Personal and Ubiquitous Computing (2023) 27:2321–2338
https://doi.org/10.1007/s00779-023-01782-5

ORIGINAL ARTICLE

Effective remote automated vehicle operation: a mixed reality contextual comparison study

Michael Gafert[1] · Alexander G. Mirnig[2] · Peter Fröhlich[1] · Vanessa Kraut[1] · Zoja Anzur[1] · Manfred Tscheligi[2]

Received: 20 December 2022 / Accepted: 26 October 2023 / Published online: 15 December 2023

Abstract
With the increasing pervasion of automated vehicle fleets, there is an equally increased need to provide effective remote operation capabilities for interventions in cases of vehicle malfunction, rough road conditions or unnavigable areas. This shift in the human operator role to that of an observer and occasional teleoperator requires appropriate interaction interfaces. In this paper, we present an in-context analysis of previously proposed teleoperation information requirements realized as concrete user interfaces in a user study (N=16). Participants completed multiple tasks in a mixed approach, controlling an actual miniature vehicle via an XR enabled headset which showed two interface manifestations (minimum and maximum requirements realized). From the results, we derive separate sets of fundamental UI elements per requirement for effective teleoperation in two phases of interaction: orientation and navigation. We conclude with further suggestions to extend teleoperation UIs regarding contexts and tasks beyond the navigation use case used in the study.

Keywords Automated vehicles · Remote operation · Teleoperation · Interfaces · Extended reality

1 Introduction and motivation

Vehicle Automation is a continuous process that strives to bring in-vehicle technologies as well as related infrastructure ever closer towards full automation. In open traffic and mixed environments, higher automation is more difficult to achieve than in closed and more controllable environments. Such environments, e.g. airports, warehouses, or harbor cargo areas, enable more control over vehicle speeds (limited distances means less need for high velocities) and traffic stakeholders (few or even no manually controlled vehicles or pedestrians). Due to this, they provide a valuable glimpse into the highly automated ubiquitous traffic of the future.

In such a high to fully automated traffic environment, there is no longer a need for a human driver behind the wheel at all times. This does not mean, however, that the human will no longer be in interaction with the vehicle. Rather, the role will shift from an (active) driver to that of an operator, as far as interaction with the vehicle controls is concerned. Instead of being behind the wheel of a single vehicle and intervene when needed, an operator supervises a fleet of vehicles, with said fleet operating mostly autonomously (in the sense that routing, maneuvers etc. are performed via automation, base parameters are still likely to be defined by human operators). Remote operation refers to management of one or more automated vehicles. It includes both driving and non-driving related tasks, such as monitoring (constant) and intervention (when needed). Remote operation is a synonym for teleoperation. Furthermore, remote driving refers to the real-time driving of a vehicle by a remote driver and remote assistance refers to provision of information to the vehicle by a remote operator [19]. In addition, the workplace of an operator is outside the vehicle, possibly even from a remote location.

✉ Alexander G. Mirnig
alexander.mirnig@ait.ac.at; alexander.mirnig@plus.ac.at

Michael Gafert
michael.gafert@ait.ac.at

Peter Fröhlich
peter.froehlich@ait.ac.at

Vanessa Kraut
vanessa.kraut@ait.ac.at

Zoja Anzur
zoja.anzur@ait.ac.at

Manfred Tscheligi
manfred.tscheligi@ait.ac.at

1 AIT Austrian Institute of Technology, Giefinggasse 4, Vienna 1210, Austria

2 AIT Austrian Institute of Technology & University of Salzburg, Giefinggasse 4, Vienna 1210, Austria

As such, the interaction of an operator is drastically different from monitoring and controlling in-vehicle, because the operator is physically disconnected from the vehicle [33]. Moreover, the vehicle is beyond visual line-of-sight of the operator [19]. Since the operator is monitoring the fleet from an outside perspective, there is a greater need for additional contextual information (vehicle task, vehicle status, driving environment) if an intervention is required. In order to address the challenges brought about by the remote monitoring and operation context, several sets of requirements for remote operation interfaces have been identified and proposed [14, 21, 22].

Due to the often stark differences in application contexts, novelty of full vehicle automation in closed environments, and general difficulty of evaluating User Interfaces (UIs) in a field setting, there is still little information on actually implemented UIs that realize these requirements (or subsets thereof) and how well they work in practice. The goal and intended contribution of this paper is to address this gap by implementing a teleoperation UI based on the requirements outlined by Graf and Hussmann [15] and evaluating it in terms of how well it enables a remote operator to complete a driving task if an intervention is required. After an introduction into the related work in Section 2, we demonstrate the consolidation of the requirements and the operationalization towards a contextualized research study prototype (Section 3). We then present in Section 4 the results from a comparison study conducted in a laboratory setting (N=16). In the study, we compared two versions of the UI (one with a minimal, one with a maximal set of in-UI indicators) in order to assess which elements were more or less useful and in which phase of the intervention (orientation vs. actual driving task). Furthermore, we were interested in understanding whether difficulty of the driving task had any influence on user reactions to the system and its features (see the concrete research questions in Section 4). The study used an extended reality setup, where the remote operator controlled the remotely operated vehicle via a headset from a first person perspective via cameras mounted on the vehicle. The vehicle was a miniature remote controlled vehicle which drove around a real miniature environment. This allowed us to create a more realistic interaction context compared to a fully digital simulation.

2 Related work

In literature [24], teleoperation is mentioned as a way of dealing with certain drawbacks of automated vehicles as it allows the vehicles to be supported by human operators. Human factors aspects of teleoperation have been studied since decades, for different application domains [2, 23, 32, 38], but there is a general consensus that many issues remain [29]. In the case of complex scenarios, such as the teleoperation of automated vehicles, new challenges arise due to their unique characteristics (e.g., the driver not physically occupying the vehicle). [7, 25]. Furthermore, there are many various teleoperation concepts (e.g., direct control, shared control), each with certain advantages and disadvantages [24].

The literature in this field of research also mentions cooperation driving (e.g., [3, 36, 37]) with its advantage being that the driver does not have to be in control of the vehicle all the time, but only when needed. In their article, Colley et al. [3]. Present ORIAS - a system for cooperative driving, with two aspects: object identification and action selection. The first aspect relates to the help the system receives from the operator with identifying encountered objects. The second aspect relates to the operator's possibility of selection of actions that should be done in response to the environment. The latter means more control in the hands of the operator and more required situational awareness. Furthermore, Walch et al. [36] present another cooperative driving system, CooperationCaptcha. This system also cooperates with the driver when encountering unrecognizable objects. The method was recognized as feasible and related to low cognitive load and high usability.

In our research, however, we will focus on teleoperation. Many challenges arise from teleoperation, mostly because the operator is physically disconnected from the vehicle. Tener and Lanir [33] conducted a study and identified six major categories of teleoperation challenges based on interviews and observations, namely lack of physical sensing, human cognition and perception, video and communication quality, remote interaction with humans, impaired visibility, and lack of sounds. In their work, they additionally presented certain UI suggestions to address the issues, e.g., adding special visual effects, projecting the future trajectory of the vehicle. Furthermore, several specific HMI design proposals have been made with the aim of addressing similar issues. For example, Naujoks et al. [27] discussed challenges related to teleoperation HMI design, and they presented a pilot evaluation of specific schematic takeover requests and warnings to be displayed in the teleoperator's cockpit. Kadavasal et al. [20] presented a mixed-reality based setup for teleoperation and Georg et al. [10] describe a demonstrator implementation for a VR-based teleoperation simulator setup in a high-fidelity simulation environment. Furthermore, Gafert et al. [9] presented a TeleOperationStation - a prototyping environment for testing automated vehicles operations, their challenges, and possible solutions.

In teleoperation contexts, the vehicle usually is not manually controlled at all times but performs automated driving as its standard mode of driving. When the vehicle is faced with a situation where the operator is needed, a takeover request is issued. Then, the system provides information to the operator [9]. A strong determinant for the user experience and acceptance of teleoperation are technological performance limitations, most importantly latency [28], but also video quality [12]. In order to design assistive elements to cope with the latencies inherent in teleoperation interfaces, displays that indicate the projected movements of the remote vehicle and the movements of other road users to forecast the vehicle position to overcome the time delay are used [13]. Graf et al. [13] presented and evaluated Predictive Corridor (PC) in their study. PC is a system meant to tackle the challenges of teleoperation, mostly focusing on time latency and remote situational awareness. Furthermore, PC was evaluated in a user study assessing cognitive workload, locus of attention and human performance. It proved to be successful in dealing with the posed challenges and led to enhanced performance of the teleoperation driver.

Recently, first requirements for teleoperation of automated vehicles have been proposed. Kettwich and Dreßler [21] provide a task classification for remote operation scenarios in the public transportation domain, based on conceptual task analysis and card sorting. Graf and Hussmann [14] provide a comprehensive collection of user requirements for teleoperation, based on expert consultations. We gathered and preliminarily grouped into 12 categories, namely, vehicle position, status, issues, characteristics, and operations, task objectives, on-board sensor, communication state, objects and obstacles, environmental information, weather conditions and terrain features. Each of the categories also contains subcategories.

The presented requirements include the provision of information about a vehicle's position, status and characteristics and issues. Furthermore, the requirements comprise information on tasks to be fulfilled by the operator (e.g., the description and remaining time for the tasks), as well as assistance for operations (e.g., potential collisions or projected vehicle position, events and obstacles in the environment, or other road users).

