

# Introducing ROADS: A Systematic Comparison of Remote Control Interaction Concepts for Automated Vehicles at Road Works

Mark Colley

m.colley@ucl.ac.uk  
Institute of Media Informatics, Ulm  
University  
Ulm, Germany  
UCL Interaction Centre  
London, United Kingdom

Jonathan Westhauser

jonathan.westhauser@uni-ulm.de  
Institute of Media Informatics, Ulm  
University  
Ulm, Germany

Jonas Andersson

jonas.andersson@ri.se  
RISE Research Institutes of Sweden  
Göteborg, Sweden

Alexander G. Mirnig

alexander.mirnig@ait.ac.at  
Department of Artificial Intelligence  
and Human Interfaces, University of  
Salzburg  
Salzburg, Austria  
Center for Technology Experience,  
AIT Austrian Institute of Technology  
Vienna, Austria

Enrico Rukzio

enrico.rukzio@uni-ulm.de  
Institute of Media Informatics, Ulm  
University  
Ulm, Germany

## Abstract

As vehicle automation technology continues to mature, there is a necessity for robust remote monitoring and intervention features. These are essential for intervening during vehicle malfunctions, challenging road conditions, or in areas that are difficult to navigate. This evolution in the role of the human operator—from a constant driver to an intermittent teleoperator—necessitates the development of suitable interaction interfaces. While some interfaces were suggested, a comparative study is missing. We designed, implemented, and evaluated three interaction concepts (*path planning*, *trajectory guidance*, and *waypoint guidance*) with up to four concurrent requests of automated vehicles in a within-subjects study with  $N=23$  participants. The results showed a clear preference for the *path planning* concept. It also led to the highest usability but lower satisfaction. With *trajectory guidance*, the fewest requests were resolved. The study's findings contribute to the ongoing development of HMIs focused on the remote assistance of automated vehicles.

## CCS Concepts

• Human-centered computing → Empirical studies in HCI.

## Keywords

remote operation, systematic comparison

### ACM Reference Format:

Mark Colley, Jonathan Westhauser, Jonas Andersson, Alexander G. Mirnig, and Enrico Rukzio. 2025. Introducing ROADS: A Systematic Comparison



This work is licensed under a Creative Commons Attribution 4.0 International License. CHI '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1394-1/25/04

<https://doi.org/10.1145/3706598.3713476>

of Remote Control Interaction Concepts for Automated Vehicles at Road Works. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*, April 26–May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/3706598.3713476>

## 1 Introduction

The road toward fully automated vehicles (AVs) is still paved with numerous challenges that must be overcome. Full automation, i.e., without the need for human intervention for regular intervention or fallback-readiness, is more readily achievable in controlled environments like airports or warehouses, where vehicle speeds and traffic stakeholders can be regulated. While AVs are already available in cities like Phoenix and San Francisco with state restrictions [27], widespread adoption is not yet in sight. The technology is advancing rapidly and has anticipated significant advantages in safety, reliability, passenger comfort, and reducing economic and environmental costs [11].

However, achieving highly automated driving (SAE level 4 [29]) in mixed-traffic urban environments is challenging. When an environment can not be fully controlled, there will be situations that the automated system does not anticipate and will likely not possess the correct solution or exit strategy. This is why, during the transition phase towards full automation, for example, leveraged by delivery companies to benefit from anticipated lower costs, human safety drivers or operators are used to close this gap and be able to respond to unforeseen circumstances intuitively. Restrictions on the operational design domain (ODD), e.g., when local authorities prohibit the operation of AVs within road works [10], further restrict the operation potential of fully automated AVs. To address these and similar challenges, remote operation systems have been proposed. These provide human monitoring and intervention capabilities, potentially in a one-to-many interaction setting, without requiring the physical presence of a human in the vehicle, where a remote

operator (RO) from a control center guides the AV [13, 14, 22, 23, 26]. This was also shown in a pilot study referenced by Vreeswijk et al. [37], which presents a real-world application. This strategy can apply to various vehicle types, including personal cars, vans, trucks, and buses.

While several parts of remote operation, such as remote driving (Tener and Lanir [34]), HMI requirements (VINNOVA [36]), and (technical) challenges and limitations (Kang et al. [20]), have been evaluated, a systematic comparison of user interfaces (UIs) for RO intervention when a vehicle needs assistance has not yet been conducted. In addition, it is commonly assumed that a teleoperator will monitor – and potentially have to intervene on – more than a single vehicle [13]. Yet, the question of how many assistance events one operator can handle, given a specific interaction design, is not well studied for regular assistance scenarios [24]. Therefore, we implemented an integrated remote operation system called REMOTE OPERATION AUTOMATED DRIVING SUITE (ROADS). ROADS allows the RO to use the three concepts of *waypoint guidance*, *trajectory guidance*, and *path planning*, as described in the overview of Majstorović et al. [26]. In the remote operation scenario, between one and four AVs have to be controlled and steered through road works, as this could constitute the end of an ODD. In a within-subject design with N=23 participants, we found that the participants preferred *path planning*. The worst rating was given to *trajectory guidance*.

*Contribution statement:* This work contributes (1) the implementation of three interaction concepts for remote operation of AVs, which will be openly available to the broader community, and (2) results of a within-subjects study with N=23 participants evaluating the interaction concepts with one to four parallel requests showing that multiple requests are feasible with degraded performance and that *path planning* was preferred.

## 2 Related Work

This work builds on prior work on teleoperation and, in general, on cooperation between humans and automation.

### 2.1 Scenarios for Teleoperation

Kettwich et al. [22] analyzed possible scenarios for teleoperation in road traffic. These scenarios were gathered via interviews and observations with control center staff, video analyses from real-world road events, and interviews with AV safety operators. Kettwich et al. [22] categorized use cases into eight clusters, including passenger interaction, technical malfunctions, and interactions between the environment, the RO, the AV, and others.

In our study, the scenario consists of road works too complex or out of the legal boundary for our AV to handle independently and, therefore, is out of its ODD, relating to the unmapped construction site [22].

### 2.2 Teleoperation Concepts

AV teleoperation intensities can be divided into three categories. Bogdoll et al. [4] defined *remote monitoring*, *remote assistance*, and *remote driving* in ascending order of the interaction intensity (see also Andersson and Söderman [2]). *Remote monitoring* refers to a one-sided stream of information from the vehicle to the RO. *Remote*

*assistance* requires the RO to aid the AV in its decisions but not take over to perform actual driving tasks, contrary to *remote driving*. Bogdoll et al. [4] define *remote driving* "as 'real-time performance of part or all of the DDT [dynamic driving task] and/or DDT fallback [...] by a remote driver.'" [4, p. 2], while directly controlling the vehicle "in the form of steering and throttle/brake commands" [4, p. 10].

Using Bogdoll et al. [4]'s taxonomy, Majstorović et al. [26] surveyed prior work. They split *remote driving* and *remote assistance* into concepts. We evaluated concepts of *trajectory guidance* (building a path by dragging the mouse cursor), *waypoint guidance* (generate a connected path by placing single points), and *interactive path planning* (user chooses one of multiple system-offered paths) and compared them. We exclude other concepts as they are either strongly tied to *remote driving* or too use case specific. We specifically exclude the concepts of explicit *remote driving* due to their limitations in teleoperation (see Section 2.5).

Although Majstorović et al. [26] classify *trajectory guidance* as a form of *remote driving*, based on the practical implementation we are proposing, the fluid transition between those categories and the corresponding definition in the taxonomy by Bogdoll et al. [4], for further discussion, we categorized it as *remote assistance*. Additionally, the interaction type for actions in remote assistance is very different compared to active control due to the different levels of interaction abstraction (direct movement vs., e.g., point guidance) and the event-driven interaction requests in remote assistance [26].

Brecht et al. [5] evaluated several teleoperation concepts (direct control, shared control, trajectory guidance, waypoint guidance, collaborative planning, and perception modification) from an ego perspective with eight experts in the automotive and teleoperation industry. Such experts might not be the final users of these teleoperation concepts, so we had participants without a background in these industries. They found that "a holistic teleoperation system should be composed of implementations of the Shared Control, Collaborative Planning and Perception Modification concepts" [5, p. 639].

Colley et al. [8] conducted a literature survey regarding the "final 100 meters problem" [8], which relates to the variability of the final destination in a journey based on user preferences and technical limitations. The survey was conducted on "possible interaction modalities and modes" [8, p. 2], creating an overview of possibilities to communicate with the vehicle. The results of a VR study indicated a preference for a steering wheel, followed by using a joystick. While Colley et al. [8] evaluated interaction concepts from inside the AV, we compare concepts for remote assistance. This opens new possibilities and interaction perspectives (e.g., handling multiple requests simultaneously or simplified representations of the environment).

