

Interaction Methods for Teleoperation: A Usability Study on Remote Assistance in Highly Automated Vehicles

1st Alexandra Nick
*Institute of Human and
Industrial Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
alexandra.nick@kit.edu*

2nd Marco Käppler
*Institute of Human and
Industrial Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
marco.kaeppler@kit.edu*

3rd Barbara Deml
*Institute of Human and
Industrial Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
barbara.deml@kit.edu*

Abstract—The development of automated driving systems is currently limited by narrowly defined operational design domains, restricting their deployment in real-world conditions. One promising solution to overcome limitations is teleoperation, in particular remote assistance. Remote assistance enables human intervention in ambiguous traffic situations and incorrect sensory perceptions of the vehicle where guidance or aid in decision-making is needed. This allows us to solve problems that exceed the vehicle’s capabilities. In order to develop a functioning overall system and allow effective teleoperation, it is crucial to consider human factors for potential remote operators. Currently, there are a number of concepts about the interaction with such remote assistance systems. However, it is unclear how various interaction methods differ in terms of usability, workload, required interaction time, and which types of information are really important for remote operators.

To this end, this study investigates three interaction concepts for remote assistance. Perception modification, trajectory guidance, and setting waypoints are evaluated in terms of workload (NASA-TLX), usability (UEQ-S), and preference ranking. Eye tracking data were collected to analyze the visual attention distribution of the provided information. Eighteen participants took part in a controlled lab study using an interface to provide maneuver-based solutions for six real-world driving scenarios in which the automated vehicle requested assistance. The results show that participants prefer the waypoint-based interaction method the least, while usability and perceived workload do not differ significantly between the interaction methods. All show comparable levels of performance in terms of interaction time. The remote assistance system used for this study requires improvements regarding its overall acceptance, which will be considered in the ongoing development process.

This research contributes to designing an optimized remote assistance workplace, focusing on interface design, interaction efficiency, and workload reduction. The insights gained will guide further refinements in system design, ensuring that teleoperation enhances the safety and efficiency of highly automated driving in uncertain real-world conditions.

Index Terms—teleoperation, usability, highly automated vehicles, remote assistance, workload

I. INTRODUCTION

The vision of fully autonomous vehicles has been in people’s minds for a long time. Commuters and travelers dream

of being able to use the time they spend inside their vehicles.

In addition to this vision, highly automated vehicles (HAVs) have the potential to improve road safety by reducing the number of accidents [1]–[3]. Efficient mobility concepts, such as on-demand shuttles without safety drivers, can contribute to the expansion of public transport. On the one hand, this could compensate for the predicted shortage of staff in public transport, and on the other hand, the economic efficiency of transport companies could be reconciled with better access to rural areas. These developments may also reduce the demand for parking space, the number of privately owned vehicles, and overall greenhouse gas emissions [4].

Despite various advances in automated driving, these developments remain largely confined to showcase scenarios. In an ideal world of full automation, autonomous vehicles would be able to handle all situations seamlessly on their own. Full automation is indeed possible in closed systems with suitable infrastructure, such as the London Underground or restricted company premises. In the real world, however, we are still a long way from full automation due to mixed traffic on public roads and vulnerable road users.

These problems have led to a promising approach to teleoperation for Level 4 automated driving [5], [6]. Level 4 vehicles operate fully autonomously (e.g., motorways or urban shuttle routes) within their operational design domain (ODD). Passengers in the vehicle are not responsible for disengagement, but a remote operator intervenes in problematic, rare, unforeseen or unclear situations and fulfills the purpose of a fallback solution [7]. So far, HAVs are not able to function outside their ODD. Even simple problems lead to disengagement of the automated driving functions [5], [8]. Frequent problems of HAVs are often related to perception problems, unfamiliar situations, problems with decision-making, or ambiguity with contradictory information. Frequently mentioned examples include roadworks with changed street alignments, manipulated road signs (e.g., with stickers), road obstacles combined with solid road marks, poor weather conditions, and poor visibility [9], [10]. Another key problem in mixed traffic is dealing

with vulnerable road users due to ambiguous information or dynamic obstacles. Among others, these situations need human support to provide security and efficiency in road driving.

One potential solution to these challenges proposed by academia and industry is teleoperation, an umbrella term for remote driving, remote assistance, and remote monitoring, with various realization concepts [7], [11]. According to [5], remote assistance is defined as a remote-controlled support function in which a human operator provides situational assistance to an HAV without taking over the continuous driving task. This assistance can be information allocation, decision support, or control commands but remains limited to specific scenarios that cannot be routinely controlled. The vehicle remains responsible for the driving task and autonomously implements the assistance it receives [7], [12] unless it explicitly enables the operator to directly control it. In other words, remote assistance provides high-level decision support in edge cases of automation [13] by using human information processing capabilities and the ability to understand situations semantically. So far, only a few examples of prototypical systems for remote assistance exist. One approach established a user interface with a simple tablet-based solution. Only the most critical vehicle information is displayed without much situational context. Suggested maneuvers are proposed so they can be selected easily [14]. A more complex interface with a multi-monitor setup displays detailed information, including simulated driving situations, map information, and multiple status indicators that can be used to analyze complex traffic situations such as construction zones or urban environments [15]. Possible concepts for the input of control commands by the remote operator include waypoint or trajectory guidance, interactive path planning, and changing perception [7], [16].

There is a significant research gap regarding which interface is suitable for remote assistance and how the usability of different interaction methods compares. The relevance of various types of information can be linked to studies on remote driving [17], [18] and expert discussion results [6]. Despite the increasing importance of remote assistance, there is a lack of empirical research on which types of information are essential and how they should be presented to human operators.