The authors performed a preliminary expert-based prioritization of these requirements, and in subsequent work they present a design space, related to their research [15]. In their work, Graf et al. [15] consider the influence of various factors on remote operation, namely the user, context, interaction, and technology. However, so far, concrete design proposals referencing to a framework of requirements are not available. While first evaluation studies with interactive prototypes have been presented, realistic application scenarios have yet to be addressed. [1, 11, 22].

Despite these advances in rigorous evaluations, the amount of such empirical accounts is still insufficient, and therefore a comprehensive validated set of design recommendations for automated vehicle teleoperation is not yet available. Graf et al. [15] mention that the field of design of teleoperation user interfaces is still evolving and further investigation needs to be done.

3 Requirements-driven design of an XR-based teleoperation user interface

In order to address the identified gap of many existing teleoperation requirements yet few empirically verified ones, we decided to

(a) implement a teleoperation interface that covers a wide range of existing requirements, and
(b) evaluate the interface and provide empirical grounding for its respective requirements.

To this end, we developed a teleoperation research environment that could be used for all application domains of remote automated vehicle management, but which was contextualized for this user study in the area of automated transport logistics (e.g. hub-to-hub transport or intralogistics). In order to validate the above mentioned requirements (see [14]), we consolidated these into a list of 54 items (not considering sub-categories of these requirements, such as vehicle size, length, width or height). This list was then presented to 10 experts in the field of automated transport logistics, as part of a technical workshop in an international research project. The overall relevance of the requirements was confirmed by these experts and further prioritized via a rating for each feature. Rating, inclusion, and exclusion depended on importance for regular and emergency operations, feasibility, and relevance within application contexts (niche vs. broad application expectancy). The end results was a selection of 36 requirements, which are listed in Table 1. For these, we created prototypical designs (Section 3.1) and then implemented them within an XR simulation framework (Section 4.1) in order to enable their evaluation in a user study presented in Section 4.2.

3.1 User interface design

The user interface elements discussed in this paper mostly satisfy the requirements presented in [14]. These requirements are used as a basis for the developed UI elements. In order to enable efficient comparison of interface features, we decided – instead of realizing and testing each requirement as an individual UI element – to instead develop two versions of a general control interface: a **minimal interface** and a **maximal interface** version (**MIN** and **MAX**). MIN is defined to contain only the necessary and/or already established standard information elements and indicators required to perform

Table 1 This table shows the mapping between proposed interfaces as shown in Figs. 2 and 3 to individual UI elements evaluated by the participants shown in Fig. 8

Category	Interface	UI Element	Requirement	Variant
Basic	Camera streams (surrounding user)	Four live camera feeds	360° remote view	MIN
			Camera state	MIN
			Camera orientation	MIN
			Vehicle heading	MIN
	Navigation system	Navigation system	Vehicle location	MIN
	Real steering wheel	Real steering wheel	Steering wheel orientation	MIN
	Speedometer	Speedometer	Vehicle speed	MIN
Status and controls	Vehicle status and controls	Charge level	Charge level	MIN
		Light controls	Light status	MIN
		Vehicle status	Tire damage	MAX
			Light damage	MAX
			Tire pressure	MAX
			Last vehicle inspection	MAX
			Overall vehicle damage	MAX
Vehicle information	Blueprint	Blueprint	Vehicle size (length, height, weight)	MAX
	Loaded cargo	Loaded cargo	Vehicle weight	MAX
Environment	Proximity indicator	Proximity indicator	Distance to obstacle	MAX
	Event list, Notification	Event list	Past/Current/Future vehicle actions	MAX
			Past/Current/Future vehicle issues	MAX
			Other sensors status functionality	MAX
			Object location	MAX
	Event list, Notification	Event list (collision information)	Potential for collision	MAX
	Event list, Notification	Event list (road users collision)	Projected other road user trajectory	MAX
Assisted driving	Projected driving lane (embedded in front camera stream)	Projected driving lane	Wheels orientation	MAX
	Projected vehicle position (embedded in front camera stream)	Vehicle position	Projected vehicle position w/ communication latency	MAX
	Projected vehicle stop position (embedded in front camera stream)	Stop position	Projected vehicle stop position w/ communication latency	MAX
Limitations	Connection status	Connection status	Bandwidth available	MAX
			Bandwidth requirements	MAX
			Mobile carrier (current, available)	MAX
	Weather	Weather	Weather (current, predicted)	MAX
Task information	Task information	Task information	Time constraints	MIN
			Tasks to complete	MIN
			Status of tasks/progress	MAX
			Task impact on plan	MAX
			Projected time to task completion	MAX
			Projected ability to complete the task	MAX

Furthermore, the table shows the 36 selected requirements by Graf and Hussmann [14] listed in the "Requirement" column. Additionally, the variant describes if the given UI element was used in the MAX or MIN interface. MAX includes all elements from MIN. The category is the basis for the structure in the results section

Fig. 1 A user in the XR teleoperation environment on the left remote controlling the miniature vehicle on the right in a miniature physical environment

regular driving tasks. MAX contains all 36 of the prioritized and selected requirements in addition to those already present in MIN. The requirements are listed in Table 1 in the "Requirement" column. Each requirement is implemented via a corresponding unique UI element. Fulfilment of requirements can, therefore, be mapped in a 1:1-manner to the interaction performance of the individual UI elements. The UI elements were combined and presented to the user in the form of interface panels (see Figs. 1 and 2). To provide an example: The "Vehicle status" panel consists of multiple UI elements, such as charge level, light controls, etc. All interfaces are shown to the operator at the same time and placed in the virtual environment as floating interfaces. Some of them have dynamic information depending on the task which changes over time (Task information, Navigation system, etc.), some are interactable via hand-tracking and virtual buttons to enable interaction (only used to press the virtual "Start" button and the virtual "Complete task" button shown in Fig. 5 as triggers) and some have randomized information like the status of the vehicle to simulate problems.

To structure the interfaces they are categorized into the following list which is also the basis for the structure of the results section.

Basic Basic UI features include the four camera feeds, the steering wheel, navigation system and speedometer. These elements are shown in both interface variants.

Status and controls Status of the vehicle including state of charge level, vehicle status, and light controls. This interface has two variants with an extended vehicle status present only in the maximum UI variant.

Vehicle information The vehicle information includes static information for a given remote operation task. A blueprint and the loaded cargo of the vehicle should help operators to get a better feeling for the size, weight and the resulting effects on steering and handling. The size, weight and type

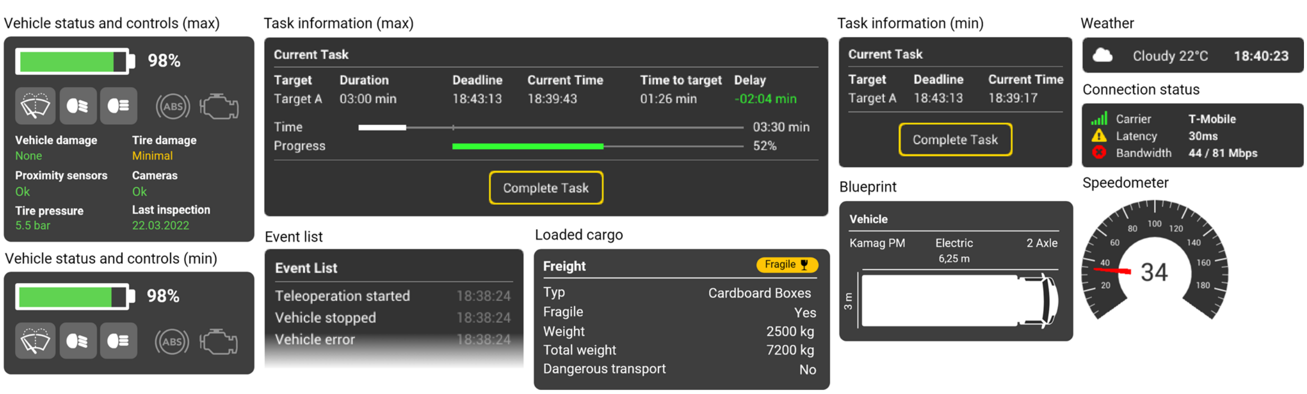


Fig. 2 UI elements of the MIN and MAX interface. Vehicle status and controls and task information have one version per variant as they are featured in both interfaces, with a reduced version in MIN. Not in the figure are overlayed elements like proximity indicators, projected driving lane with projected stop and vehicle position line and notification

of cargo is randomized and shows different information at the start of each task.

Environment Elements which have information about the environment include: Three proximity indicators showing the distance between vehicle and other objects, which change size and color. Additionally, three-dimensional sound cues are added as beeps depending on the distance. Information about potential collisions and information about road users are shown in the event list and displayed as a notification in the center of the screen.

Assisted driving A visualization of the driving lane depending on the angle of the wheels and the steering wheel is overlayed as an augmented interface over the front camera stream. Along with a line showing the projected vehicle position with communication latency and another line showing the projected vehicle stop position with communication latency.

Limitations Limiting interfaces include the connection status which has information about carrier, latency, current bandwidth, and available bandwidth.

