Why is the Automation Not Available and When Can I Use It?

Simon Danner[(✉)] [iD], Tobias Hecht [iD], Benjamin Steidl, and Klaus Bengler [iD]

Chair of Ergonomics, Technical University of Munich, Garching, Germany
`simon.danner@tum.de`

Abstract. To this point, research regarding Human-Machine Interfaces (HMIs) and information needs in automated vehicles has focused on the period when the automation is active whereas this study investigated information needs regarding an AutomSSSated Driving System (ADS) while the same is not available. A driving simulator study ($N = 34$) was conducted to measure the effects of displaying the time until the automation will be available as well as reasons for non-availability on subjective ratings of the system and perceived system understanding. No effects were found for subjective ratings as usability, acceptance and workload, but there was a significant increase in perceived system understanding. Even though acceptance and usability showed no significant increase, participants ranked HMI concepts containing additional information higher than the baseline concept. The results indicate that HMIs giving information on the automation while it is not available are perceived favorably by participants and can increase transparency of the automation.

Keywords: Automated driving · Driving simulator study · Human-machine interface · Automation availability · Mental model

1 Introduction

Lower levels of automated driving are available today, while higher levels such as SAE Level 3 and 4, which allow users to conduct non-driving related activities (NDRAs) [1] are estimated to hit the market in near future. In fact, the possibility to conduct NDRAs seems to be one of the main reasons to use such ADS [2]. Even though future users have no experience with automated vehicles (AVs), their mental models regarding AVs are based upon their expectations towards the usage. It is important to investigate those expectations and mental models and therefore ways of correcting them, since in some cases they might not correspond to the reality and the correctness of mental models could affect the quality of human-machine interaction [3]. Giving information on the availability of ADS while driving manually is a possibility to correct users' mental models, which is investigated in this study with special focus on periods of non-availability.

1.1 System Understanding and Mental Models

Automated vehicles will be more complicated than advanced driver-assistance systems (ADAS) [4] and human operators will not always understand when an ADS is not available [5] and could even believe that SAE level 3 automated driving functions have to be available continuously, at least on highways [2]. Future users might initially use an ADS without experience but with having expectations leading to an internal representation of the system's functionalities, which might not correspond to the actual ADS capabilities [6]. This is in accordance with the finding that drivers over-rely on automated driving functions [7] resulting from a lack of understanding of how the systems work and what their limitations are [8]. The expectations towards and the understanding of the ADS are part of the users' mental models, as a mental model is the representation of a process or a system with its functionalities and dependencies in a person's mind [9]. When it comes to new experiences, existent mental models are compared to the perceived reality and in case the model was incorrect, it is adapted [3, 10]. Therefore, functionalities or limitations of a system as part of the mental model are not necessarily constant, since not experienced limitations can be erased from the model. Beggiato & Krems (2013) [3] found that known limitations of an ACC got forgotten after five rides without experiencing this limitation. This can lead to a mismatch between the expectations and the actual situation potentially causing ADS surprise [5] and frustration [11].

1.2 Information on the ADS When Not Available to Enhance System Understanding

Future users of SAE Level 3 and 4 automated driving functions will most likely lack of a deep understanding of the ADS and its limitations due to a lack of training [8] and reasons for non-availability of the function will not always be obvious for drivers, analogue to reasons for requests to intervene (RtIs) [12].

In an exploratory driving simulator study investigating potential information needs regarding an ADS, participants asked for a display of the reasons for non-availability as well as for a display of the time until the ADS will be available again [2]. Danner, Pfromm & Bengler (2020) [13] conducted a study showing that the display of the availability duration of a SAE Level 3 automation before activating the same influences drivers' behavior and rating of the system. Wandtner, Schömig & Schmidt (2018) [14] investigated information on the duration of availability and non-availability periods while driving manually and automatically, but the effect of this display while driving manually was not questioned in this study.