It is known from various human-machine interaction research that effective usage relies on interface designs that support intuitive, efficient, and reliable collaboration between humans and automated systems [19]. This also applies to the interaction between remote operators and HAVs [11], where operators must assess, decide, and intervene based on limited visual and contextual information. Key concepts for investigating remote assistance interfaces should include usability, mental workload, and situational awareness. An intuitive design can reduce mental demand and prevent cognitive overload [20]. The interaction method itself plays a crucial role, as its design can contribute to transparency, reduce operational friction, and support intuitive use. Furthermore, situational awareness can be positively influenced by a well-structured interaction flow [21], [22], which in turn affects reaction times [23], decision-making processes [24], and error

rates.

In order to understand and develop remote assistance interfaces, it is essential to understand the types of information required for effective interaction. Visual attention measured with eye tracking can provide objective insights into visual perception, workload, and situational awareness [25], [26], complementing subjective evaluations. In addition, individual differences such as affinity for technology [27], immersive tendency [28]–[30], gaming experience, multitasking ability, and responsiveness [31], [32] may influence user performance and system perception.

Investigating these dimensions is therefore essential for developing interfaces that are both usable and cognitively efficient. This approach contributes to the field of remote assistance for SAE Level 4 vehicles by presenting an empirically grounded experimental setup to evaluate interaction quality, usability, workload, and user acceptance, thereby supporting the development of user-centered teleoperation interfaces for future mobility systems. The study systematically investigates the usability of remote assistance systems and addresses two key research questions: 1) Do different interaction methods have varying levels of usability, operator workload, and required interaction time? and 2) What types of information are important in remote assistance?

II. METHODS

A. Stimuli and Apparatus

The experimental setup of the remote assistance station is built in a laboratory room at the institute of the authors. It is a multi-monitor setup consisting of three screens arranged vertically one above the other, illustrated in Fig. 1. The upper screen (screen one) displays the recorded driving videos (LG, 24 inches). The screen in the middle is referred to as the dashboard and presents relevant information about the vehicle and its surroundings (LG, 24 inches). The bottom screen is a touch screen that is vertically mounted on the table for pen-based interaction (Asus ProArt, 15,6 inches). It represents the interactive unit where the driving maneuvers are performed. In order to measure participants' fixation on the different screens, a SmartEye stationary eye tracker with four cameras was installed around the screens.



Fig. 1. Experimental setup of the remote control station with a vertical triple monitor configuration (video screen, dashboard, interactive unit) and a four-camera based eye tracking system.



Fig. 2. Representative camera images from the six driving scenarios recorded on a rainy day were used in the study. The scenarios cover various urban and suburban conditions, including intersections, narrow streets, and road construction zones.



Fig. 3. Custom dashboard interface for remote vehicle supervision. The layout includes (from top to bottom) 1) problem notification, 2) communication status indicators, 3) satellite map with vehicle position, 4) contextual driving information, and 5) vehicle and sensor data.

The video screen displays the driving scenarios from a front view, and the side mirrors are in the left and right bottom corners. In total, six different driving scenarios, illustrated in Fig. 2, were presented during the study on the upper screen. All videos were recorded with three GoPro action cameras mounted on the roof and the side mirrors of a car. The videos were edited and cut with DaVinci Resolve 19 [33].

The dashboard and the interactive unit were created using the Axure [34] prototyping tool, and all the information provided was in German. The dashboard (Fig. 3) contains various static information. The displayed content is custom-made and oriented on recommendations [15], [17], [18]. Furthermore, this screen was used to show incoming vehicle requests (green screen with textbox "incoming vehicle request") and the break screen (blue screen with textbox "break").

The interactive unit (Fig. 4) on which the users make their input to solve the traffic situation consists of an abstract view of the current driving situation from above and is oriented on previous work [7], [22]. It contains the road, the obstacle as a blue rectangle, a small map section, and a top bar showing the current interaction method, editing tools, and a finished button.

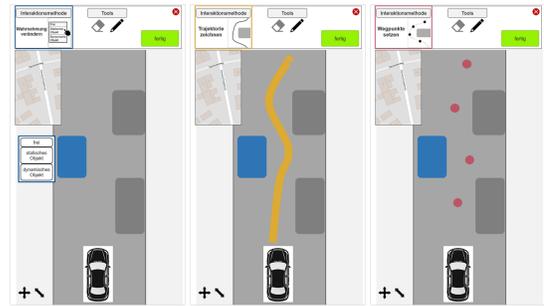


Fig. 4. Overview of the three interaction methods provided by the interactive unit: perception modification (left), trajectory guidance (center), and waypoint setting (right).

All interactions are accomplished with the provided touch pen. The participants can interact with the interactive unit in three different ways. In the variant *perception modification*, there are three selection buttons for 'free', 'static object', and 'dynamic object'. The participants have the opportunity to indicate the most suitable variant for the traffic situation. In the *trajectory guidance* variant, the path to be taken by the HAV can be drawn in; in the *setting waypoints* variant, waypoints to avoid the obstacle can be tapped as dots on the road.

B. Procedure

The controlled lab study is designed as a within-subject lab experiment in which the interaction method (perception modification, trajectory guidance, setting waypoints) is the independent variable. Workload, usability, and visual attention are the dependent variables. In addition, an overall evaluation of the provided setup for remote assistance is conducted.