Task information Includes information about task to complete, time constraints (deadline), progress of task (percentage), projected time to task completion, projected ability to complete task (delay). Task progress and time constraint are further portrayed as progress bars. The task progress bars can start from negative values in order to enable buffer/accommodation times between takeover and navigation start (for the eventual study, this was set to -20s). A button named "Complete task" is interactable and can be clicked by the operator once he/she deems the driving task as completed. This interface has two variants as basic information also needed to be provided to the minimal interface variant described in section.

4 User study

Our experimental study systematically evaluated the above-described user interface implementations of Graf and Hussmann's requirements [14]. The following **research questions for the study** were formulated:

RQ1 Which of the defined requirements and resulting user interface indicators need to be displayed for effective remote operation?
  (a) Before driving (orientation phase)
  (b) While driving (navigation phase)

RQ2 In which ways are the user interface features impacting the experience and performance regarding an easy and a challenging remote operation task?

4.1 Experimental system setup

The experimental system setup consisted of an extended reality headset (Varjo XR-3 [35]), a physical steering wheel (Fanatec Clubsport Wheelbase 2.5V [6] together with Fanatec Clubsport Steering Wheel Round 1 V2 [5]), physical gas pedal and brake (Clubsport Pedals V3 [4]), a custom VR environment, a miniature vehicle and a miniature physical environment. The miniature vehicle was driven remotely in the physical environment by an operator wearing an extended reality headset. The operator was physically in another room and drove the vehicle remotely from his/her virtual teleoperation station while wearing the XR headset. Four camera feeds were streamed to the operator and driving commands issued by the operator via a physical steering wheel and pedals were sent to the vehicle. The surrounding XR environment can be seen in Fig. 3.

4.1.1 XR environment

The XR-based interaction environment was built in Unity to allow for a more flexible UI element placement and development. UI elements can therefore be placed arbitrarily in the air as floating interface panels and do not need to be limited by the placement of physical touch-screens around the operator. Four virtual screens surrounded the user with one on top for the back camera (see Fig. 3). An extended reality headset was chosen to allow the operator to see the real steering wheel used for driving. Apart from the hands and steering wheel all elements were virtual. No other extended reality functionality was utilized.

4.1.2 Miniature vehicle

To allow operators to make errors while driving a physical miniature vehicle was custom built (see Fig. 4b). The vehicle is 21 cm long, and 13 cm wide. It is mostly 3D printed and controlled by a Raspberry Pi 4 [8]. A servo controls the steering of the front wheels and a Nema 8 8HS15-0604S stepper motor controlled by a TMC2209 stepper driver rotates the back wheels. The vehicle behaves like a truck and has a slow acceleration curve. However, when the acceleration pedal input is incorrect in a given scenario the vehicle does not stall like a normal car would. The commands are sent from Unity over WebSockets to the Raspberry PI. Four Arducam B0261 120° Field-of-View cameras are mounted on each side of the vehicle. Each camera streams a MJPEG stream to Unity over WiFi. The network setup has a latency of under 20ms with averaging 28 frames-per-second. At the front of the vehicle three Time-of-Flight sensors are mounted to measure the proximity of close objects. The distance is displayed in the proximity indicator element shown to the operator.

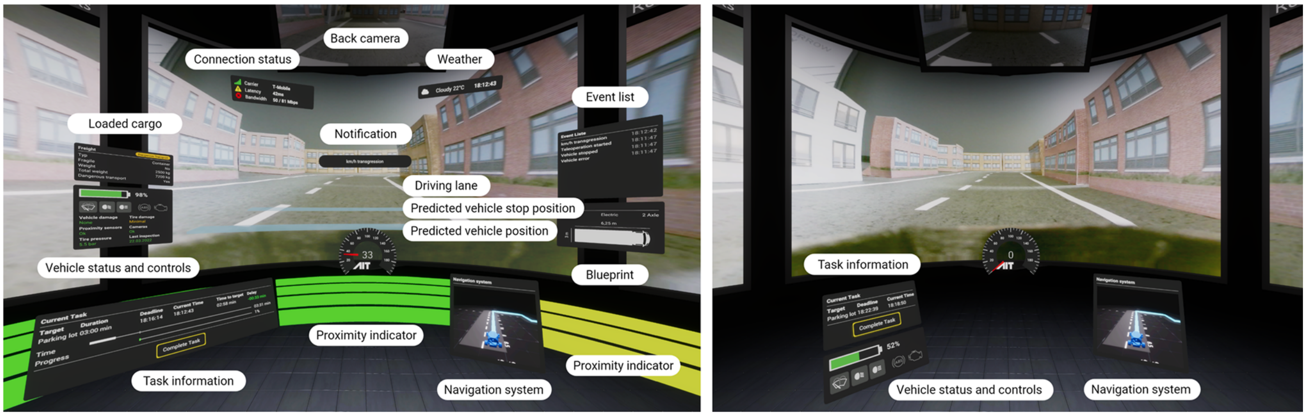
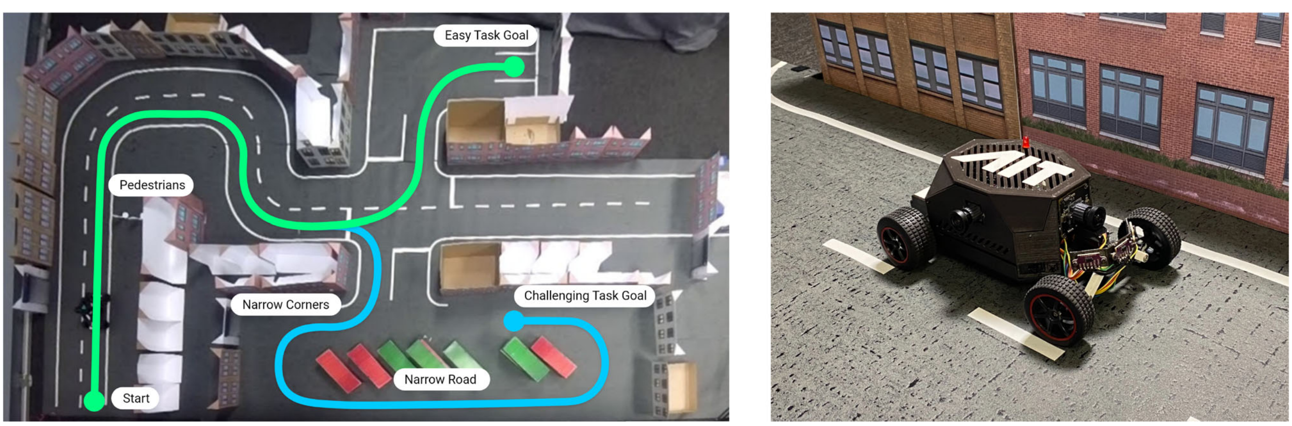


Fig. 3 Teleoperation XR environment presented to the operator surrounded by four cameras streams and floating UI elements shown in Fig. 2. Left the MAX interface variant, right the MIN interface variant. Left and right cameras extend to surround the user on each side. The operator is positioned in front of the speedometer with the real steering wheel overlayed into the virtual environment

4.1.3 Physical environment

The physical environment consists of multiple paper cutouts representing different facades of buildings placed on the track (see Fig. 4a). The physical environment is highly modular and can represent multiple tracks. In addition to house facades, street signs, and road markings, pedestrians were also added at the side of the road. For this study one start position and two goal positions are chosen. The start position can be seen in Fig. 4a in the bottom left corner. We defined two different track types in order to reflect different levels of difficulty. While it is not feasible to cover the full range of potential intervention causes and actions that need to be performed post-takeover, we can introduce different levels of difficulty for similar task categories to then identify whether a requirement is of different importance depending on how involved the intervention is. For the purpose of this study, we decided to focus on navigational tasks, as it can be expected that intervention after a standstill will often entail (at least partial) manual navigation to a safe position with error or drivability state diagnostic beforehand. In addition, restrictions imposed by unfamiliar physical environments in combination with vehicle size are particular to the remote operation context, where one operator may oversee multiple physical environments and vehicles.

The first track and corresponding task is defined as *easy* and consists of three wide right turns and two wide left turns with one intersection. The second task is defined as *challenging* and starts at the same starting position as the easy task but involves two wide right turns, one wide left turn, two



(a) Track used in the study with one start position and two defined goals. The blue line shows the easy track and the green the challenging track. Noteworthy elements are the pedestrians and narrow road which trigger a notification in the interface.

(b) The vehicle which is operated remotely with four cameras and three proximity sensors.

Fig. 4 Physical study setup with environment built with paper cutouts and the 3D printed custom vehicle narrow right turns, and four narrow left turns. Furthermore, the operator is required to reverse in the container area as the vehicle's turning radius is too big. The easy difficulty track is estimated to take about one minute to navigate if driven carefully and according to the speed limit. The challenging difficulty task is estimated to take about three minutes to navigate. The exact timing is presented in the results section of the navigation phase (Section 5.2.1).

4.2 Experimental design, sample and measures

We conducted the study as a within-subjects **experimental design** (N=16). We compared the two versions of the UI (MIN and MAX) in two different driving environments (Easy and Challenging), which resulted in four conditions for comparison as shown in Table 2. There was no additional control condition without an UI, as controlling a vehicle without basic indicators is not reflected in any relevant application context. The participants were recruited via an online platform in Austria, Europe. The **experimental sample** comprised 16 subjects with a mean age of 36.8 years (SD=8.9), 10 male and 6 female. All participants had a driver's license, but did not have previous experience with teleoperation. Subjects received a monetary incentive. We screened for susceptibility to motion sickness during recruitment (via self-reporting). In addition, each participant was asked for motion sickness symptoms at key points during the trial (i.e., after each round of interaction). Regards **experimental measures**, we used a mix of qualitative methods (questionnaires and interviews) and objective performance metrics in order to ensure validity and explainability of the results.