Information on the reasons for RtIs were shown to enhance perceived system understanding during a transition to manual driving [12]. Furthermore, the display of the availability duration before and after activating the ADS can have positive effects on subjective ratings of the system [13, 15–17]. Thus, these types of information shall be adapted to design an HMI giving information on the ADS while the same is not available. Hence, the newly designed HMI gives information on when the ADS will be available and displays the reasons for the current non-availability. Consequently, giving these information is hypothesized to correct the mental model and therefore might enhance the perceived system understanding and the subjective ratings of the ADS and its HMI.

2   Research Objectives

Based on considerations outlined above, this study aims at investigating the effect of additional information regarding an ADS when it is not available. Therefore, three HMI concepts were designed containing different amounts of information. One baseline concept (BC) is compared to two advanced concepts. Advanced concept one (AC1) contains information on how long it takes until the ADS will be available. Advanced concept two (AC2) contains the same information as AC1 but additionally informs about the reasons the ADS is not available. Furthermore, as reasons for non-availability are not always obvious, these concepts are compared on routes where the reasons are either visible for the driver or not. The concepts are investigated regarding acceptance, usability, workload, which are important measures for evaluating HMIs [18], and perceived system understanding.

Hence the following research questions are aimed to be answered in this work:

1. Is there a difference in the perceived system understanding in dependency of the presence or absence of information on the ADS when it is not available?
2. Is there a difference in the subjective ratings of the system in dependency of the presence or absence of information on when the ADS will be available?
3. Is there a difference in the subjective ratings of the system in dependency of the presence or absence of information on the reasons for non-availability of the ADS?
4. Is there a difference in the subjective measures (acceptance, usability and workload) in dependency of the obviousness of reasons for non-availability of the ADS?

3   Methods

3.1   Experimental Design, HMI Concepts and Procedure

To answer the research questions a 2×3 mixed design was chosen. The three HMI concepts served as within-factor while the obviousness of the reasons for non-availability served as between factor. The study participants were randomly assigned to either the group "obvious reasons" (OR) or to the group "non-obvious reasons" (NOR). The obvious reasons were a construction work on the highway and afterwards a missing lane marking. The non-obvious reasons were sensor error and missing map data. Moreover, in both conditions an additional reason (connection problem) was displayed. The participants experienced three rides with three different HMI concepts, always on the same route. The order of the HMI concepts was randomized to minimize sequence effects. The participants started on a highway resting area and joined the highway. The ADS was not available at the beginning. The BC gave no information on the ADS, AC1 displayed the time until the ADS will be available and AC2 displayed the duration as well as the reasons for non-availability. The manual drive took about 5 min. Subsequently, the ADS became available and the participants could drive automatically. When driving automatically, the participants were instructed to conduct an NDRA, which consisted of the game "2049" presented on a tablet computer simulating the central information display (CID). The automated drive took about three minutes. Afterwards, a highway exit appeared and an RtI requested the participants to take over the control over the

vehicle again. The reason for this RtI was the highway exit, which should be taken to be able to follow the route. The rest of the track led over a country road into a small city, where the participants stopped the car at a parking lot. After each experimental drive, subjective questionnaires were administered.