### 2.3 Human Factors

**2.3.1 Human Factors in Remote Operation.** Kari and Steinert [21] identified the human factors related to remote operations. Although they focus on remote *ship* operations, many factors apply in an automotive context. Especially situational awareness (SA) [14, 16, 23, 26], the lack of physical sensing and soundscape [34] and the implied delays [9, 20, 34] are important for ROs.

Additionally, Hashimoto et al. [18] raised the question of what the maximum feasible count  $N$  of vehicles managed concurrently is. Further challenges involved are determining the difficulty levels when one RO has to grasp the situation for multiple vehicles in case of emergencies or determining "how many additional operators (up to  $N-1$  carts) are required when one cart is in the takeover?" [18, p. 6] (cart  $\hat{=}$  vehicle). Additionally, switching between vehicles and switching costs shall be considered [18]. Thus, our study addresses the applicability of one operator managing multiple vehicles concurrently. It takes performance measurements on parallel incoming vehicle requests, tied with the currently used interaction modality. To support parallel request handling, our proposed UI provides an additional secondary display for a selected vehicle to display its front camera to monitor the vehicle.

**2.3.2 Organization and Roles in Remote Operation Centers (ROCs).** Schrank and Kettwich [30] proposed a "first conceptual step" for an organizational structure for ROCs. While being a preliminary concept, it lays a foundation for future discussions. This proposal defines three central roles, the Remote Coordinator (RC), the Remote Driving Operator (RDO), and the Remote System Operator (RSO), along with multiple Peripheral Roles that support and complement core operations. The primary task of an RC is to maintain an overview and monitor all incidents and how they are handled. RDOs are responsible for remotely controlling individual AVs, while RSOs manage a fleet of AVs.

## 2.4 Interface Requirements

Several studies were conducted to determine the requirements for remote control interfaces for AVs, for example, the HAVOC Project [36]<sup>1</sup>. 15 participants monitored 10 automated trucks and had to assist in five occurring events actively. Two of those events were active control tasks conducted with a steering wheel, pedals, and a computer screen. The feedback focused on the lack of information to consciously control the vehicle, including the need for a 360° overview of the surroundings and the uncertainty of when the AV can take over again, which we aimed to resolve in our prototype. To achieve this within the interface, we included a projected birds-eye view of the vehicle, added information on collision avoidance detection, and notified the RO when the vehicle took over again.

Comparably, Tener and Lanir [34] studied remote driving. The requirements we identified and adapted for our use case concerning remote assistance included emphasizing the takeover reason, giving contextual road information, and giving AI suggestions for operations [34]. We included automated suggestions only in *path planning* for differentiation.

Gafert et al. [14] used a demonstrational VR remote driving setup, whose interface was based on the requirements of Graf and Hussmann [15]. To gather these requirements, Graf and Hussmann [15] conducted 18 interviews. Half of them had some remote controlling experiences. 80 requirements were collected. In a second interview with participants from industry, these requirements were clustered and rated, resulting in a final set of 20 requirements (e.g., 360-degree

remote view, vehicle position, vehicle issues, traffic rules), where we adopted the relevant ones.

Kettwich et al. [23] designed a Human Machine Interface (HMI) for the teleoperation of vehicles in a public transport control center context. They wanted to receive "a first impression on whether the development is in the right direction, particularly whether its overall setup is valid" [23, p. 19]. The findings included that the proposed HMI establishes a "suitable interface design for the teleoperation of highly automated vehicles in public transport" [23, p. 15]. However, some suggested improvements were related to overwhelming information. Participants dealt with 6 monitors and operated with a keyboard and mouse and an additional touchscreen. Thus, we created a more minimalist design to maintain focus on the most important elements for the given task (see Section 3).

Further, Kettwich et al. [23] defined seven HMI evaluation criteria: (1) Features, (2) Information, (3) SA, (4) Usability, (5) User Acceptance, (6) Attention, (7) Capacity. We used questionnaires to evaluate our prototypes. We used the SUS [6] to measure user usability, the van der Laan acceptance scale [35] to measure User Acceptance, and NASA-TLX [17] to measure user capacity. Additionally, participants were asked for liked and missed features. However, we did not include dedicated scales for measuring the remaining criteria: Information, SA, and Attention, but instead relied on logged data to provide recognition to them.

## 2.5 Limitations of Remote Controlling

While controlling AVs remotely brings many advantages, limitations also exist. Cummings et al. [9] and Tener and Lanir [34] raise critical issues addressing the remote driving of AVs. Delays concerning the technical side of communication and the human side of reorientation and reaction time narrow the applicability of remote operations. Therefore, Cummings et al. [9] point out that requiring an RO to take over the vehicle at high speeds in a short time frame in the range of seconds would be unsafe. Taking these limitations into account, the expected safe situations that remain open for a takeover are those in which the vehicle is traveling at slow speeds, less than 10 mph [9], or has already pulled to the shoulder by performing a DDT as specified by the SAE [29] in a critical situation. The technical delay in communication was further elaborated by Kang et al. [20]. They present a case study further investigating the imposed delay by comparing the effect of different levels of resolutions in LTE [32] and Wi-Fi [33], which represent "wireless networks used by vehicles" [20, p. 22]. Their results show that higher resolutions, i.e., higher image sizes, lead to higher image latencies.

Concerning the given limitations on remote driving operations, we propose that the effects of limitations on remote assistance are less intensive and, therefore, bearable. As the task of speed and brake control is delegated to the vehicle, many involved security issues caused by latency and human factors (see Section 2.3) are minimized or entirely mitigated. Additionally, enabling the RO to plan vehicle routes for limited distances within remote assistance also enables them to assist multiple vehicles simultaneously, contrary to basic remote driving. Therefore, our study focuses on remote assistance to take advantage of the benefits an SAE 4 vehicle presents to mitigate security risks.

<sup>1</sup><http://tinyurl.com/havocproject>; Accessed by: 29.11.2024

### 3 User Study

We systematically explored three interaction concepts based on a previous overview by Majstorović et al. [26]: *waypoint guidance*, *trajectory guidance*, and *path planning*. Additionally, we were interested in the effect of multiple parallel requests. While the apparatus allows up to seven parallel requests, internal pilot tests before the main study showed that four parallel requests are already highly demanding. Therefore, the independent variables were the *interaction concept* and the *number of requests*. We did not include overview maps and technical details of the requesting vehicle. While these are relevant data for a real-life application, for our research question (RQ), these were irrelevant due to the focus on a self-contained scenario without references to other locations. The two RQs that guided this study were:

RQ1: What are the effects of different *interaction concepts* on potential remote operators in terms of (1) usability, (2) acceptance, (3) performance, (4) efficiency and effectiveness?

RQ2: What are the effects of scaling the *number of requests* with the *interaction concept*?

#### 3.1 Apparatus



Figure 1: Scene overview of the road works.

The apparatus comprises a Unity application for the driving and interaction environment (see Figure 1). We designed a standalone application using Unity 2022.3.7. This application simulates one or more AVs approaching road works. These road works are also mirrored to avoid participants becoming too accustomed to the environment without changing the actual task’s difficulty. In the scenario, there are three lanes with other vehicles. The driving environment aligns with previous research on takeovers on highways [7]. We used the Mobile Traffic System asset<sup>2</sup> to simulate traffic behavior. The traffic and AV speed at the start of the interaction is  $\approx 120$ km/h.

The traffic alongside the assisted AV behaved according to German traffic regulations, which prohibit overtaking within this road works area due to the present lane markings. Consequently, if a

<sup>2</sup><https://assetstore.unity.com/packages/tools/behavior-ai/mobile-traffic-system-194888>; Accessed on 14.08.2024



Figure 2: Workspace of the participants. The image shows the monitor used for the experiment with a mounted webcam.

request is neglected (e.g., waiting for instructions) for too long, it will cause a traffic jam in its lane.

When a path is processed, the AV begins to follow it without prior effectiveness checks (e.g., we did not check whether this is the fastest path or whether the path would lead to a successful exit of the construction site). Therefore, planning a suitable path is entirely delegated to the RO, while the AV retains responsibility for proper path following, including collision avoidance and cruise control. This follows the application requirements of at least an SAE-Level 4 vehicle to perform *remote assistance*, as specified by Bogdoll et al. [4].

To resolve an incoming request, a total distance of 600 meters had to be traversed, of which the vehicle had an initial path of 200 meters already assigned when issuing the request, which could optionally be overwritten by the RO. This path distance is implied by the specified view range of our remote assistance view, which is set to 200 meters, starting from the vehicle. This aligns with the minimum range requirement for *LiDAR* sensors, as specified by Bastos et al. [3]. Current *LiDAR* systems also already meet this requirement as Liu et al. [25] mention. To safely travel the distance required to resolve the request without changing the initial zoom level, an RO would have to make a minimum of ten separate path inputs for *trajectory guidance*, seven for *waypoint guidance*, or three for *path planning*. However, zooming was enabled, therefore, these values are just for reference.