Each participant was welcomed, informed about data protection and experimental procedure, and consented to voluntary participation. The participants first completed questionnaires on demographics, driving experience, affinity for technology and gaming experience. The task of interacting with the remote assistance interface was then explained to the participants and the eye tracking system was calibrated. This was followed by three test trials to familiarize the participants with the interaction methods. This phase included accepting incoming requests from the vehicles and confirming the interaction. The 18 experimental trials were divided into three blocks. In each block, the same six traffic situations were solved using one of the three interaction methods. The order of the six videos varied between the blocks. The sequence of the interaction methods was randomized. The vehicle requests occurred consecutively. The pauses between the submission of a maneuver and the subsequent vehicle request varied randomly between three and eight seconds. After each block, a short break was taken during which the participants answered the NASA-TLX [35] and UEQ-S [36] questions to evaluate the interaction method. Afterward, they proceeded to the following blocks of trials. Following the completion of all three conditions and their subsequent evaluation, the overall system was evaluated with regard to its acceptability and

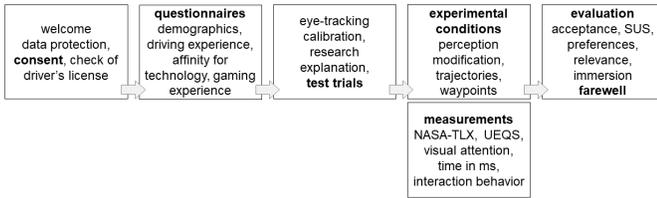


Fig. 5. Overview of the study procedure with a planned total duration of 45 to 50 minutes.

usability. Then, the relevance of the information provided was assessed, and the interaction methods were ranked. There was an opportunity for open feedback. Finally, the immersive tendency questionnaire [28] was completed. In the end, the participants were rewarded. An overview of the procedure is provided in Fig. 5.

C. Measurements

This study uses different data types to evaluate the systems' performance, the interaction methods regarding the associated workload and usability, and socio-demographics. Empirical performance metrics, validated questionnaires, and self-created items are used. All scales are checked regarding internal validity using Cronbach's α .

The system's and participants' performance is measured objectively with reaction times to incoming vehicle alerts as well as the interaction time in each trial from acceptance of the request to the submission of the maneuver. The experiment runs in a custom Python script and records data in milliseconds. Visual attention in terms of fixations in areas of interest [25], defined as the three screens, is measured with the four-camera SmartEye system [37]. The three maneuver-based interaction methods are evaluated regarding the perceived workload with the NASA-TLX [35] ($\alpha_{PM} = .62$, $\alpha_T = .77$, $\alpha_W = .78$) and usability with the UEQ-S [36], [38] ($\alpha_{PM} = .79$, $\alpha_T = .69$, $\alpha_W = .69$)¹. The setup as a whole is evaluated in terms of its acceptance, usability, and the relevance of presented information. Acceptance is measured on the subscales usefulness (without item 3, $\alpha = .58$) and satisfaction ($\alpha = .62$) [39]. Usability is measured with the SUS questionnaire [40] (without item 7, $\alpha = .74$). Open feedback questions were asked to gain qualitative insights into user experience, encountered limitations, and improvement possibilities.

To describe the sample age, gender, education level, driving frequency, vehicle usage and gaming experience (frequency and genres) [41] are measured. Affinity for technology [27] ($\alpha = .93$) and immersive tendency [28], [30] with the two subscales amount of involvement ($\alpha = .48$) and emotional involvement ($\alpha = .68$) are measured to understand possible trends in the results. All survey data is collected within a LimeSurvey questionnaire (version based on PHP 8.2.28) [42] on an additional tablet.

¹Cronbach's α values for PM = perception modification, T = trajectory guidance, W = setting waypoints

D. Sample

The sample consisted of 18 participants (11 male, 7 female) recruited as potential remote operators of automated vehicles. The age of the participants ranged from 19 to 35 years ($M = 26$, $SD = 4.38$). All participants were native speakers living in Germany and had a valid German driving license. All had normal or corrected-to-normal vision. Participants reported a moderate level of affinity for technology ($M = 4.01$, $SD = 1.01$) on the ATI scale [27]. In terms of immersive tendency, the sample showed a medium amount of involvement ($M = 26.24$, $SD = 5.25$) and a high amount of emotional involvement ($M = 30.76$, $SD = 4.55$). Given the small number of test subjects, this is a rather heterogeneous sample.

E. Data Processing

All questionnaire and time evaluation data are analyzed with R in RStudio [43], [44] using the packages dplyr, ggplot2, tidyr, psych, readr, ez, tidyverse, patchwork, ggpubr, and rstatix for computations and visualizations. In addition to descriptive analyses, the assumptions for parametric testing were evaluated using the Shapiro–Wilk test for normality and Mauchly's test for sphericity. Depending on the outcome, either repeated-measures ANOVAs or Friedman tests were conducted to assess whether descriptive differences between interaction methods are significant. Where applicable, post hoc comparisons were performed using Bonferroni-corrected paired t -tests (parametric) or Wilcoxon signed-rank tests with continuity correction (non-parametric).

III. RESULTS

A. Interaction methods

The participants' responses to the NASA-TLX items for workload measurements are illustrated as boxplots in Fig. 6. A repeated-measures ANOVA showed no significant effect of interaction method on mental demand, $F(2, 34) = 1.48$, $p = .243$, with a small effect size ($\eta^2 = .029$) and no significant effect of interaction method on frustration, $F(2, 34) = 3.03$, $p = .062$, with a small effect size ($\eta^2 = .059$). For the other four dimensions, the assumption of normality was violated. Accordingly, the Friedman test was computed for those four dimensions. There were no significant effects of the interaction method on physical demand ($\chi^2(2) = 5.93$, $p = .052$), on temporal demand ($\chi^2(2) = 1.24$, $p = .538$), and on perceived performance ($\chi^2(2) = 1.76$, $p = .414$). There is a significant effect of interaction method on perceived effort ($\chi^2(2) = 6.58$, $p = .037$). Pairwise comparisons using Wilcoxon signed-rank tests with continuity correction and Bonferroni adjustment revealed no significant differences in perceived effort between any pair of interaction methods (P-T: $p = .345$, P-W: $p = .075$, T-W: $p = .740$).