We used a self-created demographics questionnaire, the NASA-TLX [26], and the UEQ+ [31]. The NASA-TLX was utilized to verify the task load for the *Easy* and *Challenging* task and to understand if the MIN and MAX variant had any impact on the task load. The UEQ+ was used to verify whether the UI elements impacted the operator's performance or had other impacts, which could skew the results. The used scales were: Efficiency, Perspicuity, Usefulness, Intuitive Use, Trustworthiness of Content, Quality of Content and Clarity. At the end, participants were asked to rate the usefulness of the elements in the MAX UI, both for the orientation and navigation phases (1 = low usefulness; 5 = high usefulness). We used two types of interviews: short, structured interviews that were conducted after each interaction round and one semi-structured interview across all rounds of interaction at the end. The objective measures employed in the study were eye tracking (dwell times and dwell counts) to capture participants' attention towards specific elements in the UI during the orientation phase. The dwell data was generated by defining areas of interest in iMotions [17]. In addition to eye tracking, we also captured task completion time to capture the relative dwell time in relation to the overall time it took the participant to orient themselves and furthermore get objective measures on the duration of the navigation phase when comparing the interface variants.

4.3 Task

The task of the operator was to take over the vehicle in the case of an emergency. The emergency was initiated by the study conductor vocally and always meant stopping of the automated vehicle. The operator had to put on the VR headset once notified of the emergency, identify the problem which induced the emergency and then drive the vehicle remotely to a position marked on the navigation system. The speed limit was 30 km/h (in relation to vehicles size which was 1:13 which resulted in an absolute speed of 0.6 m/s). If the speed limit was exceeded a warning was shown in the event list and notification interface. The problem which triggered the emergency was shown in the vehicle status interface and event list and was either an engine control light, vehicle damage,

Table 2 Experimental design setup. Each participant drove four rounds in total: two with the MIN interface, which includes 11 requirements (see Table 1), and two rounds with the MAX interface, which includes 36 requirements

Round	Interface	No. of included requirements	Difficulty	Route elements
1	MIN	11	Easy (1 min)	6 turns (w)
2	MAX	36	Challenging (3 min)	3 turns (w), 6 turns (n), 1 reversal
3	MIN	11	Challenging (3 min)	3 turns (w), 6 turns (n), 1 reversal
4	MAX	36	Easy (1 min)	6 turns (w)

The *Easy* track was estimated to take about one minute and the *Challenging* track around three minutes to navigate. The order of the interface variant and difficulty was randomized for each participant. The route elements describe the obstacles the participant has to navigate trough on the track (see Fig. 4a). (w) = wide, (n) = narrow

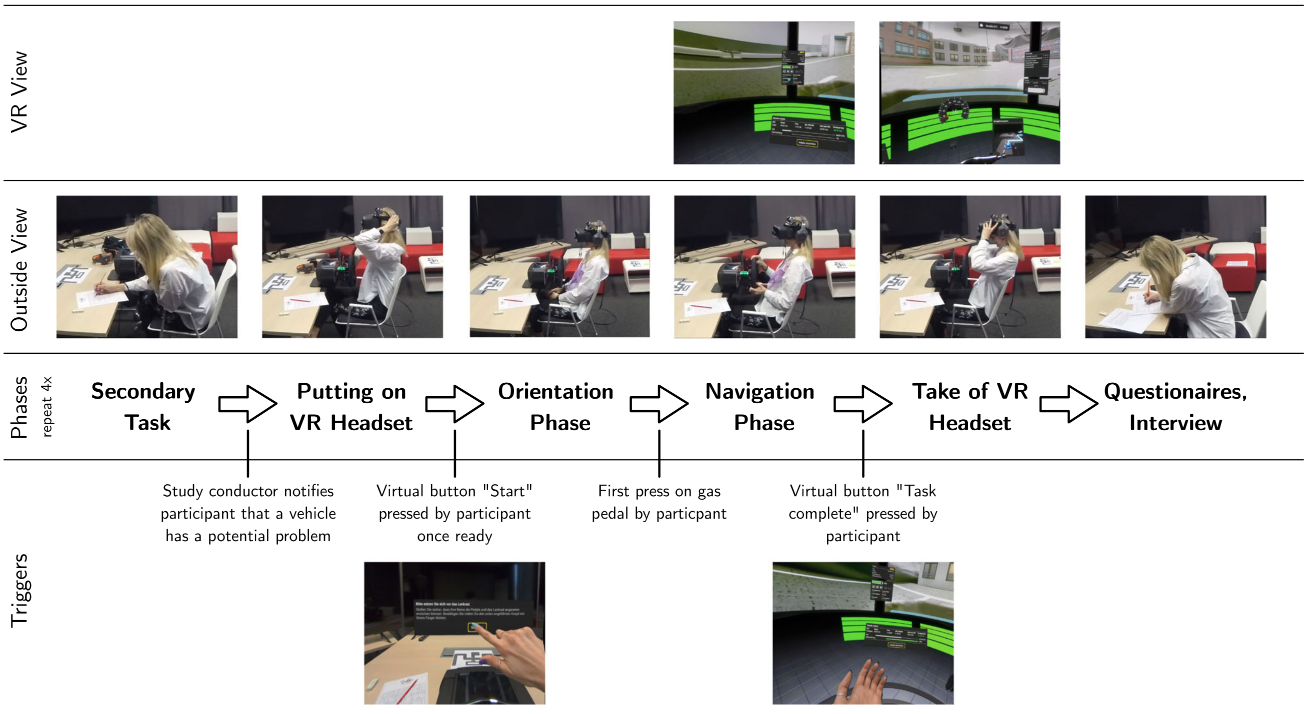


Fig. 5 The procedure of one round starting with a secondary task (we chose Sudoku), the participant entering the virtual cockpit by putting on the Varjo XR-3 headset and starting the scene with a virtual button, orientating in the virtual world, navigating the vehicle and exiting with another virtual button. At the end of each round the UEQ+ and NASA-TLX was handed our and a short interview was held. The next round started without a break with the secondary task

tire damage, proximity or camera failure, too low tire pressure or an overdue inspection. The problems were chosen automatically at random by the system. The detailed problem description was only visible in the MAX interface (see Fig. 2 “Vehicle status and controls” for MIN and MAX).

4.4 Procedure

The whole study procedure took approximately 90 minutes per participant. After a welcome phase (10 minutes), which consisted of the briefing, procedure and data collection information, and signing the informed consent, the familiarization phase followed (10 minutes), in which participants drove a trial round with the vehicle, using the MIN interface variant. This phase was kept brief in order to enable interface familiarity during the trials but minimize potential learning effects.

In the experimental phase (60 minutes), each participant drove 4 rounds. The procedure for one round is shown in Fig. 5. Each round began with the participant solving Sudoku puzzles for approximately five minutes, to model the distraction of other tasks not related to the teleoperation of the vehicle. The participant was then notified by the study conductor that a vehicle has a potential problem and that they need to put on the XR headset to investigate it. Once the participant put on the headset they could start the **orientation phase** with the click of a virtual button. By pressing the button the real world darkened and the virtual world as shown in Fig. 3 became visible. Depending on the current variant either the MAX or MIN interface was visible. Before beginning the driving task, the participant was asked to get an overview of the situation regarding the vehicle, its state and its surroundings and to verbally report their assessment once they had achieved it (orientation phase). The first push of the gas pedal then marked the start of the **navigation phase** in which the participant had to follow the navigation system and navigate around the miniature environment. Once the track was cleared the participant had to stop the navigation phase with another virtual button which completes the task. After each round a NASA-TLX and UEQ+ questionnaire was filled out by the participants and the cycle continued by continuing the Sudoku.

After the experimental phase (all four rounds), the participant filled a final weighted NASA-TLX and demographics and the final interview was conducted with the rating of the individual UI elements’ importance as listed in Table 1. This final phase had a duration of 15 minutes.

5 Results

The results presentation is structured into two main parts: the orientation phase (Section 5.1) and the navigation phase (Section 5.2). For each of these two phases, we provide an overview and we report on specific results related to the teleoperation HMI features, the behavioral data (eyetracking and task completion time for MIN and MAX), as well as the questionnaire and interview results (NASA-TLS and UEQ+). For statistical comparisons, we ran a general linear model repeated measures analysis of variance with SPSS to analyze main and interaction effects, as well as to derive pairwise differences (based on Bonferroni-adjusted p-values). We treated the questionnaire scales as interval scales [16], [30]. The eye-tracking data was recorded with the Varjo XR-3 headset in Varjo Base [34] and imported into iMotions [18] to get individual and aggregated gaze data. We analyzed the aggregated gaze data and used the dwell time and dwell count of each UI element on which we defined areas of interest. Additionally, we used the mean fixation order evaluated by iMotions to understand which elements have been noticed first and which have been focused on last before starting the driving task. Only one participant reported motion sickness during the interaction, which resulted in cancellation of that trial altogether and excluding the partial data collected from that trial form the analysis. All other 16 participants did not report or show any signs of motion sickness during any part of the study. We, therefore, exclude potential effects of motion sickness on performance or quality of interaction.