The RO was given 120 seconds to resolve all incoming requests in the study. When multiple requests were to be assisted, all of them came in at the same time.

Regarding hardware, we used a QHD 27-inch monitor (i.e., the Dell S2721D) and a standard cable mouse and keyboard (see Figure 2). In this work, we deliberately omitted technical limitations of RO such as latency to test the *interaction concept* effectively.

## 3.2 Procedure

Each participant experienced **12** conditions, representing a  $3 \times 4$  design (*interaction concept* with three levels: *trajectory guidance*, *waypoint guidance*, and *path planning* and *number of requests* with values ranging from one to four). Each condition (i.e., combination of *interaction concept* and *number of requests*) had one trial. Therefore, every participant had to handle 30 requests in total.

Each session started with a brief introduction and signing of the consent form. Afterward, participants had time to get accustomed to the systems. Each participant was introduced to all system capabilities, including the user interface and the interaction concepts. Furthermore, each participant was encouraged to familiarize themselves with the system and particularly with the interaction concepts until becoming confident with them. For this purpose, the same scenario was used as in the real conditions with two concurrent requests. The conditions were then presented in counterbalanced order using a Latin square. The conditions with the same *interaction concept* were not placed together. No participant mentioned any frustration by constantly switching interaction modalities. The introduction to the capabilities of the system was given as follows:

*You are a remote operator for automated vehicles that are not capable to operate in all possible conditions. Therefore, if they encounter scenarios in which they cannot or are not allowed to operate, you are asked to take control from afar. Depending on the condition, you will use different interaction concepts to guide the automated vehicle through the scenario.*

After all conditions, a demographic questionnaire was filled out by the participants. On average, a session lasted 80 min. Participants were compensated with 15€.

## 3.3 User Interface

**3.3.1 Request Management.** At the beginning of the interaction, the left panel shows the requests of the vehicles. The middle is still empty and at the bottom, there are fields for information on the emerging requests (see Figure 3). The possible keystrokes on the right apply to the interaction concept used.

The requests can be dragged to the center or to the right edge via drag-and-drop. The areas for this are highlighted. Up to two requests can be active at the same time; while only one of them can be controlled (the large one in the middle), the other one (right at the edge) can only be observed.

In the lower panel, the reason for the required takeover is described. Just above it is the speed of the vehicle, and to the right of it, a representation of a vehicle, from which the front and rear are illuminated in yellow as soon as something is near the vehicle. This indicates that the vehicle cannot drive in that direction.

In the input panel for the interaction, there are three buttons. From left to right, first, the "Vehicle Focus" button. When pressed, the view constantly follows the vehicle. The second is the "Path End Focus". Here, the view jumps to the end of the path as long as this button is active. The focus is lost as soon as the operator moves the view by pressing and dragging the left mouse button. Third is the "lock button". When this is active, the operator restricts the manual dragging direction of the view to only forward or backward

movement. Otherwise, all directions are possible. One can also drag opened requests back to the list on the left anytime. The goal is to answer the requests in such a way that the vehicle can drive independently again. The vehicle will announce itself by means of a text in the middle as soon as this is the case. The request disappears shortly thereafter.

**3.3.2 Interaction Concepts.** In general, starting and stopping were handled by the AV as long as collision avoidance did not trigger and was no task of the RO. Compared to previous work, e.g., by Brecht et al. [5], our interface provided a bird-view visualization compared to an ego perspective. Importantly, these concepts also work **without** clear lane separation; therefore, simplifications like buttons for "move along" or "switch lane" maneuvers were avoided.

**Waypoint Guidance.** The user can form the path by sequentially clicking the right mouse button and the resulting points (see Figure 4a; also called waypoint guidance by Brecht et al. [5]). Points that would form an internal angle of  $\leq 90^\circ$  will result in no waypoint set, as this maneuver would not be possible by a vehicle.  $90^\circ$  turns are possible with advanced steering systems<sup>3</sup>. Importantly, the final movement of the AV was always determined by the Polarith AI asset<sup>4</sup>. This incorporates calculations for semi-realistic movements. The RO can add waypoints between others and move points after pressing the mouse. In addition, pressing the Shift key and clicking the mouse can delete a waypoint again. Holding the control key when placing and moving points sets the position to the center of a lane, similar to the *trajectory guidance* interaction.

**Trajectory Guidance.** The RO can draw the path by dragging the mouse (see Figure 4b; also called trajectory guidance by Brecht et al. [5]). The right mouse button must be pressed during this process. This generates equidistant waypoints. Hectic movements in opposite directions are blocked. These are detected by taking the next three consecutive waypoints and forming an (imaginary) triangle. If the internal angle at the second point is  $\geq 90^\circ$ , the input is blocked. When the mouse button is released, the newly drawn path is accepted and the vehicle starts to follow it. This interaction also enables to extend and fully or partially replace paths.

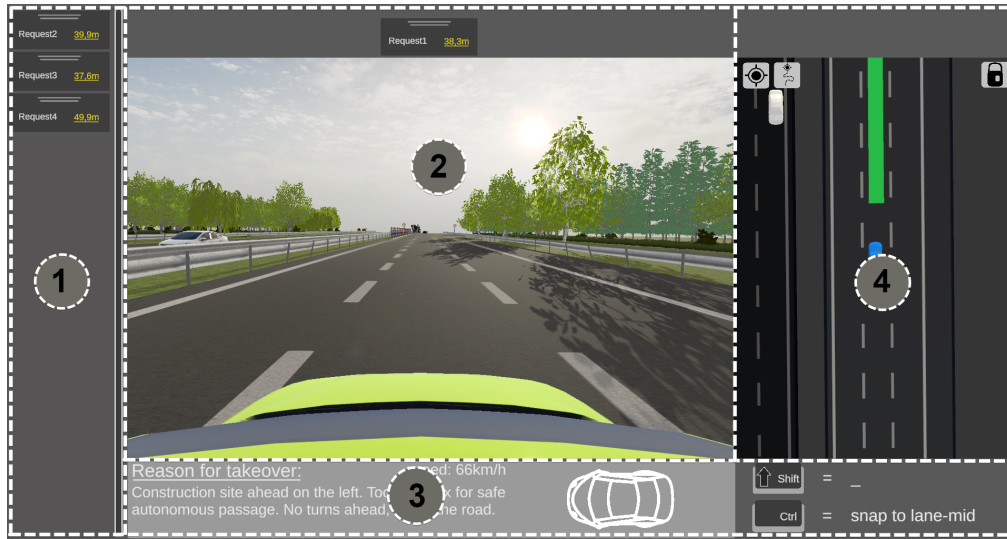
What happens in the end when drawing new paths depends on the distance of the start and end points to the last drawn path:

- **Extension:** If only one point of the new path is close (i.e., 3.5m) to the existing path, the old path is extended. The extension direction depends on the distances to the start and end points and the vehicle involved.
- **Replacement:** If both the start and end points of the new path are close to the old one, the segment in between on the old path is replaced with the new path.
- **Parallel Replacement:** If the new path is close enough to the old one and runs parallel, the parallel part is replaced with the new path. This rule is rarely used and has a low threshold to prevent interference with extensions.

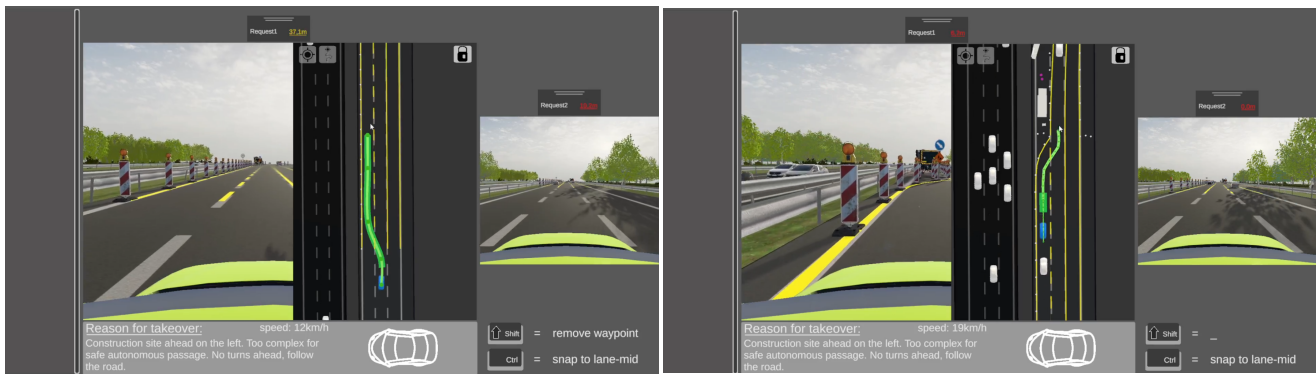
Holding down the control key automatically draws in the middle of the lane next to the cursor, simplifying the process of drawing