Regarding the HMI, the availability duration of the ADS was given when active [15, 17] and when available but still not activated [13]. Therefore, during the automated drive all concepts were the same. Figure 1 shows the HMI-concept displayed while the ADS was not available for AC2. The duration until the ADS will be available was shown in the top left corner for AC1 and AC2, explained by the words "automation available in X min". For AC2 the reasons for non-availability were displayed in the top right corner. An icon and a descriptive text was given, as in this example "impediment due to construction site", along with the text " +1 further". Participants had the chance to press the ADS button on the steering wheel while the ADS was not available to access a pop-up window displaying further reasons (also shown in Fig. 1). Since the participants were not instructed to try this button when the ADS was not available, it was anticipated that not all participants would see the pop-up. For this reason, a short drive was conducted after filling in the last standardized questionnaire and rating the perceived system understanding, where participants were instructed to try the button when the ADS was not available. Afterwards the pop-up as well as the interaction was rated.
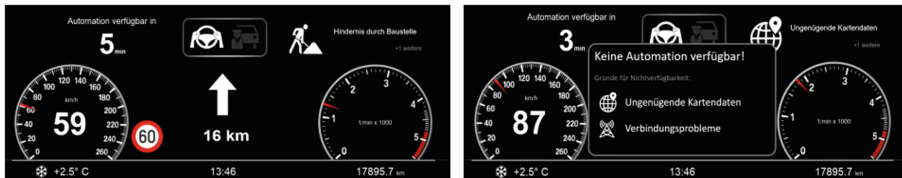

Fig. 1. The left picture shows the HMI for AC2 during a period of non-availability. AC1 was the same only without the reason for non-availability in the upper right corner. BL was the same as AC1 but without the time until availability in the top left corner. The right picture shows the pop-up in AC2 ("no automation available, due to insufficient map data and loss of connection").

3.2 Apparatus and Measures

Driving Simulator: The study was conducted in the driving simulator of the Chair of Ergonomics offering a 120° view due to three 55″ displays. The simulator is equipped with a motion platform, capable of simulating pitch and role motions. Side mirrors were displayed by smaller displays while the view of the rare mirror was integrated in the top of the middle screen. An additional display behind the steering wheel served as the Instrument Cluster (IC). For this study, the driving simulator was equipped with an ADS of SAE Level 3, which was only available on highways. To activate the ADS, the participants had to press a button on the steering wheel. To deactivate the ADS, the same button as for the activation could be used as well as the accelerator or brake paddle.

Acceptance: Acceptance was measured by means of the acceptance scale [19], using a semantic differential. The questionnaire consists of nine items and two dimension, satisfaction and usefulness.

Usability: Usability was measured using the System Usability Scale (SUS) [20]. This scale consists of 10 items, answered on a 5-point Likert scale. The overall score is built by summing the single answers (having scores from 0 to 4) and multiplying the sum by 2.5. Therefore, the highest possible value is 100.

Workload: NASA-rTLX is used for measuring the perceived workload. This is the short form of the NASA TLX. It has shown to be equally sensitive like the original. It consists of 6 scales answered from 0 to 20. The overall score is built by building the mean of the 6 different scales [21].

Perceived System Understanding: Perceived system understanding is investigated by administering two questions answered on a 5-point Likert scale. The first question referred to the time until the ADS will be available again ("I asked myself when the automation will be available again") and the second one referred to the reasons of non-availability ("It was clear at any point of time why the automation was not available").

3.3 Sample and Statistical Analysis

The sample consisted of $N = 34$ participants with an average age of $M = 30.59$ ($SD = 6.96$) years. Thirteen participants were female (38%). The mean duration of possession of a driver's license was $M = 12.35$ ($SD = 14.7$). Ten of the participants (29%) had already taken part in a driving simulator study and eight (24%) had already taken part in a driving simulator study concerning automated driving.

To answer the aforementioned research questions, mixed ANOVAs were conducted if the assumptions for parametric testing were not violated. If the condition of normal distribution was not fulfilled, the ANOVA was still performed, since this method is considered robust against this violation [22]. The alternative for ANOVA was the Friedman-Test. The statistical analysis was conducted using JASP.

Two questions regarding the perceived transparency or system understanding were asked. To answer them, a Friedman-Test was conducted for each question, since single item Likert scales cannot be considered interval scaled.