5.1 Results for the orientation phase

The subsequent section describes the results of the orientation phase staring with the general eye tracking results and then describing each unique UI element grouped by the category shown in Fig. 8. This phase started as soon as the participant pressed the virtual button to enter the teleportation environment as shown in Fig. 3 and stopped with the first press on the gas pedal. As the usefulness questionnaire for all UI elements (MAX) was filled out at the end of the task, the results of this section are not separating MIN and MAX, but describe the usefulness of every UI element individually. The results also do not separate *Easy* and *Challenging* difficulty as these are only related to the difficulty of the track and therefore relevant in the navigation phase.

5.1.1 Eye tracking

As the operators' task was to identify the potential problem we used eye tracking to analyze which UI elements the operator fixated on during this phase. As we wanted to see the difference for all UI elements the results focus on the MAX variant as MIN had much less interfaces. The numeric dwell time and dwell count results are shown in Fig. 6, an additional heatmap can be seen in Fig. 7.

The first element fixated is **loaded cargo**, with an average dwell time of 4 seconds, followed by **vehicle status**, with an average dwell time of 6.4 seconds which is the highest dwell time of any UI element. The UI elements on the left side of the

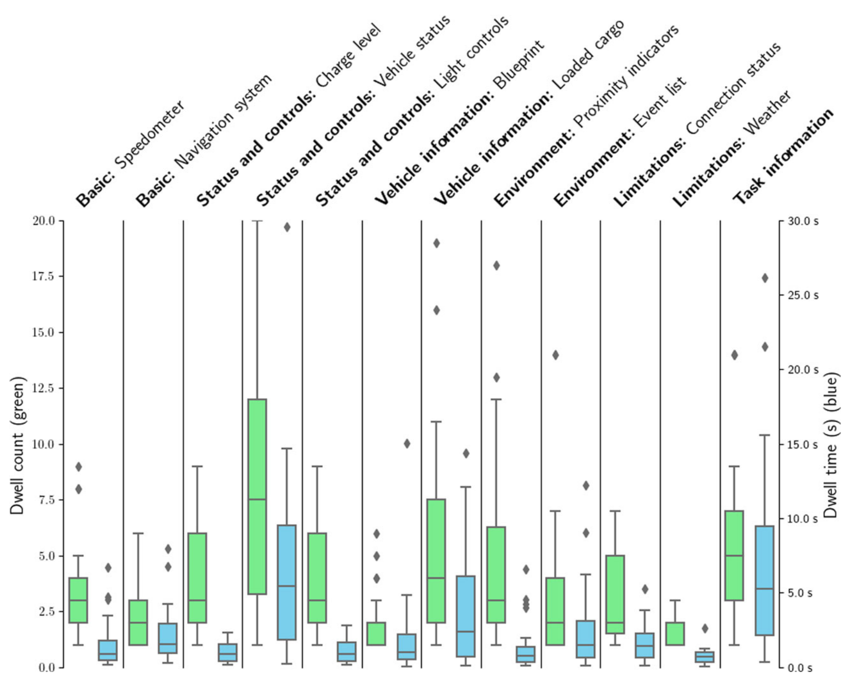


Fig. 6 Dwell count (green) and dwell time in seconds (blue) of each unique UI element visible in the orientation phase. Only UI elements visible in the orientation phase from Table 1 are shown

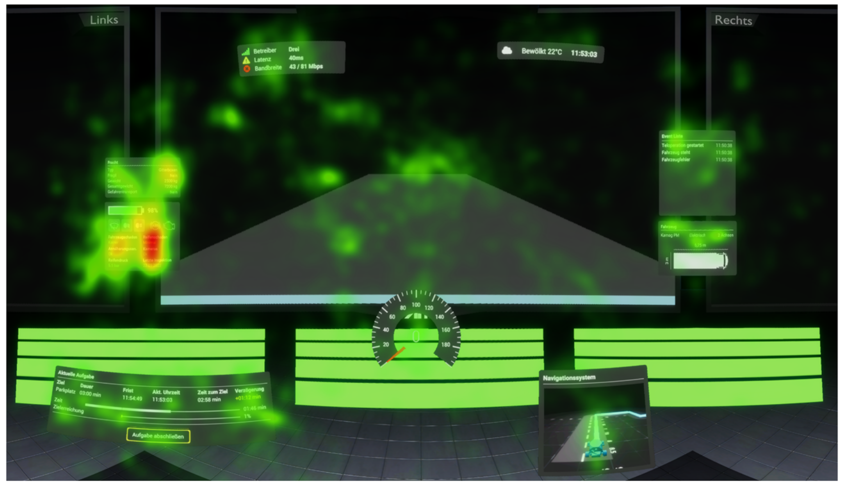


Fig. 7 Fixation count heatmap of the orientation phase. Red areas suggest a higher level of interest

steering wheel are fixated on for longer and sooner compared to elements on the right side of the screen. This correlates with the answers from the interviews, where participants stated that the most relevant information in the orientation phase is the vehicle's status and its potential problems along with the cargo which influences the driving behavior. According to the fixation sequence the last elements to be fixated are **weather** ($\bar{x}$ dwell time = 0.7s) and the **navigation system** ($\bar{x}$ dwell time = 2.2s). The full sequence (starting with the first element to be fixated on) of UI elements ordered by their average time to first fixation is: (1) Basic: Loaded cargo, (2) Status and controls: Vehicle status, (3) Basic: Speedometer, (4) Status and controls: Charge Level, (5) Task information, (6) Status and controls: Light controls, (7) Limitations: Connection status, (8) Environment: Event list, (9) Vehicle information: Blueprint, (10) Environment: Proximity indicators, (11) Limitations: Weather and (12) Basic: Navigation System in the last place.

5.1.2 Basic

The subsequent sections describe the results of each interface category starting with Basic. The results are also shown in Fig. 8. The basic features from the MIN condition received comparatively high usefulness ratings in our questionnaire: the **camera feeds** ($\bar{x}$ = 4.2), the **real steering wheel** ($\bar{x}$ = 3.5), the **camera status and orientation** ($\bar{x}$ = 3.6) and the **navigation system** ($\bar{x}$ = 3.7) which was regarded as one of the most important UI elements to see before the driving task to gain a general understanding of the current position. However, as the vehicle was not moving, it was generally not looked at for long in this phase. Contrary to the other features, **speedometer** received lower usefulness ratings ($\bar{x}$ = 2.2) for the orientation phase, as the car would then usually not be in movement. Likewise, the gaze map shows relatively low dwell times.

5.1.3 Status and controls

Many status and control information features were rated as useful, especially the **vehicle status** ($\bar{x}$ = 3.8), but also the **charge level** ($\bar{x}$ = 3.8). Participants assessed the status of the vehicle by checking for damages and charge level, which also resulted in the highest dwell time of any UI element with 6.4 seconds for the vehicle status UI element. The **light controls** ($\bar{x}$ = 2.7) received usefulness ratings lower than 3 and their relevance was considered limited for completing the test task.

5.1.4 Vehicle information

Participants reported that information about **loaded cargo** with respect to fragility of the cargo impacted the driving style. The usefulness rating was slightly above average ($\bar{x}$ = 3.3), while a high gaze dwell time was observed in contrast to other design elements. Also, the **blueprint** was rated with medium usefulness (($\bar{x}$ = 2.9), since participants considered it to be interesting information but not meaningful to their task.

5.1.5 Environment

Information about road users received higher ratings of medium usefulness ($\bar{x}$ = 2.9), but the other design elements related to the environment were consistently rated as having lower than medium usefulness: **Proximity indicators** ($\bar{x}$ = 2.25), **collision information** ($\bar{x}$ = 2.3), **event list** ($\bar{x}$ = 2.4375). The results were reflected in the interviews, stating that before the driving task started, the vehicle was is