Based on the NASA-TLX scale score calculation, each interaction method is comparably low on a scale from 1 (low) to 21 (high). The means for the interaction methods are $M = 4.98$ ($SD = 2.34$) for perception modification, $M = 6.91$ ($SD = 3.38$) for trajectory guidance, and $M = 6.21$ ($SD = 2.94$) for setting waypoints. The results

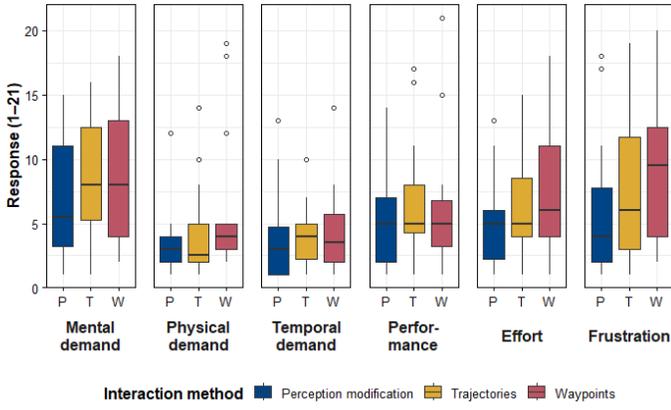


Fig. 6. Subjective workload (NASA-TLX) ratings across interaction methods. Ratings range from 1 (low) to 21 (high). Each box represents the median, IQR, and range for one workload dimension and interaction method.

of the Shapiro-Wilk test showed a violation of the normal distribution ($p < .05$). A Friedman test revealed a significant effect of interaction method on overall subjective workload, $\chi^2(2) = 8.34$ ($p = .015$). The executed post hoc Wilcoxon signed-rank tests with Bonferroni correction showed that participants reported significantly lower subjective workload in the perception modification condition compared to trajectories ($p = .042$) and waypoints ($p = .030$). No significant difference was found between trajectory guidance and setting waypoints ($p = .828$).

The participants reported medium UEQ-S values for all interaction methods. The perception modification was rated with a mean of $M = 4.94$ ($SD = 0.85$), the trajectory guidance with a mean of $M = 5.26$ ($SD = 0.73$), and setting waypoints with $M = 4.81$ ($SD = 0.70$). A repeated-measures ANOVA revealed a significant effect of interaction method on the usability ($F(2, 34) = 4.08$, $p = .026$), with a very small effect size ($\eta^2 = .061$). Pairwise comparisons using Bonferroni-corrected paired t -tests showed a significant difference between the waypoint and trajectory guidance conditions ($p = .012$). The other differences were not significant.

The results of the preference ranking illustrate that setting waypoints is the least favored input method. The distribution of participants' preferences is provided in Fig. 7.

The mean interaction time in seconds was lowest for perception modification ($M = 28.11$, $SD = 6.84$), followed by trajectory guidance ($M = 33.56$, $SD = 13.50$), and highest for setting waypoints ($M = 36.26$, $SD = 14.84$). A Friedman test indicated no significant differences in the interaction time across the three interaction methods, $\chi^2(2) = 4.00$, $p = .135$, but a descriptive trend is visible. The time analysis over all 18 trials grouped for each interaction method is illustrated in Fig. 8. It needs to be considered that the video order is different for each interaction method. Over all trials it shows a decline in the interaction duration. Further, some videos, especially where other vulnerable road users are involved, took longer to complete during the first interaction.

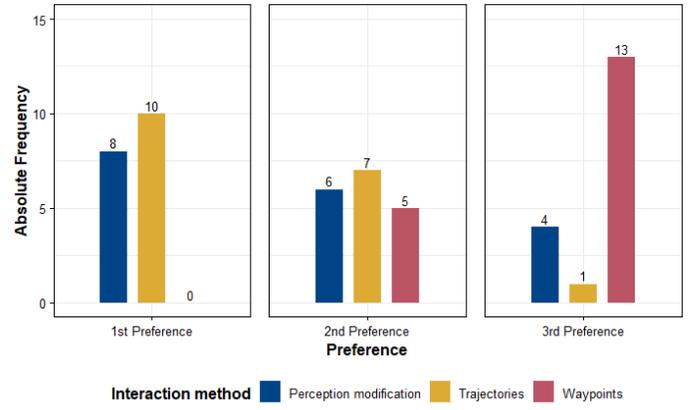


Fig. 7. Absolute frequencies of participants' ranked preferences for the three interaction methods. Bars indicate the number of participants who ranked each method as their first, second, or third preference.

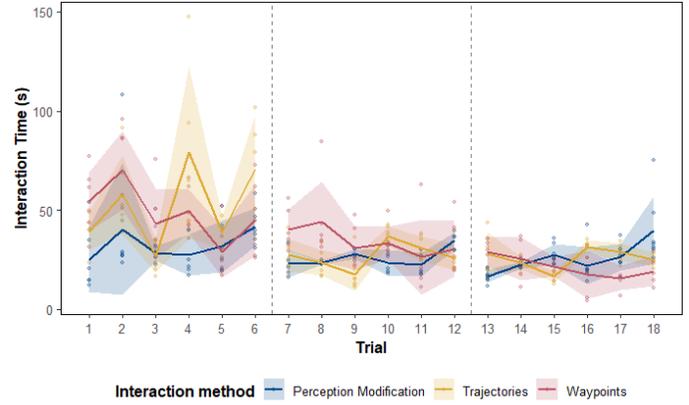


Fig. 8. Interaction time per trial across the three interaction methods. Lines represent mean interaction times (in seconds), with shaded areas indicating ± 1 standard deviation. Dashed lines separate experimental blocks (Trials 1-6, 7-12, and 13-18).