<sup>3</sup><https://www.youtube.com/watch?v=uZas2YCV3JY>; Accessed: 29.11.2024

<sup>4</sup><https://assetstore.unity.com/packages/tools/behavior-ai/polarith-ai-movement-with-3d-sensors-71465>; Accessed on 08.12.2024



**Figure 3: The main screen of the application. Here, four requests have to be handled in parallel. UI elements overview: (1) request list, (2) main request camera, (3) information panel, (4) remote assistance input panel.**



**(a) Waypoints. 4 individual waypoints (shown in dark green) are automatically connected. A second request is shown to the right. (b) Trajectories. By moving the pressed mouse cursor, points (shown in dark green) are combined into a trajectory.**

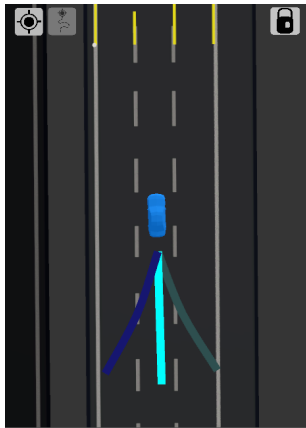


**(c) Path planning. Three potential paths are visible and color-coded. A second request is in the queue don't he left.**

**Figure 4: Overview of employed interaction concepts for remote operation. The ego-vehicle is shown in blue. The green part shows the future trajectory that is already planned.**

along the roadway when no special maneuvers, such as changing lanes or driving outside regular lane boundaries, are required.

**Path Planning.** The user can set the path by selecting from suggested paths with the right mouse button (see Figure 4c; called collaborative planning by Brecht et al. [5]). In the evaluated scenario, up to three paths (two while driving through the construction site) are continuously generated depending on the situation (i.e., the number of lanes). By holding the shift key, one can select from three static generated reverse paths, as seen in Figure 5. These do not align with any roadway shapes and represent a fallback option to reposition in case of blocking objects in front of the vehicle or required orientation adjustments. This interaction concept includes the assumption that the AV will still be able to plan possible paths but either might not be allowed to execute them due to a legal boundary or the confidence in the applicability might be too low. Waymo claims to “identify [...] stop signs greater than 500 meters away” [19], thus, we cautiously showed trajectories up to 185m in front of the AV.



**Figure 5: Backward paths in *path planning*. Three options are possible, as for forward paths. For visual distinction, other colors are used.**

### 3.4 Participants

We computed the desired sample size for the main experiment via an a-priori power analysis using G\*Power [12]. To achieve a power of .8, with an alpha level of .05, theoretically, 21 participants should result in an anticipated medium effect size (*Effect Size f*=0.27) in a repeated measures ANOVA.

We recruited N=23 participants (Mean age = 23.9, SD = 2.60, range: [18, 28]; Gender: 34.8% women, 60.9% men, 4.35% non-binary; Education: High school, 56.52%; College, 43.48%). Participants played video games on average  $M=12.91$  ( $SD=15.43$ ) hours per week. On 5-point Likert scales ( $1 = Strongly Disagree - 5 = Strongly Agree$ ), participants showed high interest in AVs ( $M=4.57$ ,  $SD=.51$ ), believed AVs to ease their lives ( $M=4.22$ ,  $SD=.85$ ), but were skeptical about whether they become reality by 2033 ( $M=3.57$ ,  $SD=1.34$ ). Participants had no prior expertise in remote control.

### 3.5 Measurements

**Objective dependent variables:** During each session, the system logged the eye gaze with 60 Hz for the four areas of interest: request panel, info panel, main panel, and secondary panel. We also logged the current lane deviation in meters and mouse movements (only after the third participant due to technical problems) in cm with 10 Hz.

**Subjective dependent variables:** After each condition, we measured the Task Load Index via the NASA-TLX [17] on 20-point scales and usability via the system usability scale (SUS) [6] which is a 10 item questionnaire with five response options: from Strongly agree to Strongly disagree. These items are: “I think that I would like to use this system frequently.”, “I found the system unnecessarily complex.”, “I thought the system was easy to use.”, “I think that I would need the support of a technical person to be able to use this system.”, “I found the various functions in this system were well integrated.”, “I thought there was too much inconsistency in this system.”, “I would imagine that most people would learn to use this system very quickly.”, “I found the system very cumbersome to use.”, “I felt very confident using the system.”, “I needed to learn a lot of things before I could get going with this system.”. Finally, we asked participants about their acceptance using the van der Laan acceptance scale [35] consisting of nine semantic differentials rating a system. It measures acceptance on the two subscales Usefulness (averaging the five items: Useful ↔ Useless, Bad ↔ Good, Effective ↔ Superfluous, Assisting ↔ Worthless, Raising Alertness ↔ Sleep-inducing) and Satisfying (averaging the items: Pleasant ↔ Unpleasant, Nice ↔ Annoying, Irritating ↔ Likeable, Undesirable ↔ Desirable).

After all conditions, participants rated their preferences regarding the interaction concepts from greatest (*ranking = 1*) to lowest (*ranking = 3*). Open questions regarding improvement proposals were also asked (i.e., “Why did you rate them that way?” and “Please describe further positive aspects and ideas for improvement.”).

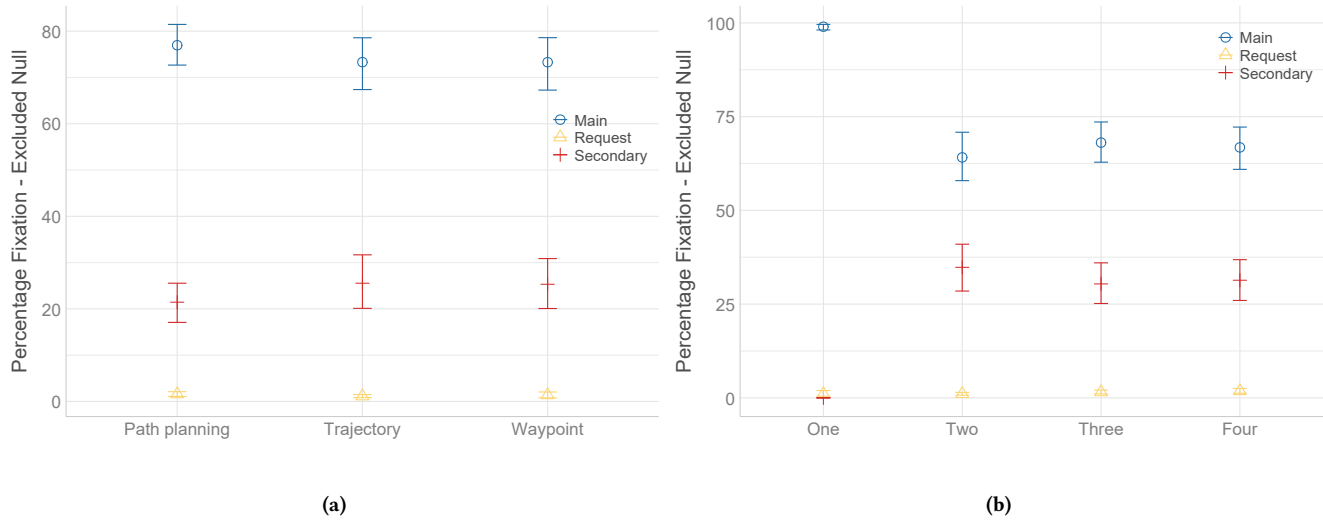
## 4 Results

### 4.1 Data Analysis

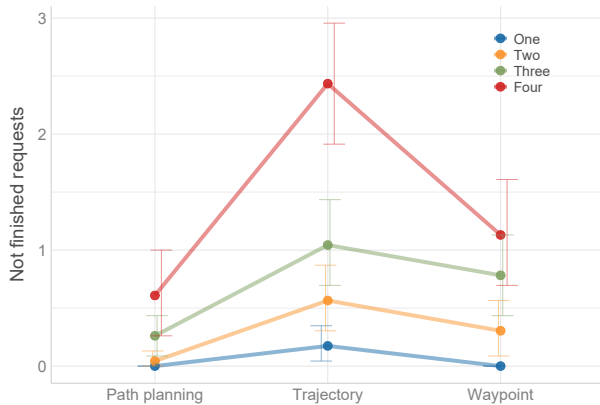
Before every statistical test, we checked the required assumptions (e.g., normality distribution). For non-parametric data, we used the ARTool package by Wobbrock et al. [38] as the typical ANOVA is inappropriate with non-normally distributed data and Holm correction for post-hoc tests (using Dunn’s test). The procedure is abbreviated, as in the original publication, with ART. R in version 4.4.1 and RStudio in version 2024.09.0 was employed. All packages were up to date in December 2024.