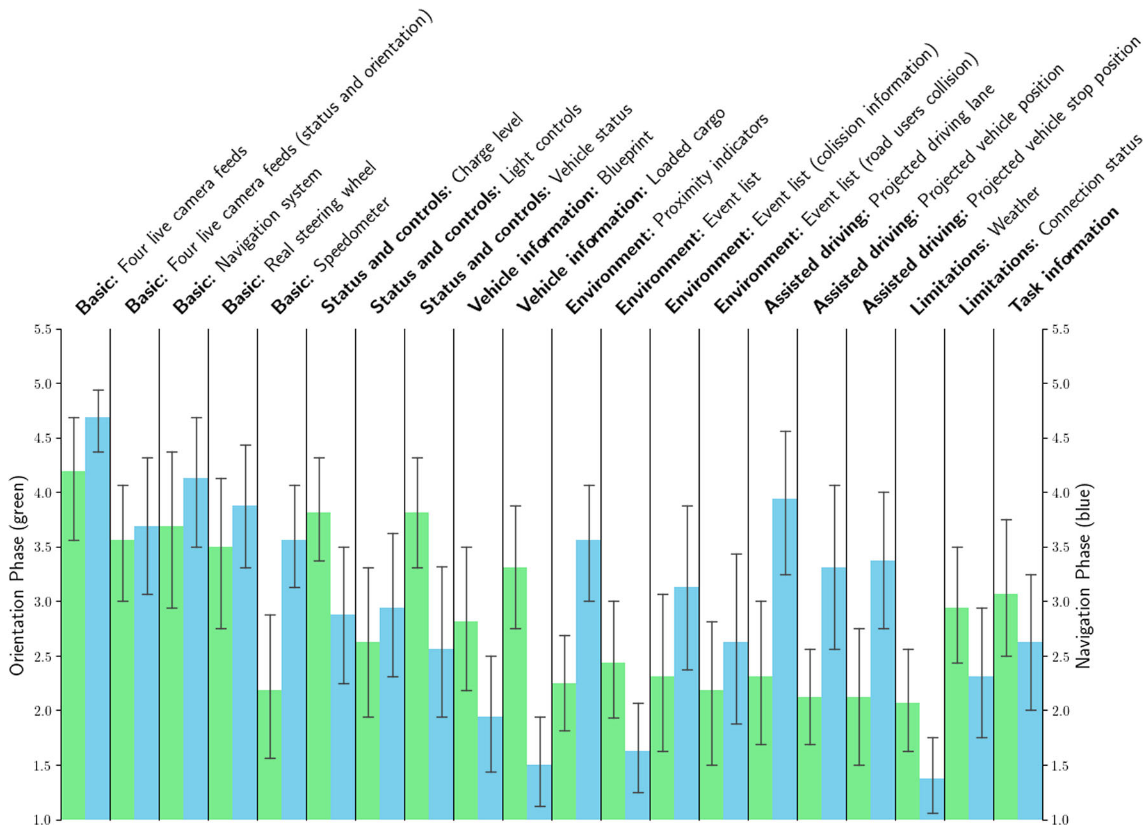


Fig. 8 Mean usefulness scores from low usefulness (1) to high usefulness (5) for each UI element in the orientation phase (green) and the navigation phase (blue). Each element is structured into the categories described in Section 3.1. Error bars represent 5% confidence intervals

in a static and safe environment, which therefore would not require any short-term environment information.

5.1.6 Assisted driving

The **driving lane** ($\bar{x}$ = 3.7) helped participants to estimate the wheel angle better than the angle of the physical steering wheel in front of the operator. Feedback about the **vehicle position and stop position indicator lines** (both $\bar{x}$ = 2.1) were not relevant before the vehicle was in motion as they could also not be seen.

5.1.7 Limitations

The usefulness of the **connection status** ($\bar{x}$ = 1.9) was rated low, since participants were not aware of a camera lag before starting the drive. Details about the **weather situation** ($\bar{x}$ = 2.1) were not regarded as a meaningful feature in the given test scenarios, but it was seen as relevant for other situations. Weather was also the least important element according to the gaze map with an average dwell count of 0.7 seconds and a dwell count of 1.4 seconds.

5.1.8 Task information

Participants used the **task information** ($\bar{x}$ = 3.06) to decide for how long they could inspect the UI and vehicle status before starting the drive, based on the available time. The task information had the highest dwell time of seven seconds. However, it was also often conceived as overwhelming as it offered many different values. The most important sub-element of the task information was the delay shown on the top-right side of the element. Once participants saw the delay they perceived more stress according to the interview and started driving, ignoring the other interfaces, apart from vehicle status.

5.2 Results for the navigation phase

This section describes the results of the navigation phase which started once the gas pedal was pressed and ended when the operator clicked the "Task complete" virtual button in the task description interface. The results are structured into task completion time, UEQ+, NASA-TLX - which describe the differences between MIN and MAX interface and *Easy* and *Challenging* difficulty - and usefulness on a per element basis structured by their category. The usefulness results of each UI element during the navigation phase, are as above, not separated in MIN and MAX as only MAX showed all elements.

5.2.1 Task completion time

For the analysis of task accomplishment time, one subject had to be excluded from the analysis, due to an aborted condition due to technical problems, resulting in a sample size of N=15. The overall mean duration of driving the tracks was 201 seconds. We found a main effect of feature richness, where the condition with all design elements (MAX) was $\bar{x}$ = 230s, and the one with only the basic feature set (MIN) was $\bar{x}$ = 175s, F(1,14)= 7.6, p<.05. As expected, we found that the accomplishment time was longer for the challenging tasks than for the simpler tasks ($\bar{x}$ = 268.8 vs. $\bar{x}$ = 136.0). This difference was significant F(1.14)=32.4, p<.01. We did not find a significant interaction between feature richness and difficulty level.

5.2.2 UEQ+

The UEQ results show that the overall mean of all remote operation conditions was on the positive side, ranging (on a scale from 1-7) between 5.6 for Intuitiveness and 5.0 for Efficiency (with standard deviation around 1.3). We did not find any main or interaction effects of the factors feature richness and difficulty on an error probability level of p<0.5. The largest indicated differences and the lowest error probabilities were found for Usefulness and Trustworthiness (with slightly higher values for the rich feature set as compared to the basic feature set ($\bar{x}$ = 5.6 vs. $\bar{x}$ = 5.0, p=.096; $\bar{x}$ = 5.4 vs. $\bar{x}$ = 4.8, p=.13).

5.2.3 NASA-TLX

For the analysis of the NASA-TLX, scales were weighted, and an average was created based on the weighted scales (i.e. also here a scale of 1 to 20). The overall mean of all scales was 7.1 (SD=3.1), indicating a low to medium workload. Feature richness did not have an effect on workload, but the difficulty level had, $\bar{x}$ = 7.5 vs $\bar{x}$ = 6.8, F(1.15)=16.3, p<.01). No interaction effect of these two factors was found. When looking at the design elements, the **cameras** ($\bar{x}$ = 4.7) received a high rating, since they were regarded essential for the orientation and navigation in the space. In the MIN-interface, participants relied strongly on the video to estimate distances to objects and to predict the driving lane.

5.2.4 Basic

The subsequent sections describe the results of each interface category starting with Basic. The **real steering wheel** ($\bar{x}$ = 3.9) received a high usefulness rating (it was mostly mentioned as a challenge when turns took longer than expected). The **navigation system** ($\bar{x}$ = 4.1) turned out to be a key element for completing the tasks. It could even compensate for perceived camera lags by allowing operators to track their vehicle position on the map, given that participants perceived the system as trustworthy. The above-medium rating of the **speedometer** ($\bar{x}$ = 3.6) can be explained by this participant's comment: "These are the most important things and thus need to be supported: how fast am I going and where do I need to go?". **Camera status and orientation** ($\bar{x}$ = 3.7) were rated high, and the rear view mirror was perceived as accurate and useful.

5.2.5 Status and controls

The **charge level** and **vehicle status** design elements were rated low for usefulness during the drive ($\bar{x}$ = 2.9 and $\bar{x}$ = 2.6), since the operator had already considered these factors in the orientation phase (where they were given higher rating scores). In addition, the elements were static and participants did not need to be updated of them while driving. This lack of change could explain why **light controls** ($\bar{x}$ = 2.9) were rated low (light was considered to be unnecessary by three users).

5.2.6 Vehicle information

Participants did not regard ascribe high usefulness to the provision of **loaded cargo** and the **blueprint** of vehicle measurements ($\bar{x}$ = 1.9 and $\bar{x}$ = 1.5) during the whole drive, since this information never would never change during the task.

5.2.7 Environment

The **proximity indicators** were received well ($\bar{x}$ = 3.6), since they helped participants to learn about distances while driving the unfamiliar vehicle. Participants reported that **notifications about road users** ($\bar{x}$ = 2.6) restricted their field of vision.

5.2.8 Assisted driving

The **driving lane** ($\bar{x}$ = 3.9) was useful during the drive and helped participants to know about their pathway, even when there were camera delays. Participants relied on feedback about **vehicle position and stop position** ($\bar{x}$ = 3.4 and $\bar{x}$ = 3.3), even though a participant felt insecure about the reliability of the vehicle position when he perceived a camera delay.

5.2.9 Limitations

Providing information about the connection status ($\bar{x}$ = 2.3) was regarded as having little usefulness, since interpreting the numerical information was not intuitive to some. Participants did not deem the **weather indication** ($\bar{x}$ = 1.4) useful. A participant added that these details would be meaningful in bad weather conditions, where an operator would have to know how to adopt their driving behavior.

5.2.10 Task information

The **task information** ($\bar{x}$ = 2.6) was stressful to some participants, as it pressed them to arrive at the destination faster. Others reported that it was not useful, since they were trying to drive as reasonably and fast as possible anyway.

6 Discussion

As already mentioned in Section 2, Graf and Hussmann provide an extensive collection of user requirements for teleoperation [14] and subsequently present a related design space [15]. In our work we tried to implement some of the suggested requirements in an actual UI with the aim of researching how effective they are in practice. Moreover, we conducted our study because one of the main issues in our field of research is the lack of studies addressing realistic application scenarios (e.g., [1, 11, 22]).

In the following, we discuss the findings and answer the initially posed research questions. We will first take a look at the results regarding workload and task accomplishment time (RQ2), and then answer the primary question, i.e., which HMI elements are important during which phase of interaction (RQ1), on that basis.