4 Results

Perceived Transparency/System Understanding
The first question referred to the understanding of when the ADS will be available. A significant effect was found for the within factor concept (Chi-Square(2) = 12.13; $p = .002$; Kendall's $W = .29$). Post-hoc tests by means of Wilcoxon-Tests revealed a moderate significant difference ($p_{holm} = .006$; $r = 0.50$) between BC ($Mdn = 4$) and

AC1 ($Mdn = 2$) and a strong significant difference between BC and AC2 ($Mdn = 2$) ($p_{holm} < .001$; $r = .63$), but no effect for the comparison between AC1 and AC2 ($p_{holm} = .25$). Higher values in this question indicate less understanding.

The second question referred to the understanding of why the ADS was not available. The Friedman test showed a significant effect for the concept factor (Chi-Square(2) = 19.61; $p < .001$; Kendall's $W = 0.54$). The post-hoc tests showed no significant effects ($p_{holm} = .056$) for the comparisons BC ($Mdn = 3$) vs. AC1 ($Mdn = 4$), but a large significant effect for BC vs AC2 ($Mdn = 5$) ($p_{holm} < .001$; $r = .70$) and a moderate significant effect for AC1 vs AC2 ($p_{holm} = .02$; $r = 0.47$). Here, higher values indicate more understanding.

Acceptance: Based on the literature regarding the acceptance scale [19], acceptance consists of two dimensions. Since our data did not fit the factor structure, an overall value for acceptance was used. This value was formed by adding the single item values and then dividing them by the number of items.

We conducted a mixed ANOVA to investigate the differences between the different HMI concepts, the difference between the two conditions (OR vs. NOR) as well as an interaction effect. The sphericity assumption was violated and therefore a Greenhouse-Geisser correction was used for the analysis. Homogeneity was given in all measures of acceptance. The ANOVA showed no significant results for the main effect of the condition (Greenhouse–Geisser $F(1.50, 46.63) = 2.57$, $p = 0.10$) and the group factor ($F(1, 31) = 0.60$, $p = 0.44$). Furthermore, no significant interaction could be found (Greenhouse–Geisser $F(1.50, 46.63) = 0.94$, $p = 0.38$). On a descriptive level is a tendency towards the AC1 ($M = 1.30$, $SD = 0.50$) concept, which received the highest rating, followed by the AC2 ($M = 1.17$, $SD = 0.65$) concept. The baseline concept achieved the lowest mean value ($M = 1.07$, $SD = 0.63$). The same order of subjective ratings can be observed when investigating the mean values dependent of the between-factor. The ratings of AC1 and AC2 are less different for the OR condition, while the difference becomes stronger in the NOR condition. The values for each condition are shown in Table 1.

Usability: For investigating effects on usability we conducted a mixed ANOVA. Homogeneity was given for every condition, while the sphericity assumption was not met and therefore a Greenhouse-Geisser correction was used. The ANOVA showed no significant main effect for the concept (Greenhouse–Geisser $F(1.4, 43.8) = 0.392$; $p = .646$) and no significant main effect for the conditions ($F(1, 31) = 3.058$; $p = .090$) and no significant interaction effect (Greenhouse–Geisser $F(1.4, 43.8) = 0.446$; $p = .575$). On a descriptive level, the averaged ratings are the highest for the BC ($M = 82.35$, $SD = 9.86$), followed by AC 1 ($M = 81.82$, $SD = 18.31$) and AC2 ($M = 79.62$, $SD = 17.40$). For the OR condition, AC1 is rated best, followed by the BC concept. For the NOR condition, the BC is rated best, followed by AC1.

Workload: A mixed ANOVA was conducted to investigate the workload in dependency of the conditions and concepts. The homogeneity and sphericity assumptions were met, so no correction was used for the analysis. The ANOVA showed no significant main effects for the concepts ($F(2, 62) = 0.270$; $p = .764$) and the conditions ($F(1, 31) = $

0.006; $p = .938$). Furthermore, no significant interaction could be found ($F(2, 62) = 1.436$; $p = .246$). Investigating the mean values on a descriptive level, workload was highest for the BC ($M = 5.32, SD = 2.61$), followed by AC2 ($M = 5.15, SD = 2.73$) and AC1 ($M = 5.09, SD = 2.83$). For the OR condition AC2 concept showed least workload, followed by BC. For NOR condition the AC1 concept evoked least workload, followed by BC. The single means for every group are shown in Table 1.

