M_{lat} = 65.21, M_{long} = 58.04$]. Fdr-corrected [2] post-hoc pairwise sign tests were conducted to assess the differences within driving function. Table 2 summarises the resulting p-values for both driving functions. Further, a significant interaction between driving function and visual variables [$F(1, 6.65) = 2.56, p = .0139$] was found, indicating that the individual visual variables do not work equally well for lateral and longitudinal uncertainty.

Response Time and Preference

The applied repeated measures design within groups prevents the use of correlation coefficients due to differences in the variance between measurements taken on the same subject and those taken on different participants. Thus, multiple regression analysis was used to investigate if the response times significantly predict preference scores, taking out subjects as a factor [4].

The results of the regression indicated that the predictors explained 22.9% of the variance in the lateral group [$R^2 = .229, F(23, 229) = 2.857, p < 0.001$] and 17.2% in the longitudinal group [$R^2 = .172, F(23, 229) = 2.062, p = 0.004$]. It was found that response time predicted preference significantly in the lateral group [$\beta = -1.5452, p < 0.001$] and in the longitudinal group [$\beta = -0.6973, p = 0.041$].

DISCUSSION

The results of the study indicate a varying suitability of the analysed visual variables in the context of uncertainty communication in automated driving.

The sorting scores suggest that particularly *hue*, *saturation*, *transparency*, *frequency*, and *value* convey a clear order. Considering that the order direction is consistent between lateral and longitudinal group, it seems that the visual variables are ordered independent of the implementation. The fact that some variables, namely *size*, *value*, *saturation*, and *crispness*, were sorted in an opposing order to that identified by MacEachren [31] can be attributed to the previously indicated differences in connotation, with uncertainty relating more to urgency rather than vague information in the context of this publication. Further, a main effect of visual variables on preference scores suggests that participants prefer some variables over others. As indicated by post-hoc pairwise sign tests, particularly *hue* seems to significantly differ from other variables in terms of preference, irrespective of the driving function. This can potentially be attributed to a difference in salience compared to other variables, as indicated by stronger preattentive effects [48]. Additionally, participants might be more familiar with the use of colour to communicate urgency, hence preferring it due to a mere exposure effect. This is affirmed by the fact that response times significantly predicted preference ratings, as this indicates that participants preferred those variables that communicated a clear order to them. The fact that *hue* has not been identified as a suitable variable in GIScience can be attributed to the fact that MacEachren [31] used a combination of hue values (olive, green, purple) with a less familiar order. The fact that both preference scores and response times were not significantly different between groups suggests that the selected layouts work equally well for communicating uncertainties using visual variables. However, given that there was a significant interaction for both response times and preference scores, the implementation of some variables worked better for one of the groups. This can be attributed to the fact that especially those variables that do not rely on fill colour or alpha values require a different implementation. For instance, the variable *position* was varied along a different axis for each group. While the lateral group was shown a variation in the y-direction, i. e. a wider or narrower road, the longitudinal group was presented with a change in the z-axis, i. e. moving the bars upwards. Similar differences in implementation apply to *grain* and *orientation*.

Limitations

The presented results are to be understood only under consideration of the following limitations. First, implementing the visual variables in a driving scene using the previously selected layouts generates various degrees of freedom. As such, several assumptions had to be made, including the selection of a neutral base colour, the determination of specific instances for each variable on three levels (e. g. blue, purple and red for *hue*), and the transfer of these variables into a driving scene. Further, the study did not assess the impact of cultural influences on the results. For instance, colour hues might have a different meaning in other cultures [15]. Additionally, the impact of prior experience with ACC or lane assist may be another factor influencing the results. While this study attempted to mimic a driving environment to test the variables in context, a further evaluation in an immersive, dynamic scenario is essential. Additionally, drivers would not be able to see the different variable levels next to each other as shown in Figure 3a. Instead, it needs to be ensured that the variable levels can be distinguished without seeing them side by side. Importantly, different implementations of the variables, for instance red-to-green hue variation instead of blue-to-red, may lead to results that deviate from those described in this paper. Finally, some visualisation methods may be rendered unsuitable in the actual usage environment due to lighting conditions. However, the presented results significantly cut down the amount of suitable variables for further exploration and thus provide a valuable first step for AR-based uncertainty displays.