B. Setup evaluation

Usability and acceptance ratings were analyzed to assess users' overall perception of the system. Participants reported the overall usability with a SUS score of $M = 66.76$ ($SD = 12.68$), based on nine items (item 7 excluded), resulting in a possible range of 0 to 90. The acceptance subscale usefulness (excluding item 3) yielded a mean score of $M = 2.06$ ($SD = 0.53$). The subscale satisfaction showed a mean of $M = 2.10$ ($SD = 0.42$). Both scales range from 1 to 5.

The relevance of the presented information was rated from 1 (not relevant at all) to 7 (very relevant) for left mirror ($M = 5.00$, $SD = 1.46$), right mirror ($M = 5.00$, $SD = 1.50$), frontview ($M = 7.00$, $SD = 0.00$), velocity ($M = 2.61$, $SD = 1.75$), sensor information ($M = 2.39$, $SD = 1.38$), map information ($M = 1.72$, $SD = 1.36$), and passenger information ($M = 1.18$, $SD = 0.39$). The reported results are illustrated as boxplots in Fig. 9.

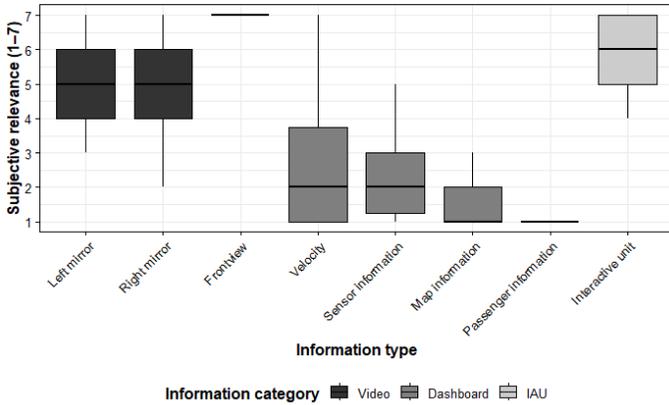


Fig. 9. Subjective relevance ratings for different information types grouped by interface category.

IV. DISCUSSION

A. Interaction methods

The three interaction methods (perception modification, trajectory guidance, and setting waypoints) were evaluated regarding user preference, interaction time, usability, and workload. The preference ranking indicates that setting waypoints is the least favored method. Perception modification and trajectory guidance are evaluated similarly. In particular, the lack of preference for waypoints is consistent with the results of expert discussions [6]. Participants reported that setting waypoints was less intuitive, which may be due to the absence of defined constraints, such as distance or the number of points. Another approach can be the use of virtual reality (VR) or extended reality (XR) to provide remote operators with more information and create a highly immersive environment of each driving situation. There are already several studies on the use of VR or XR for remote driving [17] and in vehicle take-over requests [45]. To date, this has not been investigated for the remote assistance of on-road vehicles. It is an open research question how VR or XR for remote assistance could look like and what influence this has on intuitiveness, performance, situational awareness, and mental demand. If higher immersion is suitable for more abstract assistance and frequent context switches, it warrants further investigation. Preliminary implementations for robot teleoperation [46] can guide further investigations in this direction. Furthermore, it is essential to consider that for workplaces with several-hour shifts, VR or XR could pose an additional load on remote operators.

Additionally, interaction time for maneuver-based assistance may influence input preferences. Perception modification was the fastest interaction method, although this result must be interpreted in light of its complexity. Selecting a predefined obstacle category is inherently less time-consuming than trajectory guidance or setting waypoints. Prior experience with touch pens may have influenced the preferences and should be recorded in future studies. The results for the interaction time, categorized by interaction method, showed a decline

in the time needed to complete the interaction. This decline likely reflects a learning effect, as participants understood the scenario and subsequent inputs required less time. Besides familiarity, learning effects can include motor learning and thereby increasing interface navigation efficiency [47]. The interaction order was randomized across participants to minimize carry-over effects. Still, repeated stimuli in different conditions should be avoided in further studies. Instead, different driving situations should be pre-tested regarding their complexity and distributed to the conditions in a matched manner. Another open question is how motivation and fatigue affect the evaluation of the interaction method over a more extended period of time. 18 trials are not sufficient for this. These effects could be investigated in a further study with significantly more scenarios.

Perceived workload and usability were assessed to complement the objective interaction data. Item-based analysis of the NASA-TLX revealed significant differences for perceived effort, although post hoc Wilcoxon comparisons were inconclusive. In contrast, the overall workload assessment showed significantly lower scores for perceptual modification compared to the other two methods. The different levels of abstraction can possibly explain this. Perceptual modification requires classification rather than precise spatial input, which contributes to lower cognitive effort. Usability scores differed significantly between the waypoint and trajectory conditions, though the effect size was small. This aligns with the reported workload levels and supports the interpretation that trajectory guidance was perceived as more usable and preferred in comparison to the waypoint method.