6.1 On cognitive load and time to task completion

Before we can answer the primary question, which HMI elements are important during which phase of interaction (RQ1), we will first look at the results regarding workload and task accomplishment time (RQ2), as these will highlight an interesting finding regarding RQ1: regarding workload, the NASA-TLX showed a clear trend, where the Challenging condition was rated higher than the Easy condition. This validates the setup insofar as the Challenging condition did actually require greater effort to complete than the Easy condition. Surprisingly, the TLX revealed no differences between the MIN and MAX interfaces in terms of workload, which was contrary to our assumptions, as we had assumed the MAX interface to cause at least some difference due to all indicators being present at all times (information overload potential). In the interviews, some participants even mentioned that the MAX condition was perceived as more cumbersome and distracting, though these mentions were recorded from participants who had started with MIN and then changed to MAX in round two. Conversely, participants who had started with MAX and then switched to MIN in round two, sometimes reported to feel less supported or safe as a result of the loss of information. Thus, there seem to be subjective effects due to the switch of conditions but no objective effects on workload.

This does not mean, however, that MAX is automatically the ideal interface, as there was a significant difference in time needed for task completion during the navigation phase between MIN and MAX, where MAX caused longer task completion times than MIN. So while not necessarily causing higher workload, we can interpret the participants' statements regarding added confusion as lowered efficiency for MAX. There was one additional factor that the post-analysis revealed: the participants in the MIN condition were speeding more often than in the MAX condition. While both contained a standard speedometer, only MAX also had a permanent speed limit indicator. This led to participants speeding less in the MAX condition but also, as the interviews revealed, to navigate more by instruments and mostly looking at the navigation system and speed limit indicators than the environment through the camera feeds (which they did in the MIN condition). Not only does this create an interesting contrast between providing sufficient information for safe navigation vs. drawing attention away from the driving environment, it is also relevant for determining speed limits as a recommended HMI element, contrary to what the participants stated in the interviews, as will be seen in the following section.

6.2 Which information and when - orientation vs. navigation phase

Beside aiming to implement the work of Graf and Hussman [14] we wanted to get insight into which type of information is preferred in which context. Therefore, we compared the importance of various requirements in orientation and navigation phase of our experiment.

From both quantitative and qualitative results, we can now summarize which information and resulting HMI elements were deemed most important in each phase of interaction (RQ1). As one would expect, the standard elements of camera feeds, navigation system, and speedometer were deemed important and useful across both phases and should be part of any teleoperation HMI as a result. As far as the additional UI-elements in MAX are concerned, there are differences in importance per phase of interaction.

6.2.1 Orientation phase

In the orientation phase, we summarize the most essential information to be *orientation and proximity indication*, *lane guidance*, *error information*, *freight information*, *vehicle status*, and *vehicle dimensions* (blueprint).

Orientation and proximity information serve the general purpose of identifying one's own position in relation to the environment as well as obstacles that might not be in clear view but hinder navigation (and cause potential damage to the vehicle if ignored). Lane guidance serves a similar purpose, where the participants particularly mentioned the need to know the current wheel position and steering angle in order to safely navigate from the starting position.

Error information and vehicle status are essential to determine whether it is safe to begin navigating at all, as depending on the reason for the call to control and whether and/or for how long the vehicle could still be operated was a decision factor for the participants, even if the setup contained to (simulated) vehicle defects or damage. Cargo was important to calibrate the participants' driving stile in the beginning, so that neither the cargo (if sensitive) or the environment (if bulky or hazardous) would be endangered. Similarly, vehicle dimensions were necessary to get a feel for the vehicle's size and calibrate one's driving accordingly before resuming control.

6.2.2 Navigation phase

During navigation phase, we summarize *proximity information*, *lane guidance*, *speed limit information* as essential, with *connection status / latency information* as moderately important in addition. Once a remote operator has gotten their bearings, determined the task, vehicle, and contextual constraints, these aspects become secondary and the main emphasis is on supporting the navigation task *on the mechanical level*. Lane guidance and proximity indication are carried forward from the orientation phase, both serving to both understand and learn the vehicle's driving dynamics, which are (a) dissimilar in terms of feedback to a regular hands-on-wheel driving situation, and (b) can vary from vehicle to vehicle even within the same fleet (brand and type differences or age-related differences within the same vehicle type). We include connection status as moderately important during navigation instead of in the orientation phase. The reason for this is that the interviews showed that participants were often not sure if a given latency was a control latency (vehicle response delayed) or visual latency (camera visuals delayed), since we only showed a single latency value. As a result, the participants had difficulties calibrating their steering in accordance with the latency information, and we assume that this is the primary cause for latency information being rated high in navigation yet lower than in the orientation phase.

From the task completion times and the observed speeding in the MIN condition, keeping within the speed limit seems particularly tricky, as it is less noticeable while teleoperating and subsequently not expressed as a requirement. While it might seem trivial to conclude that showing the speed limit helps an operator staying within it, it is important to keep in mind that the usual way to operate a vehicle is to monitor the environment first, then the instruments, with speed limit information in the UI (such as it is often implemented in modern navigation systems) being a secondary resource. In the teleoperation context, especially during navigation, we see a strong argument to reverse this order of priority, as the (visual) context information and lack of haptic feedback (inertia) seem to make what is usually secondary information mandatory for safe operation.

6.3 On additional indicators and room for improvement

Finally, we want to discuss information that was identified as important but was not implemented (or not to a sufficient degree of detail) as well as aspects of the MIN or MAX UI elements that were not perceived to be implemented in an ideal way to derive further recommendations for design and improvement.

Regarding missing information, vehicle status was an initially derived requirement and was also identified as essential in the orientation phase. Here, however, participants mentioned to require more detailed status information, including predictive information, especially in cases of vehicle damage or malfunction. In such cases, an operator needs to know whether any incurred damage is severe enough so that it must be repaired before operation can be assumed or if the vehicle is damaged but can be operated. Even in the latter case, the operator must know for how long or under which conditions the vehicle can be operated (e.g., max./recommended distance, max. incline, steering radius if damage to the axis, etc.).

Environmental information, particularly weather information, was not rated as important in any of the quantitative assessments. It was, however, mentioned as relevant in many of the interviews. Weather conditions were not a variable in the study setup, the quantitative result is, therefore, not surprising. In a more realistic operation context, weather has a direct influence on contextual constraints and vehicle handling, so it is likely that it is either essential or at least highly important during the orientation phase. Still, since the data can not support this with the data collected, we do not put it forward as a requirement but do want to highlight it as a very likely candidate and subject for further investigation. Despite the eye tracking showing a high dwell time on task information during the navigation phase, we did not put it forward as an essential element for that phase. The reason for that is that the qualitative data showed these to be mainly caused by the UI element to be difficult to comprehend, leading to participants glancing over often and for prolonged periods of time in order to understand what it meant and how it changed in relation to their actions as they drove. While we can not generalize this for all possible contexts and tasks, at least for the navigation tasks investigated in this study, we saw standard navigation info and time indication to be sufficient with additional task information to not be necessary, and thus do not propose it as a general requirement.

Lastly, there was the simple matter of positioning. While the UI elements were clustered logically, they were also spread out to not obstruct the driving view. The exploratory nature of the setup meant that if individual indicators that were not close to each other turned out to be more important than others (e.g., connection status and blueprint), then this would cause participants to need to rapidly focus on distant areas of the UI and diffusion of attention as a result. This is largely a side-effect of the setup and can, in a fully developed application, be solved by simply using two element clusters - one for the orientation and one for the navigation phase - and then only showing the one that is relevant for the phase the operator is currently in.

6.4 Limitations

Apart from the unified UI element clustering and no comparison between different element clustering positions due to the study focus on information elements rather than position, there are additional limitations to the approach described in this paper. In terms of the interaction environment, opting for a physical driving environment meant that we were limited to the available lab space, resulting in a driving environment with limits to its size and complexity. The physical properties of the vehicle also limited its steering angle to 45°, which meant that corners had to be modelled larger in order to accommodate this comparably wide steering angle. These limitations resulted in the eventual design of shorter courses with the maximum amount of corners that could be included while still having two distinct courses with (a) different levels of difficulty and (b) different end points, which (c) could not be seen from the starting position.

The course was designed to feature only static challenges and obstacles. Dynamic events, such as pedestrians walking alongside the road, pedestrians crossing the road, cargo falling onto the road, or other moving vehicles, were not within the scope of the study and, therefore, not modeled in the setup. We should note that the inclusion of dynamic events would introduce significant additional effort and would, in terms of cost vs. effects likely be beneficial for scenarios with low to medium numbers of such events. Highly complex scenarios with a large number of dynamic events and/or actors are likely to still be better suited for simulation full VR.

We chose to pursue a within subjects design with a moderate number of participants (N=16). With this study, we aimed to identify and verify the in-practice necessity and suitability of the teleoperation requirements defined in prior literature. Since the purpose was not to compare or refine production ready interfaces and evaluate UI performance at a very low level of detail, we were more interested in “large” effects that signify overall viability and performance of indicators and information elements. Such effects are typically visible already with smaller sample sizes and, accordingly, we were able to gather a sufficient yield with the chosen approach and sample size, as the results and analysis showed.

7 Conclusions and future work

In this paper, we investigated two versions of a teleoperation UI with a minimal and maximal set of additional indicators based on requirements extracted from related literature. We were able to confirm a number of these requirements and could additionally determine their relevance for the orientation phase and the navigation phase.