CONCLUSION AND OUTLOOK

The objective of this study was to evaluate a set of abstract visual variables regarding their suitability for communicating uncertainties using augmented reality displays in automated driving. With previous publications focusing on uncertainty communication using digital, analogue or representational display formats, the outlined study presents the first implications for communicating the same information in an AR display. The results indicate that particularly colour-based variables, especially hue, as well as animation-based variables communicate a clear order and receive high subjective preference ratings by participants.

Further research efforts are needed to focus on evaluating the impact of communicating function-specific uncertainties, potentially under consideration of partial takeovers (e. g. operators control the vehicle only laterally). Additionally, the presented results inform the development of a prototype AR display for assessment in a dynamic driving environment. Future research efforts could also focus on different ways of communicating uncertainties, for instance using peripheral awareness displays [32, 24], to enable users to perceive changes in system uncertainty while performing NDRTs.

ACKNOWLEDGMENTS

We thank all the volunteers who participated in the experiment for their time and valuable feedback.

REFERENCES

1. Johannes Beller, Matthias Heesen, and Mark Vollrath. 2013. Improving the Driver Automation Interaction: An Approach Using Automation Uncertainty. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 55, 6 (2013), 1130–1141. DOI: <http://dx.doi.org/10.1177/0018720813482327>
2. Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B* 57, 1 (1995), 289–300.
3. Jacques Bertin. 1967. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin.
4. Martin Bland. 2015. *An Introduction to Medical Statistics*. Oxford University Press. 448 pages.
5. Heiner Bubb, Klaus Bengler, Rainer E. Gruenen, and Mark Vollrath. 2015. *Automobilergonomie*. Springer Vieweg, Wiesbaden. 663–684 pages. DOI: <http://dx.doi.org/10.1007/978-3-642-13505-7>
6. Mary T. Dzindolet, Scott A. Peterson, Regina A. Pomranky, Linda G. Pierce, and Hall P. Beck. 2003. The role of trust in automation reliance. *International Journal of Human Computer Studies* 58, 6 (2003), 697–718. DOI: [http://dx.doi.org/10.1016/S1071-5819\(03\)00038-7](http://dx.doi.org/10.1016/S1071-5819(03)00038-7)
7. Mica R. Endsley. 1995. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37, 1 (1995), 32–64. DOI: <http://dx.doi.org/10.1518/001872095779049543>
8. European Road Transport Research Advisory Council. 2015. Automated Driving Roadmap. (2015). <http://www.ertrac.org/uploads/documentsearch/id38/ERTRAC>
9. Jos Feys. 2016. Nonparametric Tests for the Interaction in Two-way Factorial Designs Using R. *The R Journal* 8, 1 (2016), 367–378. <https://journal.r-project.org/archive/2016/RJ-2016-027/RJ-2016-027.pdf>
10. Richard Finger and Ann M. Bisantz. 2002. Utilizing graphical formats to convey uncertainty in a decision-making task. *Theoretical Issues in Ergonomics Science* 2, 1 (2002), 1–25. DOI: <http://dx.doi.org/10.1080/14639220110110324>
11. Kasha Geels-Blair, Stephen Rice, and Jeremy Schwark. 2013. Using System-Wide Trust Theory to Reveal the Contagion Effects of Automation False Alarms and Misses on Compliance and Reliance in a Simulated Aviation Task. *The International Journal of Aviation Psychology* 23, 3 (2013), 245–266. DOI: <http://dx.doi.org/10.1080/10508414.2013.799355>
12. Łukasz Halik. 2012. The analysis of visual variables for use in the cartographic design of point symbols for mobile Augmented Reality applications. *Geodesy and Cartography* 61, 1 (2012), 19–30. DOI: <http://dx.doi.org/10.2478/v10277-012-0019-4>