Table 1. Means and standard deviations for every group.

	Acceptance		Usability		Workload	
	OR	NOR	OR	NOR	OR	NOR
BC	1.17 (0.64)	0.94 (0.69)	83,38 (12.31)	83.25 (6.58)	25.45 (14.13)	25.73 (13.64)
AC1	1.29 (0.61)	1.32 (0.36)	86.03 (14.28)	77.34 (21.36)	26.91 (14.74)	24.51 (12.54)
AC2	1.28 (0.62)	1.06 (0.69)	83.09 (14.32)	75.94 (19.98)	23.91 (13.75)	27.03 (15.02)

Qualitative Statements: The participants had the possibility to comment on the different concepts. After each ride with the different concepts, they were asked if they had any thoughts regarding the ADS or if something was not clear during the ride. After the ride with BC, 8 participants stated they wondered when the ADS will be available and 2 stated they would have wanted to know the reasons for non-availability, while 4 participants stated that everything was clear. After the ride with AC1, 4 participants stated they would have wanted to know the reasons for non-availability displayed while 5 participants stated that everything was clear. After the ride with AC2 5 participants stated that everything was clear, two participant noted, that the non-availability was easily understandable. One of those participants remarked that the user interface was overloaded due to the display of the reasons.

After driving the three experimental drives the participants were asked to rank the experienced concepts by putting them into an order; the first rank for the concept they considered the best and the third rank for the concept they considered the worst. For the analysis, a concept on the first rank received 3 points, on the second rank 2 and on the last rank one point. The more points a concept received in sum, the better it was ranked. AC1 received the best rating with 78 points, followed by AC2 (75) and BC (45).

The answers on the question if the participants considered the ADS activation button on the steering wheel as adequate for accessing the pop-up with further reasons for non-availability revealed overall approving rating with a median of $Mdn = 5$. The question, if the display " +1 further" was adequate was also answered approvingly with a median of $Mdn = 4$.

5   Discussion and Limitations

The results of this study show, that additional information has a positive significant effect on the perceived system understanding. Transparency, whose increase is shown through the increased system understanding, is associated with trust in automation [23]. In this study, the displays which contribute to transparency do not contain safety relevant information and therefore no correlation with trust is investigated, since the definition of trust includes a vulnerability aspect [24]. Furthermore, there is no relation between the perceived system understanding and the subjective ratings of the system, which is why the effect of transparency on users' attitudes should be investigated in future studies.

Even though the additional information on the system aims on enhancing usefulness of the system, no significant effect on acceptance, which contains the aspect usefulness [19], could be found, but as described in the results section there is a small tendency for better acceptance ratings for the advanced concepts on a descriptive level. Even if both advanced concepts are rated higher, the acceptance does not seem to be increased in dependency of the amount of information, since concept AC2 is rated worse than AC1. An explanation could be that participants perceive the information of when the ADS will be available actually as useful as it helps for example with the planning of NDRAs, while the information on why the ADS is not available does not contribute to this construct. This is in accordance with the findings of [12], who have shown that displaying reasons for RtIs has no significant effect on acceptance and trust, but on perceived system understanding. Future research should investigate, if giving information on the reasons for non-availability helps users calibrate their trust in automation or if it is helpful to ensure a safe human-machine-interaction due to the formation of an adequate mental model of the system [25]. Furthermore, displaying reasons for non-availability might have a negative influence on the aspect of satisfaction, which is also contained in the construct of acceptance. Especially in the condition of non-obvious reasons for non-availability, the subjective ratings are worse in comparison to the other condition. Participants might simply accept when the ADS is not available and do not have the need to know why, especially in an artificial study context. The steady confrontation with technical reasons for why the ADS does not work might seem not helpful but could even be annoying. Therefore, it can be concluded that this display should not be shown permanently. Nonetheless, for users who want to learn about system limitations, this information should be accessible.