### 4.2 Eye Gaze

The descriptive eye gaze data in Figure 6 shows clearly that the main panel was always viewed the most. The actual request panel received very little attention from the participants. If multiple requests were present, participants did focus > 25% of their visual attention on the secondary panel (see Figure 6b).



**Figure 6: Fixations regarding the *interaction concept* (a) and the *number of requests* (b). Displayed are the gaze areas in percentages.**



**Figure 7: Number of missing requests per *interaction concept*.**

### 4.3 Number of Missed Requests

The ART found a significant main effect of *interaction concept* ( $F(2, 38) = 46.97, p < 0.001$ ), of *number of requests* ( $F(3, 57) = 37.34, p < 0.001$ ), and a significant interaction effect of *interaction concept*  $\times$  *number of requests* on number of missed requests ( $F(6, 114) = 9.09, p < 0.001$ ; see Figure 7). While the number of missed requests naturally increased with more *number of requests*, the difference between the *interaction concept* increases drastically between *trajectory guidance* when four requests are present. This indicates that the effort for *trajectory guidance* increases faster with more requests.

### 4.4 Lane Deviation

Lane deviation is a key metric for evaluating path quality during the RO of AVs, as it reflects how well the operator's inputs align with the optimal path, which is in the middle of a lane. Every lane

switch is somewhat risky. It also measures operator performance, highlighting their ability to maintain situational awareness and provide effective guidance. Since the AV handles safety-critical functions like collision avoidance, lane deviation becomes a proxy for potential operational risks and system usability.

**Time-dependent lane deviation** (see Figure 8a): This value is calculated by summing up the current lane deviation (distance to the next lane middle) for every active request in the condition, 10 times per second. The distribution of these values for every participant is visualized in Figure 8a, with the averages for each condition highlighted (lower is better).

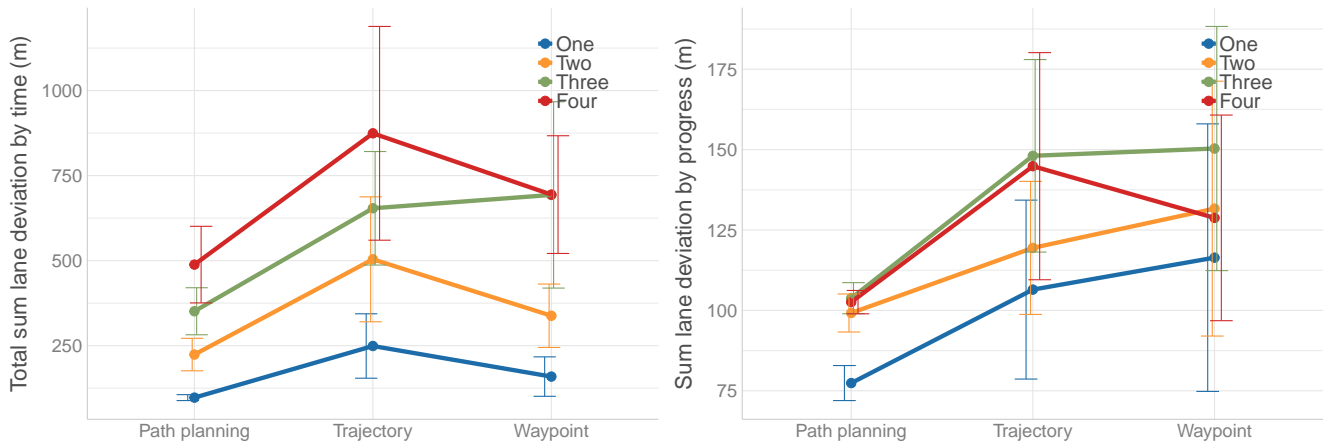
The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 28.89, p < 0.001$ ) and of *number of requests* ( $F(3, 66) = 82.07, p < 0.001$ ) on *absolute sum of lane deviation*.

A post-hoc test found that Four was significantly higher ( $M=685.56, SD=519.48$ ) in terms of *absolute sum of lane deviation* compared to Three ( $M=566.05, SD=458.19; p_{adj}=0.032$ ), compared to Two ( $M=355.21, SD=301.27; p_{adj} < 0.001$ ), compared to One ( $M=168.39, SD=159.37; p_{adj} < 0.001$ ). A post-hoc test found that Three was significantly higher ( $M=566.05, SD=458.19$ ) in terms of *absolute sum of lane deviation* compared to Two ( $M=355.21, SD=301.27; p_{adj} < 0.001$ ) and compared to One ( $M=168.39, SD=159.37; p_{adj} < 0.001$ ). A post-hoc test found that Two was significantly higher ( $M=355.21, SD=301.27$ ) in terms of *absolute sum of lane deviation* compared to One ( $M=168.39, SD=159.37; p_{adj} < 0.001$ ).

A post-hoc test found that Trajectory ( $M=570.26, SD=520.80; p_{adj} < 0.001$ ) and Waypoint ( $M=471.05, SD=453.22; p_{adj}=0.006$ ) were significantly higher in terms of *absolute sum of lane deviation* compared to Path planning ( $M=290.10, SD=216.81$ ). A post-hoc test found that Trajectory was significantly higher ( $M=570.26, SD=520.80$ ) in terms of *absolute sum of lane deviation* compared to Waypoint ( $M=471.05, SD=453.22; p_{adj}=0.047$ ).

**Progress-dependent lane deviation** (see Figure 8b): This value is calculated by first computing each request separately. For every





(a) Total sum of time-dependent lane deviation of all active requests. (b) Sum of average lane deviation per request, weighted by progress

Figure 8: Effects on AV lane deviation.

request, all records of lane deviation are mapped to and averaged over the respective current progress in meters (floored to integers). Subsequently, these averages are joined within conditions and participants based on the respective progress value and averaged again. Finally, these averages are summed up over progress, resulting in the sum of the average lane deviation per condition and participant weighted by request progress. The distribution of these values for every participant is visualized in Figure 8b, with the averages for each condition highlighted (lower is better).

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 6.60, p=0.003$ ) and of *number of requests* ( $F(3, 66) = 30.63, p<0.001$ ) on *average sum of lane deviation*.

A post-hoc test found that Four ( $M=125.40, SD=65.25; p_{adj} < 0.001$ ), Three ( $M=134.07, SD=67.46; p_{adj} < 0.001$ ) and that Two was significantly higher ( $M=116.76, SD=60.83; p_{adj} < 0.001$ ) in terms of *average sum of lane deviation* compared to One ( $M=100.09, SD=68.29$ ).

A post-hoc test found that Trajectory ( $M=129.71, SD=68.05$ ) and Waypoint ( $M=131.79, SD=87.19$ ) were significantly higher in terms of *average sum of lane deviation* compared to Path planning ( $M=95.74, SD=15.72$ ; both  $p_{adj} < 0.001$ ).

#### 4.5 Neglected Time

Neglected time was calculated by adding up the durations during which requests were neglected (i.e., a request is issued but not worked on in the main windows) and then calculating the average of this sum (i.e., cumulative neglected time total divided by neglect counts). A request is considered neglected when the RO is required to take action to enable a stopped vehicle to resume driving but does not as soon as a request is emitted.

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 60.38, p<0.001$ ), of *number of requests* ( $F(3, 66) = 89.67, p<0.001$ ), and a significant interaction effect of *interaction concept*  $\times$  *number of requests* on neglected time ( $F(6, 132) = 8.86, p<0.001$ ; see Figure 9)

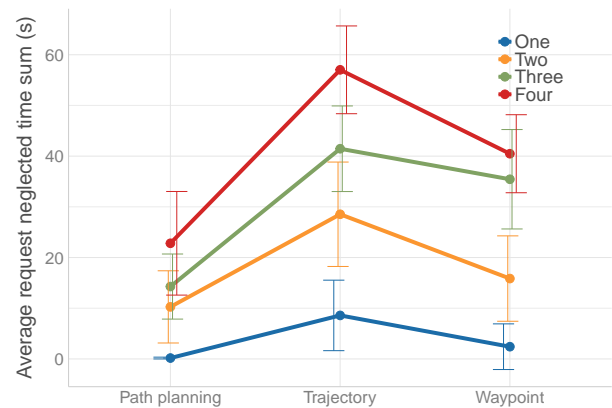


Figure 9: Interaction effect on neglected time.

The average neglected time per request naturally increased with more *number of requests*. At each level of *number of requests*, *path planning* incurs the least amount of neglect time, while *trajectory guidance* incurs the highest.