13. Christopher Healey and James Enns. 2012. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics* 18, 7 (2012), 1170–1188. DOI: <http://dx.doi.org/10.1109/TVCG.2011.127>
14. Christopher G. Healey and James T. Enns. 1999. Large datasets at a glance: combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics* 5, 2 (1999), 145–167. DOI: <http://dx.doi.org/10.1109/2945.773807>
15. Rüdiger Heimgärtner, Alkesh Solanki, and Helmut Windl. 2017. Cultural User Experience in the Car – Toward a Standardized Systematic Intercultural Agile UI/UX Design Process. In *Automotive User Interfaces: Creating Interactive Experiences in the Car*. 143–184. DOI: <http://dx.doi.org/10.1007/978-3-319-49448-7>
16. Tove Helldin, Göran Falkman, Maria Riveiro, and Staffan Davidsson. 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13 5* (2013), 210–217. DOI: <http://dx.doi.org/10.1145/2516540.2516554>
17. Sebastian Hergeth, Lutz Lorenz, Roman Vilimek, and Josef F. Krems. 2016. Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 58, 3 (2016), 509–519. DOI: <http://dx.doi.org/10.1177/0018720815625744>
18. Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 3 (2015), 407–434. DOI: <http://dx.doi.org/10.1177/0018720814547570>
19. A Hosseini, D Bacara, and M Lienkamp. 2014. A system design for automotive augmented reality using stereo night vision. *IEEE Intelligent Vehicles Symposium, Proceedings Iv* (2014), 127–133. DOI: <http://dx.doi.org/10.1109/IVS.2014.6856484>
20. Daniel E. Huber and Christopher G. Healey. 2005. Visualizing data with motion. *Proceedings of the IEEE Visualization Conference* (2005), 67. DOI: <http://dx.doi.org/10.1109/VIS.2005.125>
21. Boris Israel. 2012. *Potenziale eines kontaktanalogen Head-up Displays für den Serieneinsatz*. Ph.D. Dissertation. Technical University Munich. <https://mediatum.ub.tum.de/doc/1172451/1172451.pdf>
22. B. Julesz and J. R. Bergen. 1983. Human Factors and Behavioral Science: Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures. *Bell System Technical Journal* 62, 6 (1983), 1619–1645. DOI: <http://dx.doi.org/10.1002/j.1538-7305.1983.tb03502.x>
23. David Keller and Stephen Rice. 2010. System-Wide versus Component-Specific Trust Using Multiple Aids. *The Journal of General Psychology* 137, 1 (2010), 114–128. DOI: <http://dx.doi.org/10.1080/00221300903266713>
24. Alexander Kunze, Russell Marshall, Stephen J. Summerskill, and Ashleigh J. Filtness. 2017. Enhancing driving safety and user experience through unobtrusive and function-specific feedback. In *Adjunct Proceedings of the 9th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)*. ACM, Oldenburg, Germany, 183–189. DOI: <http://dx.doi.org/10.1145/3131726.3131762>
25. David R. Large, Gary E. Burnett, Andrew Morris, and Arun Muthumani. 2017. A Longitudinal Simulator Study to Explore Drivers' Behaviour During Highly-Automated Driving. In *Advances in Intelligent Systems and Computing*, Vol. 0. DOI: <http://dx.doi.org/10.1007/978-3-319-60441-1>
26. John D. Lee and Neville Moray. 1992. Trust, Control Strategies and Allocation of Function in Human-Machine Systems. *Ergonomics* 35, March (1992), 1243–1270. DOI: <http://dx.doi.org/10.1080/00140139208967392>
27. John D. Lee and Katrina A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors* 46, 1 (2004), 50–80.
28. M. S. Carpendale. 2003. *Considering Visual Variables as a basis for Information Visualisation*. Technical Report. University of Calgary. DOI: <http://dx.doi.org/10.5072/PRISM/30495>
29. Alan M. MacEachren. 1992. Visualizing uncertain information. *Cartographic Perspective* 13 (1992), 10–19. DOI: <http://dx.doi.org/10.1.1.62.285>
30. Alan M. MacEachren, Anthony Robinson, Susan Hopper, Steven Gardner, Robert Murray, Mark Gahegan, and Elisabeth Hetzler. 2005. Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science/Information Science* 32, 3 (2005), 139–160. DOI: <http://dx.doi.org/10.1559/1523040054738936>
31. Alan M. MacEachren, Robert E. Roth, James O'Brien, Derek Swingley, and Mark Gahegan. 2012. Visual Semiotics and Uncertainty Visualisation: An Empirical Study. In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18. IEEE, 2496–2505. DOI: <http://dx.doi.org/10.1109/TVCG.2012.279>
32. D. Scott McCrickard, C. M. Chewar, Jacob P. Somervell, and Ali Ndiwalana. 2003. A model for notification systems evaluation – assessing user goals for multitasking activity. *ACM Transactions on Computer-Human Interaction* 10, 4 (2003), 312–338. DOI: <http://dx.doi.org/10.1145/966930.966933>