In addition, no significant effects could be found for usability, but a tendency on a descriptive level was observed. Independent of the condition (OR vs. NOR), BC was rated the best, followed by AC1 and then AC2. When the condition is taken into account, AC1 was rated best when the reasons for non-availability were obvious. In the NOR condition the usability decreased for AC1 and AC2. Participants might have felt an inconsistency in the NOR condition, since they perceived a road where they would expect an ADS to work while the displays in the IC signalized that it did not. Furthermore, the rating could be in favor of BC, since, regarding the information content, it is more similar to ICs in present vehicles. Nonetheless, especially for usability, the differences in the means were very small and the standard deviations in AC1 and AC2 were rather large, which is why these tendencies should be interpreted very cautiously.

There were also no significant effects for workload and only little differences between the means. This could be explained by the HMI concepts not being too distracting. For AC1 and BC, workload was lower when the reasons for non-availability were not obvious. This might be due to the more comfortable ride, as no road works were passed. On the other hand, for AC2 workload was higher in the NOR condition. This could be explained again by the discrepancy between the subjectively perceived suitability of the road for driving automated and the displays giving reasons for non-availability. This would be in accordance with findings indicating that discrepancies between a mental model and actual events – and as result updating the mental model – lead to cognitive effort [26]. Future research should investigate how non-obvious reasons affect the workload as well as the subjective ratings of the system over time.

It could be shown that the HMIs giving additional information were more popular than a basic HMI. One reason for missing significant effects could be that the information were not really useful to the participants, since they could not use the ADS the way they might would have used it in reality, meaning choosing NDRAs themselves which might be important to them. Due to the artificial setting in the driving simulator and the dictated NDRA, the participants might not have felt the urge to use the ADS and consequently, the information when they could use it and why it was not available had no effect on acceptance und usability. In future studies, participants should be in a situation where they really want to use or have to use the ADS to investigate effects on these standardized constructs again, since this research has shown that participants prefer the HMIs with additional information.

References

1. SAE: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016_201806) (2018)
2. Danner, S., Pfromm, M., Limbacher, R., Bengler, K.: Information needs regarding the purposeful activation of automated driving functions - an exploratory study. In: de Waard, D., Toffetti, A., Pietrantoni, L., Franke, T., Petiot, J.F., Dumas, C., et al. (eds.) HFES Europe, Nantes, FR (2020)
3. Beggiato, M., Krems, J.F.: The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. Transp. Res. Part F: Traffic Psychol. Behav. **18**, 47–57 (2013). https://doi.org/10.1016/j.trf.2012.12.006
4. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse. Abuse. Hum Factors **39**, 230–53 (1997). https://doi.org/10.1518/001872097778543886
5. Carsten, O., Martens, M.H.: How can humans understand their automated cars? HMI principles, problems and solutions. Cogn. Tech. Work. **29**, 415 (2018). https://doi.org/10.1007/s10111-018-0484-0
6. Ebnali, M., Hulme, K., Ebnali-Heidari, A., Mazloumi, A.: How does training effect users' attitudes and skills needed for highly automated driving? Transp. Res. Part F: Traffic Psychol. Behav. **66**, 184–95 (2019). https://doi.org/10.1016/j.trf.2019.09.001
7. Naujoks, F., Kiesel, A., Neukum, A.: Cooperative warning systems: the impact of false and unnecessary alarms on drivers' compliance. Accid. Anal. Prev. **97**, 162–75 (2016). https://doi.org/10.1016/j.aap.2016.09.009
8. Walker, G.H., Stanton, N.A., Salmon, P.: Trust in vehicle technology. IJVD **70**, 157 (2016). https://doi.org/10.1504/IJVD.2016.074419