#### 4.6 Mouse Movements

The ART found a significant main effect of *interaction concept* ( $F(2, 38) = 7.30, p=0.002$ ), of *number of requests* ( $F(3, 57) = 65.80, p<0.001$ ), and a significant interaction effect of *interaction concept*  $\times$  *number of requests* on mouse travel distance ( $F(6, 114) = 5.41, p<0.001$ ; see Figure 10). While the order (from least to most movements) was consistently higher, the more requests were involved, for the *trajectory guidance*, already two requests led to very high mouse movement.

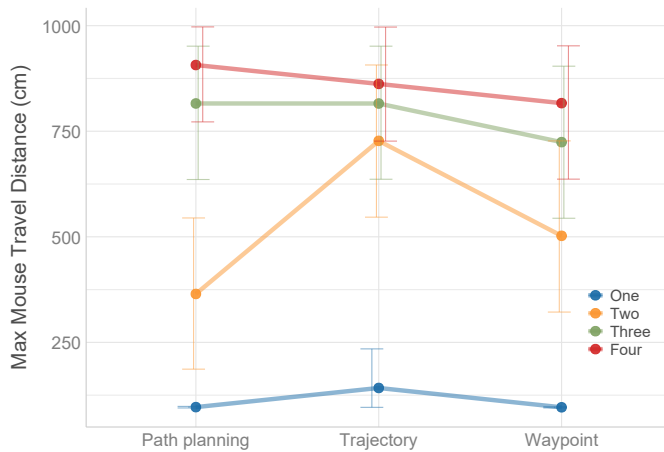


Figure 10: Interaction effect on mouse travel distance (cm).

#### 4.7 Task Load

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 20.83, p < 0.001$ ; see Figure 11a) and of *number of requests* on TLX score ( $F(3, 66) = 22.02, p < 0.001$ ; see Figure 11b).

A post-hoc test found that Trajectory ( $M=9.74, SD=4.26$ ) and Waypoint ( $M=7.95, SD=4.11$ ) were significantly worse in terms of TLX score compared to *path planning* ( $M=6.06, SD=3.30; p_{adj} < 0.001$ ) and that Trajectory was significantly worse ( $M=9.74, SD=4.26$ ) than Waypoint ( $M=7.95, SD=4.11; p_{adj} = 0.005$ ).

A post-hoc test found that Four ( $M=9.54, SD=4.41; p_{adj} < 0.00$ ), Three ( $M=8.85, SD=3.81; p_{adj} < 0.00$ ), and Two requests were significantly worse ( $M=7.77, SD=3.86; p_{adj} = 0.002$ ) in terms of TLX score than One request ( $M=5.50, SD=3.51$ ).

#### 4.8 System Usability Score

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 23.96, p < 0.001$ ; see Figure 12a) and of *number of requests* on the SUS score ( $F(3, 66) = 4.16, p = 0.009$ ; no significant post-hoc results; see Figure 12b).

A post-hoc test found that *path planning* was significantly higher ( $M=82.47, SD=13.06$ ) in terms of SUS score compared to Trajectory ( $M=58.48, SD=22.45; p_{adj} < 0.001$ ) and Waypoint ( $M=65.16, SD=21.27; p_{adj} < 0.001$ ). A post-hoc test found that Waypoint was significantly higher ( $M=65.16, SD=21.27$ ) in terms of SUS score than Trajectory ( $M=58.48, SD=22.45; p_{adj} = 0.041$ ).

#### 4.9 Usefulness and Satisfying

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 14.60, p < 0.001$ ) and of *number of requests* on Usefulness ( $F(3, 66) = 4.39, p = 0.007$  (lower is better); see Figure 13).

A post-hoc test found that Trajectory ( $M=1.27, SD=0.87$ ) and Waypoint ( $M=0.87, SD=0.79$ ) were significantly worse in terms of Usefulness compared to *path planning* ( $M=0.48, SD=0.40$ ; both  $p_{adj} < 0.001$ ) and that Trajectory was significantly worse ( $M=1.27, SD=0.87$ ) compared to Waypoint ( $p_{adj} < 0.001$ ).

The ART found a significant main effect of *interaction concept* ( $F(2, 44) = 22.25, p < 0.001$ ) and of *number of requests* on Satisfying ( $F(3, 66) = 7.01, p < 0.001$  (lower is better); see Figure 13).

A post-hoc test found that Trajectory ( $M=-0.01, SD=1.05$ ) and Waypoint ( $M=-0.52, SD=0.95$ ) were significantly worse in terms of Satisfying compared to *path planning* ( $M=-1.20, SD=0.66$ ; both  $p_{adj} < 0.001$ ) and that Trajectory was significantly worse ( $M=-0.01, SD=1.05$ ) than Waypoint ( $p_{adj} < 0.001$ ).

A post-hoc test found that Two was significantly higher ( $M=-0.47, SD=0.99$ ) in terms of Satisfying compared to One ( $M=-0.86, SD=0.99; p_{adj} = 0.048$ ).

#### 4.10 Movement Patterns

During the study's remote assistance, the position of each requesting AV was recorded. Figure 14 and Figure 15 illustrate the AVs' paths from the end of their initial route to the start of the road works. This area was specifically selected as it contains the last overtaking position and densely represents the contrasts in the steering behavior between the *interaction concepts*. For Figure 15, there are only data from two requests as for one request, the road works were always on the left side. The complete recordings of the AV positions are visualized in the appendix A. While no statistical test was conducted, a tendency to less controlled steering can be observed for *waypoint guidance* and especially for *trajectory guidance* compared to *path planning*.

#### 4.11 Ranking and Open Feedback

A Friedman rank sum test found a significant effect of *interaction concept* on ranking ( $\chi^2(2)=21.48, p < 0.001, r=0.47$ ). Participants clearly preferred *path planning*, followed by *waypoint guidance*. *trajectory guidance* was ranked the lowest. All comparisons were significant.

Users commend the system's *path planning* mechanism, describing it as "optimal," which suggests that the core functionality meets or even exceeds expectations. This is bolstered by the general agreement that the underlying concept of the system is solid, as one user explicitly stated, "The idea itself is good."

However, there are notable concerns regarding remote operation in general, particularly related to technical limitations such as latency, signal reception, and dead zones. These are considered potential hazards, as one user articulated, "Technical aspects such as delay, signal quality, signal dead zones, etc. could lead to more dangers than benefits." These technical issues raise questions about remote operation's general efficacy and safety, especially when vehicles navigate complex environments like construction zones. In this context, users desire more guidance, highlighting that "solutions measures for getting stuck in construction sites would have been very helpful."

Users find the *trajectory guidance* to be imprecise and time-consuming, suggesting a need for simplification and stabilization. This sentiment is captured in the statement, "Trajectories in comparison were generally too imprecise and time-consuming."

Regarding usability, for instance, users would appreciate the ability to have the vehicle move backward, as suggested, "As an improvement, I would suggest always having the option to drive

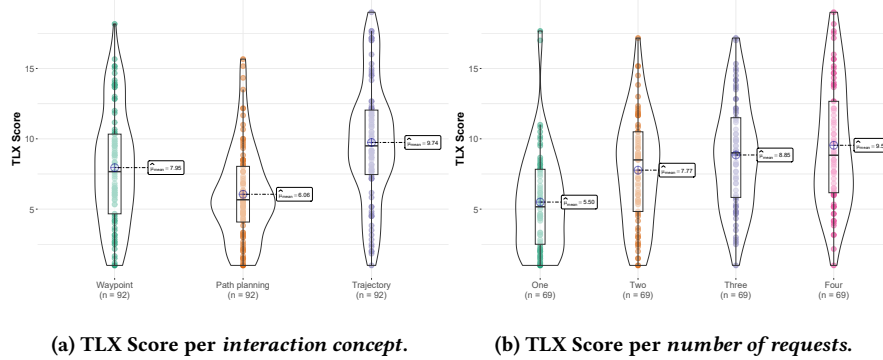


Figure 11: TLX Score per interaction concept and number of requests.

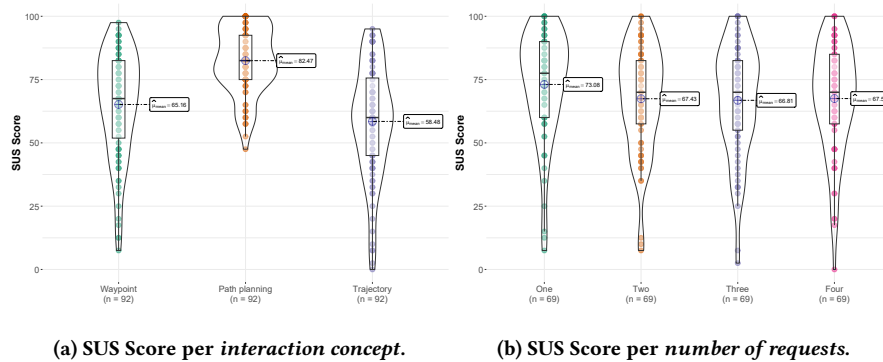


Figure 12: SUS Score per interaction concept and number of requests.

backward." This is interesting as this was **always** possible. Furthermore, there is a desire for the system to automatically focus on the most urgent cases, as mentioned in "A mode in which Main-View automatically switches to the most urgent case after a short visual warning." Additional features that could improve the system include an automatic focus on the end of the path, as one user recommended, "When using path planning, the system could jump directly to the end of the path." Again, we actually supported this via a single click. Finally, participants mentioned the desire to be able to switch between interaction concepts.