We found few overlaps between both phases and most elements that are important for the orientation phase to be no longer relevant during navigation. The orientation phase emphasizes determining vehicle location, driving behavior and physical vehicle parameters, including incurred damage and loaded cargo, if any. The navigation was identified to be less about standard navigation in terms of geolocation and instead mostly emphasizes supporting the mechanical navigation task, including lane position, proximity information, speed limits, and visual- and/or control latency.

While the extended reality setup enabled a more realistic interaction than a standard simulator study, not all aspects of teleoperation could be investigated (e.g., different weather conditions) and the tasks were limited to navigation tasks (driving from point A to point B). Thus, further investigations into additional application contexts and a wider range of teleoperation tasks (e.g., object removal, loading and unloading) for more refined results regarding necessary in-UI information for these contexts and tasks. Nonetheless, we consider the proposed information elements in this paper to be a usable basis for in-context teleoperation with further room for extension and improvement.

Acknowledgements This work is part of the project AWARD, which has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101006817. The content of this paper reflects only the author’s view. Neither the European Commission nor CINEA is responsible for any use that may be made of the information it contains.

Data Availability The datasets generated during and/or analysed during the current study are not publicly available due to data management and protection provisions this work is subject to requiring a specific cause for sharing of individual data sets. They can be made available from the corresponding author on reasonable request.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Bout M, Brenden AP, Klingeagrd M, Habibovic A, Böckle MP (2017) A head-mounted display to support teleoperations of shared automated vehicles. In AutomotiveUI 2017 - 9th international ACM conference on automotive user interfaces and interactive vehicular applications, adjunct proceedings. Association for Computing Machinery, Inc, 62–66. https://doi.org/10.1145/3131726.3131758
2. Chen JYC, Haas EC, Barnes MJ (2007) Human performance issues and user interface design for teleoperated robots. IEEE Trans Syst Man Cybern B Cybern Part C (Applications and Reviews) 37(6):1231–1245. https://doi.org/10.1109/TSMCC.2007.905819. Conference Name: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)
3. Colley M, Askari A, Walch M, Woide M, Rukzio E (2021) ORIAS: on-the-fly object identification and action selection for highly automated vehicles. 13th International conference on automotive user interfaces and interactive vehicular applications
4. Fanatec (2022) Fanatec clubsport pedals V3. https://fanatec.com/us-en/pedals/clubsport-pedals-v3
5. Fanatec (2022) Fanatec clubsport steering wheel round 1 V2. https://fanatec.com/us-en/steering-wheels/clubsport-steering-wheel-round-1-v2
6. Fanatec (2022) Fanatec clubsport wheelbase 2.5V. https://fanatec.com/us-en/racing-wheels-wheel-bases/wheel-bases/clubsport-wheel-base-v2.5
7. Feiler J, Hoffmann S, Diermeyer F (2020) Concept of a control center for an automated vehicle fleet. In 2020 IEEE 23rd international conference on intelligent transportation systems (ITSC). 1–6. https://doi.org/10.1109/ITSC45102.2020.9294411
8. Raspberry Pi Foundation (2022) Raspberry Pi 4. https://www.raspberrypi.com/products/raspberry-pi-4-model-b/
9. Gafert M, Mirnig AG, Fröhlich P, Tscheligi M (2022) TeleOperationStation: XR-exploration of user interfaces for remote automated vehicle operation. CHI conference on human factors in computing systems extended abstracts
10. Georg J-M, Diermeyer F (2019) An adaptable and immersive real time interface for resolving system limitations of automated vehicles with teleoperation. In 2019 IEEE international conference on systems, man and cybernetics (SMC). 2659–2664. https://doi.org/10.1109/SMC.2019.8914306 ISSN: 2577-1655
11. Georg JM, Feiler J, DIermeyer F, Lienkamp M (2018) Teleoperated driving, a key technology for automated driving? comparison of actual test drives with a head mounted display and conventional monitors. In IEEE conference on intelligent transportation systems, proceedings, ITSC, Vol. 2018-November. Institute of Electrical and Electronics Engineers Inc., 3403–3408. https://doi.org/10.1109/ITSC.2018.8569408
12. Georg JM, Putz E, Diermeyer F (2020) Longtime effects of videoquality, videocanvases and displays on situation awareness during teleoperation of automated vehicles*. In 2020 IEEE international conference on systems, man, and cybernetics (SMC). 248–255. https://doi.org/10.1109/SMC42975.2020.9283364 ISSN: 2577-1655
13. Graf G, Abdelrahman Y, Xu H, Abdrabou Y, Schitz D, Hußmann H, Alt F (2020) The predictive corridor: a virtual augmented driving assistance system for teleoperated autonomous vehicles 9
14. Graf G, Hussmann H (2020) User requirements for remote teleoperation-based interfaces. In Adjunct Proceedings - 12th International ACM conference on automotive user interfaces and interactive vehicular applications, AutomotiveUI 2020. Association for Computing Machinery, Inc, 85–88. https://doi.org/10.1145/3409251.3411730
15. Graf G, Palleis H, Hussmann H (2020) A design space for advanced visual interfaces for teleoperated autonomous vehicles. In ACM international conference proceeding series. association for computing machinery. https://doi.org/10.1145/3399715.3399942
16. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. In Advances in psychology. Vol. 52. Elsevier, 139–183
17. iMotions (2022) Eye tracking: the complete pocket guide. https://imotions.com/blog/eye-tracking/
18. iMotions (2022) Eye tracking: VR. https://imotions.com/biosensor/eye-tracking-vr/
19. The British Standards Institution (2022) BSI Flex 1890 v4.0
20. Kadavasal MS, Oliver JH (2009) Sensor augmented virtual reality based teleoperation using mixed autonomy. J Comput Inf Sci Eng 9(1):014502. https://doi.org/10.1115/1.3086030
21. Kettwich C, Dreßler A (2020) Requirements of future control centers in public transport. In Adjunct proceedings - 12th international ACM conference on automotive user interfaces and interactive vehicular applications, AutomotiveUI 2020. Association for Computing Machinery, Inc, 69–73. https://doi.org/10.1145/3409251.3411726
22. Kettwich C, Schrank AG (2021) Teleoperation of highly automated vehicles in public transport: state of the art and requirements for future remote-operation workstations
23. Lin T-C, Krishnan AU, Li Z (2022) Intuitive, efficient and ergonomic tele-nursing robot interfaces: design evaluation and evolution. ACM transactions on human-robot interaction
24. Majstorovic D, Hoffmann S, Pfab F, Schimpe A, Wolf M-M, Diermeyer F (2022) Survey on teleoperation concepts for automated vehicles. 2022 IEEE international conference on systems, man, and cybernetics (SMC), 1290–1296
25. Mutzenich C, Durant S, Helman S, Dalton P (2021) Updating our understanding of situation awareness in relation to remote operators of autonomous vehicles. Cogn Res: Princ Implic 6(1):1–17
26. NASA (2022) NASA TLX task load index. https://humansystems.arc.nasa.gov/groups/TLX/
27. Naujoks F, Forster Y, Wiedemann K, Neukum A (2017) A human-machine interface for cooperative highly automated driving. In Advances in intelligent systems and computing, Vol. 484. Springer Verlag, 585–595. https://doi.org/10.1007/978-3-319-41682-3_49 ISSN: 21945357
28. Neumeier S, Wintersberger P, Frison A-K, Becher A, Facchi C, Riener A (2019) Teleoperation: the holy grail to solve problems of automated driving? Sure, but latency matters. In Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications (Utrecht, Netherlands) (AutomotiveUI '19). Association for Computing Machinery, New York, NY, USA, 186–197. https://doi.org/10.1145/3342197.3344534
29. Rea DJ, Seo SH (2022) Still not solved: a call for renewed focus on user-centered teleoperation interfaces. Front Robot AI 9
30. Schrepp M, Hinderks A, Thomaschewski J (2014) Applying the user experience questionnaire (UEQ) in different evaluation sce-

narios. In International conference of design, user experience, and usability. Springer, 383–392
31. Schrepp M, Thomaschewski J (2019) Handbook for the modular extension of the user experience questionnaire
32. Sheridan TB (1993) Space teleoperation through time delay: review and prognosis. IEEE Trans Robot Autom 9(5):592–606
33. Tener F, Lanir J (2022) Driving from a distance: challenges and guidelines for autonomous vehicle teleoperation interfaces. Proceedings of the 2022 CHI conference on human factors in computing systems
34. Varjo (2022) Gaze data logging. https://developer.varjo.com/docs/v3.6.0/get-started/gaze-data-collection
35. Varjo (2022) Varjo XR-3 - The industry's highest resolution XR headset | Varjo. https://varjo.com/products/xr-3/
36. Walch M, Colley M, Weber M (2019) CooperationCaptcha: on-the-fly object labeling for highly automated vehicles. https://doi.org/10.1145/3290607.3313022
37. Walch M, Sieber T, Hock P, Baumann MRK, Weber M (2016) Towards cooperative driving: involving the driver in an autonomous vehicle's decision making. Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications
38. Whitney D, Rosen E, Phillips E, Konidaris G, Tellex S (2020) Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In Amato NM, Hager G, Thomas S, Torres-Torriti M (Eds.) Robotics research. vol. 10. Springer International Publishing, Cham, 335–350. https://doi.org/10.1007/978-3-030-28619-4_28 Series Title: Springer Proceedings in Advanced Robotics