9. Bach, N.: Mentale Modelle als Basis von Implementierungsstrategien. Deutscher Universitätsverlag, Wiesbaden (2000)
10. Stanton, N.A., Young, M.S.: A proposed psychological model of driving automation. Theor. Issues Ergon. Sci. **1**, 315–31 (2000). https://doi.org/10.1080/14639220052399131
11. Ochs, M., Pelachaud, C., Sadek, D.: An Empathic virtual dialog Agentto improve human-machine interaction. In: Padgham, L. (ed.) Richland, pp. 89–96. International Foundation for Autonomous Agents and Multiagent Systems, SC (2008)
12. Körber, M., Prasch, L., Bengler, K.: Why do i have to drive now? Post hoc explanations of takeover requests. Hum Factors **60**, 305–23 (2018). https://doi.org/10.1177/0018720817747730
13. Danner, S., Pfromm, M., Bengler, K.: Does information on automated driving functions and the way of presenting it before activation influence users' behavior and perception of the system? Information. **11**, 54 (2020). https://doi.org/10.3390/info11010054
14. Wandtner, B., Schömig, N., Schmidt, G.: Secondary task engagement and disengagement in the context of highly automated driving. Transp. Res. Part F: Traffic Psychol. Behav. **58**, 253–63 (2018). https://doi.org/10.1016/j.trf.2018.06.001
15. Hecht, T., Kratzert, S., Bengler, K.: The effects of a predictive HMI and different transition frequencies on acceptance, workload, usability, and gaze behavior during urban automated driving. Information **11**, 73 (2020). https://doi.org/10.3390/info11020073
16. Richardson, N.T., Flohr, L., Michel, B.: Takeover requests in highly automated truck driving: how do the amount and type of additional information influence the driver-automation interaction? MTI **2**, 68 (2018). https://doi.org/10.3390/mti2040068
17. Holländer, K., Pfleging, B.: Preparing drivers for planned control transitions in automated cars. In: Abdennadher, S., Alt, F. (eds.) The 17th International Conference, Cairo, Egypt, 25.11.2018–28.11.2018, pp. 83–92. ACM Press, New York (2018). https://doi.org/10.1145/3282894.3282928
18. François, M., Osiurak, F., Fort, A., Crave, P., Navarro, J.: Automotive HMI design and participatory user involvement: review and perspectives. Ergonomics **60**, 541–52 (2017). https://doi.org/10.1080/00140139.2016.1188218
19. van der Laan, J.D., Heino, A., de Waard, D.: A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. - Part C: Emerg. Technol. **5**, 1–10 (1997)
20. Brooke, J.: SUS: a "quick and dirty" usability scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B. (eds.) Usability Evaluation in Industry. Chapman and Hall/CRC, Boca Raton (1996)
21. Hart, S.G.: Nasa-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. s904–908 (2006)
22. Blanca, M.J., Alarcón, R., Arnau, J., Bono, R., Bendayan, R.: Non-normal data: is ANOVA still a valid option? Psicothema **29**, 552–7 (2017). https://doi.org/10.7334/psicothema2016.383
23. Kraus, J., Scholz, D., Stiegemeier, D., Baumann, M.: The more you know: trust dynamics and calibration in highly automated driving and the effects of take-overs, system malfunction, and system transparency. Hum. Factors **62**, 718–36 (2020). https://doi.org/10.1177/0018720819853686
24. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**, 50–80 (2004). https://doi.org/10.1518/hfes.46.1.50.30392
25. Bullinger, H.-J. (ed.): Software-Ergonomie '85: Mensch-Computer-Interaktion. Teubner, Stuttgart (1985)
26. Zwaan, R.A., Radvansky, G.A., Hilliard, A.E., Curiel, J.M.: Constructing multidimensional situation models during reading. Sci. Stud. Read. **2**, 199–220 (1998). https://doi.org/10.1207/s1532799xssr0203_2