## 5 Discussion

This work compared the three interaction techniques *trajectory guidance*, *waypoint guidance*, and *path planning* for remote operation of AVs outside their ODD. The interaction techniques were tested with one to four parallel requests. We found that participants preferred the *path planning* interaction, where they could rely most on the AVs' suggestions, which also was done in previous work [1]. This leads to questions of the necessary decision support capability level, even for remote assistance.

**Eye gaze** Gaze distributions between displays are very similar across concepts. With one request, the operator focuses 100% on the main display. With 2-4 requests, the gaze distribution is equal, with around 30% on the secondary display and slightly above 60% on the main display, and only a few percent on the request display. Except

for one request, when the operator only needs to focus on the main display, eye gaze distributions are very similar despite an increase in the number of requests. Eye gaze distribution is also very similar across interaction concepts. This could indicate that task demand does not have an effect on how the screens are used. Also, it hints that more granular analysis may be necessary to reveal differences in eye-gaze patterns between interaction concepts.

**Number of missed requests and Lane Deviation** Overall, one request consistently shows the best performance with the least neglected time, while Four requests consistently has the highest average neglected time and the most variability, indicating large individual differences in operator performance. The results show that the Trajectory concept stands out as a rather poor design for handling several requests simultaneously. While Path planning has a lower variation between the number of requests, the difficulty of handling requests with the Trajectory concept is accelerated as the number of requests increases. Primarily, this result indicates the importance and benefit of appropriate system capability, i.e., that lowering the task demand by automating manual tasks can help increase the number of vehicles an operator can handle simultaneously (REF). The results also show the importance of design for the task, where the Path planning concept seems to be a better design to allow for more vehicles to be controlled under a given workload.

**Neglected time analysis** The neglected time measurements show similar patterns as the number of missed requests. As can

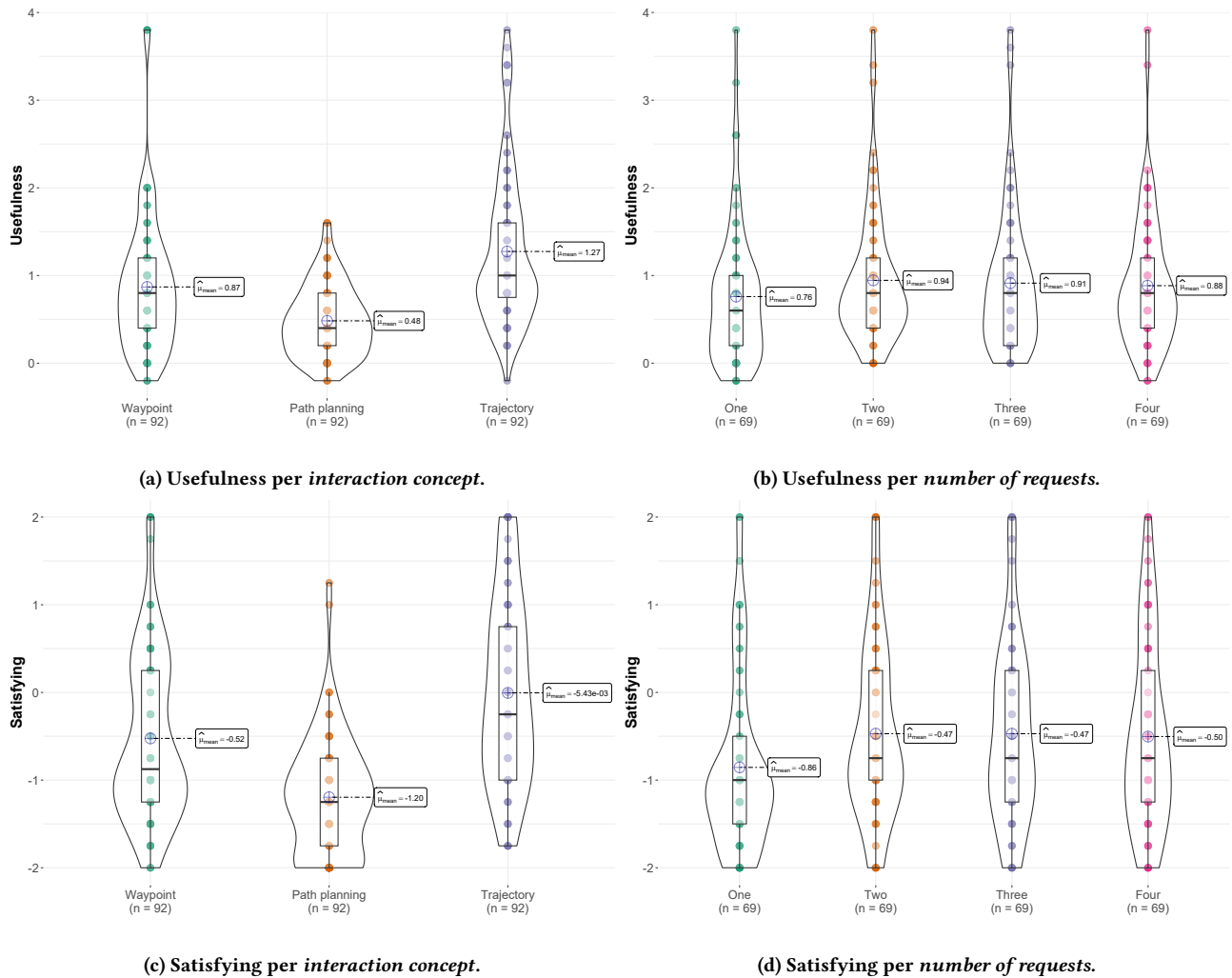


Figure 13: Usefulness and Satisfying per *interaction concept* and *number of requests*.

be expected, the data indicates that an operator’s ability to attend to requests in a timely manner becomes lower as the number of vehicles increases. It also shows that the interaction design affects how well (fast) the operators can respond. This has direct implications for the design of remote operation centers and the allocation of vehicles to operators. Specifically, the Trajectory concept shows a peak in neglected time for operators managing more than one vehicle, suggesting that this interaction concept may be particularly demanding, while Path planning gives an overall lower neglect time. The long error bars for operators managing four vehicles suggest high variability in how different operators handle this workload. This could be due to individual differences in capacity or working strategies stressing the need for training and good work procedures.

## 5.1 How Many Parallel Requests Are Feasible For Remote Operators Or Is it even Acceptable?

While it is sometimes implicitly assumed in the literature (e.g., by employing direct control via a steering wheel [1]) that an RO will only ever directly handle one request at a time, fleet management requires supervision of several vehicles [24], with an ever-present likelihood of several interventions being necessary simultaneously. The cognitive and motor demand on a human interacting with standard physical vehicle controls cannot necessarily be translated directly into a one-to-many (see Hashimoto et al. [18]) fleet management setting and is also mediated by the mode of control. Differences in input and output modulate demand simultaneous direct intervention capabilities. Our data suggests that multiple requests are feasible within certain limits. We even found that satisfaction was higher with two than with only one request (see Section 4.9). Additionally, no clear patterns are observable that indicate that a higher number of parallel requests leads to higher variability in the