B. Setup evaluation

Overall, the setup was slightly negatively evaluated in terms of acceptance and usability, indicating the need for improvements in future iterations. This may be attributed to the limited functionality and static nature of interface elements, which may have reduced the perceived realism and credibility of the system. Furthermore, the use of the generic acceptance scale proved suboptimal. Several items did not align with the specific use case of remote assistance workplace evaluations, indicating the need for a tailored measurement instrument in future studies.

The evaluation of information elements supported the initial hypotheses regarding key aspects of remote assistance. However, the static dashboard, absence of sensor errors, and limited scenario variation require cautious interpretation. The passenger-related information was displayed in this study, but was not needed to solve any scenario. In real-world deployments, a broader set of conditions (e.g., urban, rural, highway) may lead to different informational needs. Especially in high-demand contexts, such as public transportation or call-based interventions, adaptive information presentation may be beneficial [48]. However, such adaptivity should also consider the cognitive cost of changing layouts or relocating key information spatially.

Participants' open feedback reinforced the quantitative findings. Setting waypoints was often perceived as not intuitive or unclear, while trajectory guidance was considered easier to understand and control. Future research should investigate how system feedback, visual guidance, or constraint-based design can improve the clarity of waypoint interactions. Human support systems should be optimized for cognitively demanding situations, especially when fast and precise responses are required [22]. Moreover, acceptance and performance may depend on context-specific variables, such as the type of tasks, environmental complexity, and interface familiarity. Further research should investigate how the relevance of information varies depending on the task, the situational complexity, and the cognitive costs associated with task and context switching. When interpreting the results, it is essential to acknowledge that this study involved only a small sample size. To generalize the findings to real-world applications, additional studies with more participants ($n > 36$)² and a broader range of driving situation complexities are necessary [49]–[51].

V. CONCLUSION

The evaluation of the interaction methods suggests that waypoints should be disregarded as an input method. The importance of the information presented should be further investigated in future studies that examine a wider range of driving situations. Furthermore, future studies should examine the complexity of the situation and the task itself in greater detail. In different situations, the various input methods are likely to vary in their suitability for assisting the vehicle. These factors require further consideration. The usability evaluation of the remote assistance setup reveals a need for improvements in the implementation. Despite the limitations, including the simplified setup version, small effect sizes, considerable variance in the results, and a small sample size, the study provides valuable insights into the usability and behavior in realistic remote assistance workplaces. By identifying usability challenges and cognitive demands, this study provides a foundation for enhancing the role of human operators in teleoperated vehicle support.

Supplementary material is available at <https://osf.io/4mcj3>. This includes all questionnaire items, the Python code for running the experiment, the datasets, and the analysis scripts.

REFERENCES

- [1] S. Sohrabi, A. Khodadadi, S. M. Mousavi, B. Dadashova, and D. Lord, "Quantifying the automated vehicle safety performance: A scoping review of the literature, evaluation of methods, and directions for future research," *Accident Analysis & Prevention*, vol. 152, pp. 1–21, paper no. 106003, 2021. [Online]. Available: <https://doi.org/10.1016/j.aap.2021.106003>
- [2] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey;" (*Traffic Safety Facts Crash•Stats*, vol. Report No. DOT HS 812 506, pp. 1–3, 2018. [Online]. Available: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506>
- [3] N. Lübbe, H. Jeppsson, A. Ranjbar, J. Fredriksson, J. Bårgman, and M. Östling, "Predicted road traffic fatalities in Germany: The potential and limitations of vehicle safety technologies from passive safety to highly automated driving," in *Conference proceedings International Research Council on the Biomechanics of Injury, IRCOBI*, 2018, pp. 17–52. [Online]. Available: <https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/11.pdf>
- [4] B.-H. Nahmias-Biran, G. Dadashev, and Y. Levi, "Sustainable automated mobility-on-demand strategies in dense urban areas: A case study of the Tel Aviv metropolis in 2040," *Sustainability*, vol. 15, no. 22, 2023, paper no. 16037. [Online]. Available: <https://doi.org/10.3390/su152216037>
- [5] On-Road Automated Driving Committee, *Autonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, apr 2021. [Online]. Available: https://doi.org/10.4271/J3016_202104
- [6] D. Brecht, N. Gehrke, T. Kerbl, N. Krauss, D. Majstorović, F. Pfab, M.-M. Wolf, and F. Diermeyer, "Evaluation of teleoperation concepts to solve automated vehicle disengagements," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 5, pp. 629–641, 2024. [Online]. Available: <https://doi.org/10.1109/OJITS.2024.3468021>
- [7] D. Majstorovic, S. Hoffmann, F. Pfab, A. Schimpe, M.-M. Wolf, and F. Diermeyer, "Survey on teleoperation concepts for automated vehicles," pp. 1–7, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2208.08876>
- [8] F. Tener and J. Lanir, "Driving from a distance: Challenges and guidelines for autonomous vehicle teleoperation interfaces," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022, paper no. 250. [Online]. Available: <https://doi.org/10.1145/3491102.3501827>
- [9] L. Kang, W. Zhao, B. Qi, and S. Banerjee, "Augmenting self-driving with remote control: Challenges and directions," in *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*, ser. HotMobile '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 19–24. [Online]. Available: <https://doi.org/10.1145/3177102.3177104>
- [10] C. Kettwich, A. Schrank, H. Avsar, and M. Oehl, "A helping human hand: Relevant scenarios for the remote operation of highly automated vehicles in public transport," *Applied Sciences*, vol. 12, no. 9, 2022, paper no. 4350. [Online]. Available: <https://doi.org/10.3390/app12094350>
- [11] C. Kettwich, A. Schrank, and M. Oehl, "Teleoperation of highly automated vehicles in public transport: User-centered design of a human-machine interface for remote-operation and its expert usability evaluation," *Multimodal Technologies and Interaction*, vol. 5, no. 5, 2021, paper no. 26. [Online]. Available: <https://doi.org/10.3390/mti5050026>
- [12] F. Tener and J. Lanir, "Investigating intervention road scenarios for teleoperation of autonomous vehicles," *Multimedia Tools and Applications*, 2023. [Online]. Available: <https://doi.org/10.1007/s11042-023-17851-z>
- [13] H. Parr, C. Harvey, and G. Burnett, "Defining roles of the remote operator in autonomous vehicles," in *Chartered Institute of Ergonomics & Human Factors*, 2023. [Online]. Available: <https://publications.ergonomics.org.uk/uploads/Defining-Roles-of-the-Remote-Operator-in-Autonomous-Vehicles.pdf>
- [14] F. Tener and J. Lanir, "Design of a high-level guidance user interface for teleoperation of autonomous vehicles," in *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. AutomotiveUI '23 Adjunct. New York, NY, USA: Association for Computing Machinery, 2023, pp. 287–290. [Online]. Available: <https://doi.org/10.1145/3581961.3609845>
- [15] A. Schrank, F. Walocha, S. Brandenburg, and M. Oehl, "Human-centered design and evaluation of a workplace for the remote assistance of highly automated vehicles," *Cognition, Technology & Work*, vol. 26, pp. 183–206, 2024. [Online]. Available: <https://doi.org/10.1007/s10111-024-00753-x>
- [16] D. Majstorović and F. Diermeyer, "Dynamic collaborative path planning for remote assistance of highly-automated vehicles," in *2023 IEEE International Automated Vehicle Validation Conference (IAVVC)*, 2023, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/IAVVC57316.2023.10328097>
- [17] M. Gafert, A. G. Mirnig, P. Fröhlich, V. Kraut, Z. Anzur, and M. Tscheligi, "Effective remote automated vehicle operation: a