33. John M. McGuirl and Nadine B. Sarter. 2006. Supporting trust calibration and the effective use of decision aids by presenting dynamic system confidence information. *Human Factors* 48, 4 (2006), 656–665. DOI : <http://dx.doi.org/10.1518/001872006779166334>
34. Joel L. Morrison. 1974. A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *International Yearbook of Cartography* 14 (1974), 115–127. DOI : <http://dx.doi.org/citeulike-article-id:9146840>
35. Heather F. Neyedli, Justin G. Hollands, and Greg A. Jamieson. 2011. Beyond Identity: Incorporating System Reliability Information Into an Automated Combat Identification System. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53, 4 (2011), 338–355. DOI : <http://dx.doi.org/10.1177/0018720811413767>
36. Brittany E. Noah, Thomas. M. Gable, Shao-Yu Chen, Shruti Singh, and Bruce N. Walker. 2017. Development and Preliminary Evaluation of Reliability Displays for Automated Lane Keeping. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17*. ACM, Oldenburg, Germany, 202–208. DOI : <http://dx.doi.org/10.1145/3122986.3123007>
37. Kimihiro Noguchi, Yulia R. Gel, Edgar Brunner, and Frank Konietschke. 2012. nparLD: An R Software Package for the Nonparametric Analysis of Longitudinal Data in Factorial Experiments. *Journal of Statistical Software* 50, 12 (2012), 1–23. <http://www.jstatsoft.org/v50/i12/>
38. Alex T. Pang, Craig M. Wittenbrink, and Suresh K. Lodha. 1997. Approaches to uncertainty visualization. *The Visual Computer* 13, 8 (1997), 370–390. DOI : <http://dx.doi.org/10.1007/s003710050111>
39. Raja Parasuraman and Dietrich Manzey. 2010. Complacency and bias in human use of automation: an attentional integration. *Human factors* 52 (2010), 381–410. DOI : <http://dx.doi.org/10.1177/0018720810376055>
40. Byoung Jun Park, Jeong Woo Lee, Changrak Yoon, and Kyong Ho Kim. 2015. Augmented reality and representation in vehicle for safe driving at night. *International Conference on ICT Convergence 2015: Innovations Toward the IoT, 5G, and Smart Media Era, ICTC 2015* (2015), 1261–1263. DOI : <http://dx.doi.org/10.1109/ICTC.2015.7354791>
41. Robert E. Roth. 2017. Visual Variables. In *The International Encyclopedia of Geography*, Douglas Richardson, Noel Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu, and Richard A. Marston (Eds.). John Wiley and Sohns, Ltd. DOI : <http://dx.doi.org/10.1002/9781118786352.wbieg0761>
42. SAE International. 2016. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016_201609)*. Technical Report.
43. Missie Smith, Jillian Streeter, Gary E. Burnett, and Joseph L. Gabbard. 2015. Visual Search Tasks: The Effects of Head-Up Displays on Driving and Task Performance. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15* (2015), 80–87. DOI : <http://dx.doi.org/10.1145/2799250.2799291>
44. Marcus Tönnis, Christian Lange, and Gudrun Klinker. 2007. Visual longitudinal and lateral driving assistance in the head-up display of cars. *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR (2007)*. DOI : <http://dx.doi.org/10.1109/ISMAR.2007.4538831>
45. A Treisman and S Gormican. 1988. Feature analysis in early vision - evidence from search asymmetries..pdf. *Psychological review* 95, 1 (1988), 15–48. DOI : <http://dx.doi.org/10.1037/0033-295X.95.1.15>
46. Anne M Treisman and Garry Gelade. 1980. A Feature-Integration of Attention. *Cognitive Psychology* 12, 1 (1980), 97–136. DOI : [http://dx.doi.org/10.1016/0010-0285\(80\)90005-5](http://dx.doi.org/10.1016/0010-0285(80)90005-5)
47. Lu Wang, Greg A. Jamieson, and Justin G. Hollands. 2009. Trust and reliance on an automated combat identification system. *Human factors* 51, 3 (2009), 281–291. DOI : <http://dx.doi.org/10.1177/0018720809338842>
48. Colin Ware. 2013. Visual Saliency and Finding Information. In *Information Visualization*. 139–177. DOI : <http://dx.doi.org/10.1016/B978-0-12-381464-7.00005-3>
49. Philipp Wintersberger, Tamara von Sawitzky, Anna-Katharina Frison, and Andreas Riener. 2017. Traffic Augmentation as a Means to Increase Trust in Automated Driving Systems. *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter - CHIItaly '17 September* (2017), 1–7. DOI : <http://dx.doi.org/10.1145/3125571.3125600>