²Following a G*Power analysis (G*Power 3.1.9.7 for Windows), the subsequent studies with the same design should have at least 36 evaluable participants.

- mixed reality contextual comparison study,” *Personal and Ubiquitous Computing*, vol. 27, pp. 2321–2338, 2023. [Online]. Available: <https://doi.org/10.1007/s00779-023-01782-5>
- [18] O. Biletska, S. Beckmann, T. Glimm, and H. Zadek, “Key requirements and concept for the future operations control center of automated shuttle buses,” in *2021 International Scientific Symposium on Logistics: Conference Volume*, 2021.
- [19] D. Mourtzis, J. Angelopoulos, and N. Panopoulos, “The future of the human-machine interface (hmi) in society 5.0,” *Future Internet*, vol. 15, no. 5, 2023, paper no. 162. [Online]. Available: <https://doi.org/10.3390/fi15050162>
- [20] F. Walocha, A. Schrank, H. P. Nguyen, and K. Ihme, “Multimodal assessment of mental workload during automated vehicle remote assistance: Modeling of eye-tracking-related, skin conductance, and cardiovascular indicators,” *Information*, vol. 16, no. 1, 2025, paper no. 64. [Online]. Available: <https://doi.org/10.3390/info16010064>
- [21] X. Tan and Y. Zhang, “Driver situation awareness for regaining control from conditionally automated vehicles: A systematic review of empirical studies,” *Human Factors*, vol. 67, no. 4, pp. 367–403, 2025. [Online]. Available: <https://doi.org/10.1177/00187208241272071>
- [22] G. Graf, Y. Abdelrahman, H. Xu, Y. Abdrabou, D. Schitz, H. Hußmann, and F. Alt, “The Predictive Corridor: A Virtual Augmented Driving Assistance System for Teleoperated Autonomous Vehicles,” in *ICAT-EGVE 2020 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, F. Argelaguet, R. McMahan, and M. Sugimoto, Eds. The Eurographics Association, 2020. [Online]. Available: <https://doi.org/10.2312/egve.20201260>
- [23] W. Tang, Y. Tian, L. Tan, S. Xue, T. Jiang, S. Zhou, and C. Wang, “Human factors design and evaluation of china’s space manipulator teleoperation system,” *International Journal of Human-Computer Interaction*, vol. 40, no. 8, pp. 1943–1959, 2023. [Online]. Available: <https://doi.org/10.1080/10447318.2023.2250055>
- [24] E. Coskun and M. Grabowski, “Impacts of user interface complexity on user acceptance and performance in safety-critical systems,” *Journal of Homeland Security and Emergency Management*, vol. 2, no. 1, 2005. [Online]. Available: <https://doi.org/10.2202/1547-7355.1109>
- [25] Y. Guo, D. Freer, F. Deligianni, and G.-Z. Yang, “Eye-tracking for performance evaluation and workload estimation in space telerobotic training,” *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 1, pp. 1–11, 2022. [Online]. Available: <https://doi.org/10.1109/THMS.2021.3107519>
- [26] J. Theeuwes, A. Belopolsky, and C. N. Olivers, “Interactions between working memory, attention and eye movements,” *Acta Psychologica*, vol. 132, no. 2, pp. 106–114, 2009, spatial working memory and imagery: From eye movements to grounded cognition. [Online]. Available: <https://doi.org/10.1016/j.actpsy.2009.01.005>
- [27] C. A. Thomas Franke and D. Wessel, “A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ati) scale,” *International Journal of Human-Computer Interaction*, vol. 35, no. 6, pp. 456–467, 2019. [Online]. Available: <https://doi.org/10.1080/10447318.2018.1456150>
- [28] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” *Presence*, vol. 7, no. 3, pp. 225–240, 1998. [Online]. Available: <https://doi.org/10.1162/105474698565686>
- [29] H. Pongrac, “Gestaltung und evaluation von virtuellen und telepräsenzsystemen an hand von aufgabenleistung und präsenzempfinden,” Ph.D. dissertation, München, Universität der Bundeswehr, Dissertation von 2008, 2011. [Online]. Available: <https://d-nb.info/101550308X/34>
- [30] R. Scheuchenpflug, “Measuring presence in virtual environments,” in *HCI International*, vol. 2001. HCI International, 2001, pp. 56–58.
- [31] C. S. Green and D. Bavelier, “Effect of action video games on the spatial distribution of visuospatial attention,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 32, no. 6, pp. 1465–1478, 2006. [Online]. Available: <https://doi.org/10.1037/0096-1523.32.6.1465>
- [32] M. W. Dye, C. S. Green, and D. Bavelier, “Increasing speed of processing with action video games,” *Current Directions in Psychological Science*, vol. 18, no. 6, pp. 321–326, 2009. [Online]. Available: <https://doi.org/10.1111/j.1467-8721.2009.01660.x>
- [33] Blackmagic, “Davinci resolve 19,” Blackmagic Design Pty Ltd, South Melbourne, Victoria Australia, 2025. [Online]. Available: <https://www.blackmagicdesign.com/de/products/davinciresolve>
- [34] Axure Software Solutions, “Axure rp - ux prototyping tool,” Axure Software Solutions Inc, San Diego, United States of America, 2025, version 11. [Online]. Available: <https://www.axure.com>
- [35] S. G. Hart and L. E. Staveland, “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” in *Human Mental Workload*, ser. Advances in Psychology, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, vol. 52, pp. 139–183. [Online]. Available: [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [36] M. Schrepp, J. Thomaschewski, and A. Hinderks, “Design and evaluation of a short version of the user experience questionnaire (ueq-s),” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 6, pp. 103–108, 12/2017 2017. [Online]. Available: <https://doi.org/10.9781/ijimai.2017.09.001>
- [37] Smart eye GmbH, “Smart eye pro (v 9.3),” 2020. [Online]. Available: <https://www.smarteye.se/technology/eye-tracking/>
- [38] M. Schrepp, J. Thomaschewski, and A. Hinderks, “Construction of a benchmark for the user experience questionnaire (ueq),” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 4, pp. 40–44, 06/2017 2017. [Online]. Available: <https://doi.org/10.9781/ijimai.2017.445>
- [39] J. D. Van Der Laan, A. Heino, and D. De Waard, “A simple procedure for the assessment of acceptance of advanced transport telematics,” *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 1, pp. 1–10, 1997. [Online]. Available: [https://doi.org/10.1016/S0968-090X\(96\)00025-3](https://doi.org/10.1016/S0968-090X(96)00025-3)
- [40] J. Brooke, *SUS: A quick and dirty usability scale*. CRC Press, 1996, ch. 21. [Online]. Available: <https://doi.org/10.1201/9781498710411-35>
- [41] O. Koch, “Gaming-plattformen und videospiele-genres im überblick,” 2025, accessed: 18.2.2025. [Online]. Available: <https://www.verivox.de/internet/ratgeber/gaming-plattformen-und-videospiel-genres-im-ueberblick-1120271/>
- [42] Limesurvey GmbH, “Limesurvey: An open source survey tool,” LimeSurvey GmbH, Hamburg, Germany, 2012. [Online]. Available: <http://www.limesurvey.org>
- [43] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: <https://www.R-project.org/>
- [44] RStudio Team, *RStudio: Integrated Development Environment for R*, RStudio, Inc., Boston, MA, 2019. [Online]. Available: <http://www.rstudio.com/>
- [45] A. Mukhopadhyay, V. K. Sharma, P. T. Gaikwad, A. K. Sandula, and P. Biswas, “Exploring the use of xr interfaces for driver assistance in take over request,” in *Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. AutomotiveUI ’22. New York, NY, USA: Association for Computing Machinery, 2022, p. 58–61. [Online]. Available: <https://doi.org/10.1145/3544999.3552527>
- [46] Y.-P. Su, X.-Q. Chen, C. Zhou, L. H. Pearson, C. G. Pretty, and J. G. Chase, “Integrating virtual, mixed, and augmented reality into remote robotic applications: A brief review of extended reality-enhanced robotic systems for intuitive telemanipulation and telemanufacturing tasks in hazardous conditions,” *Applied Sciences*, vol. 13, no. 22, 2023. [Online]. Available: <https://doi.org/10.3390/app132212129>
- [47] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, “The role of deliberate practice in the acquisition of expert performance,” *Psychological Review*, vol. 100, no. 3, pp. 363–406, Jul 1993. [Online]. Available: <https://doi.org/10.1037/0033-295X.100.3.363>
- [48] Y. Trabelsi, O. Shabat, J. Lanir, O. Maksimov, and S. Kraus, “Advice provision in teleoperation of autonomous vehicles,” in *Proceedings of the 28th International Conference on Intelligent User Interfaces*, ser. IUI ’23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 750–761. [Online]. Available: <https://doi.org/10.1145/3581641.3584068>
- [49] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd ed. London, England: Routledge, May 2013.
- [50] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, “G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” *Behavior Research Methods*, vol. 39, pp. 175–191, May 2007.
- [51] F. Faul, E. Erdfelder, A. Buchner, and A.-G. Lang, “Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses,” *Behavior Research Methods*, vol. 41, pp. 1149–1160, Nov. 2009.