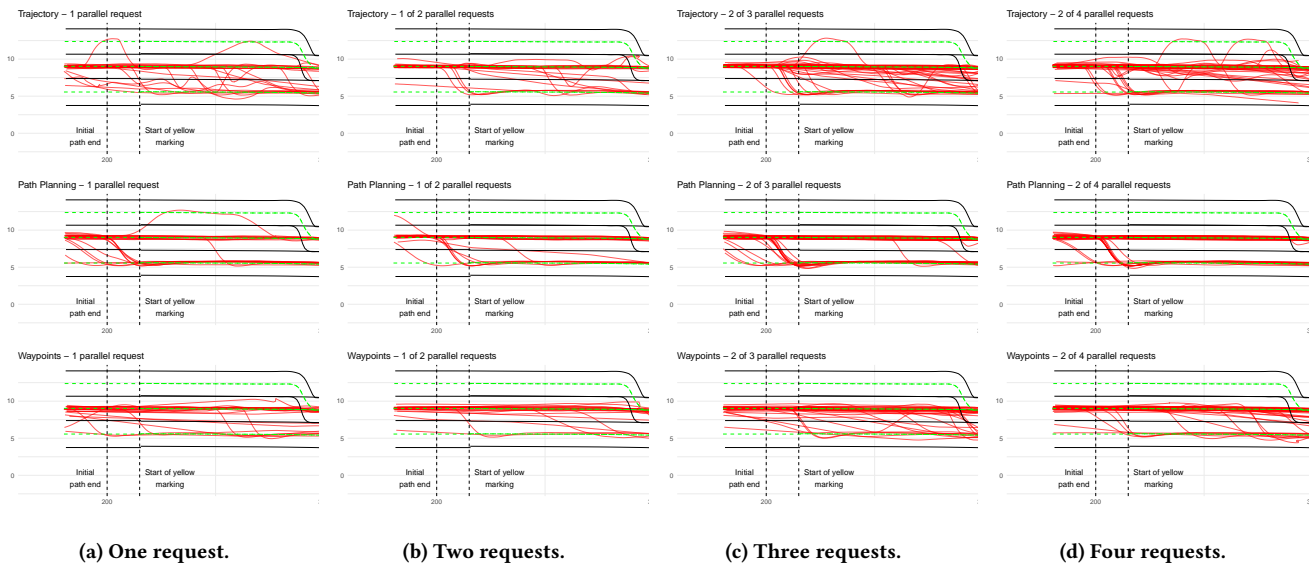


Figure 14: Movement patterns when approaching the left-sided road works.

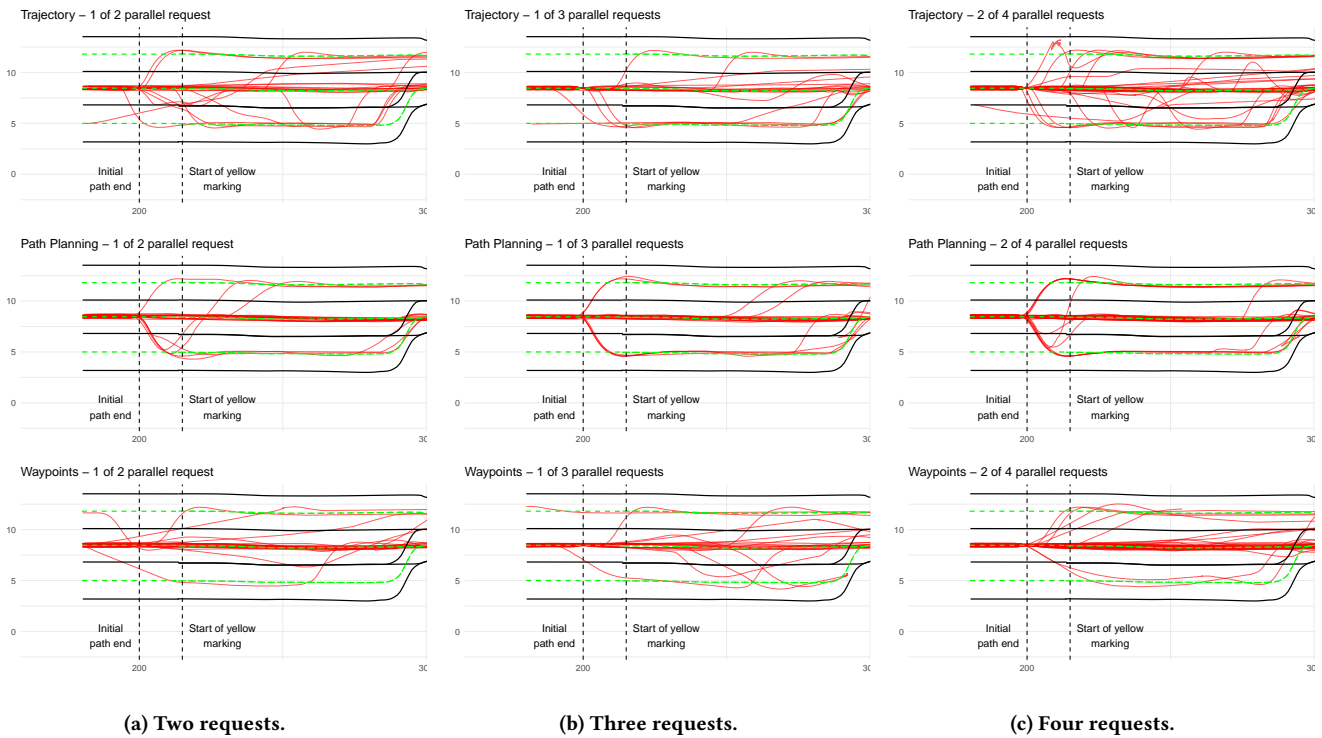


Figure 15: Movement patterns when approaching the right-sided road works.

movement patterns (see Figure 14 and Figure 15). The boundaries arise more from the question of how long AVs can be neglected and how many missed requests are acceptable or can be handled by other ROs? This emphasizes the need to also consider organizational aspects in the design of remote operation systems. Figure 7

and Figure 9 show initial data on what to expect with different techniques and number of parallel requests with alert and refreshed ROs. The question remains whether this can be generalized to all contexts or whether there are configurations (driving environment,

vehicle types, traffic density) that put increased demand on input or output capabilities and result in lower scalability potential.

## 5.2 System Capability Necessities for Remote Operation

This study found that participants could handle more than one request in parallel. While no sustained workload was induced over longer periods of time, at least for two requests, the relevant dependent variables were still in acceptable ranges. For the "ideal" number of simultaneous operations, data shows a noticeable divergence when more than three operations are involved. This suggests that 1-3 operations may be an appropriate range. However, user satisfaction peaks at two operations, narrowing the ideal range further. Future work should explore whether one operation becomes monotonous over time and whether managing three operations becomes overly demanding, irrespective of the interface or interaction technique used.

However, we highlight that the appropriateness of multiple request-handling depends not only on the interface for the RO but heavily on the capabilities of the AV controlled. Path planning is the most constrained technique in terms of input flexibility and available vehicle actions, making it the simplest to deploy for an operator. However, it requires the highest technical sophistication on the AV side. The technical capabilities of the AV necessary do not alter between *trajectory guidance* and *waypoint guidance* as a *trajectory guidance* is simply multiple waypoints taken together. The question then becomes whether it offers any distinct advantages over the reduced technical capabilities of AVs, possibly reflected in qualitative data. There is a potential for improving overall effectiveness and usability if both modes are available.

## 5.3 The Ideal Interaction Mode and the Need for Alternatives

Remote assistance is a multi-faceted interaction challenge. Therefore, the simulator to study parts of it must reflect these. Our simulator for remote vehicle operation provided simplicity and focus, allowing participants to compare operational modes easily. For this purpose, the three modalities were presented as exclusive alternatives, assuming ideal operating conditions. *path planning* received the **best** rating in **all** dependent variables, thereby, showing a clear preference. However, for *path planning* specifically, our implementation meant that the operation was predicated on there being at least one optimal or correct path to choose from. In a realistic setting, the possibility of system deficiencies leading to suboptimal path recommendations would need to additionally be considered, putting *path planning* as the lone preferred mode into perspective. Thus, *path planning* can be seen as a reasonable default but with a need for a workaround in case the default is not viable. On the basis of the results from this study, the workaround option would be the *waypoint guidance* due to possessing the necessary capabilities to more finely define a route while performing better overall than *trajectory guidance*. The eventual interaction concept that enables both scalability and finetuning- or workaround-options is one with *path planning* as the default control mode and waypoint as the supplementary mode.

An additional consideration in particular towards *path planning* is its active recommender-characteristic, which renders it most convenient. This raises additional questions regarding accountability or liability in case its recommendations are wrong [28]. Who would be held responsible if a fleet management system were to cause accidents based on suboptimal trajectory recommendations? Is it a system defect or a failure of the human operator, who should perhaps have verified the trajectory? Is it the responsibility of the third party offering the fleet management software or the OEM's who offered it to the fleet owner as part of a teleoperation-ready vehicle package? While we cannot determine these legal requirements, our study suggests that if allowed, *path planning* is an appropriate interaction mechanism for ROs.

## 5.4 Methodology

Studying remote operation is difficult due to the numerous activities, scenarios, and potentially confounding factors. Therefore, setups need to be specifically fashioned towards the specific investigative focus (vehicle control, status monitoring, routing, etc.), with one-size-fits-all setups not being feasible nor appropriate in most instances. In our case, this was the case with the focus on the interaction possibilities *path planning*, *waypoint guidance*, and *trajectory guidance*, which we decided to investigate in a comparative, single-view setting. Adding, for example, additional monitors with other information, such as a map, would have been easily doable and would be a sensible extension in a follow-up to explore a combination with spatial planning or traffic load management. However, this would reduce the focus on the intervention. The multifacetedness of remote interaction is also reflected in relevant dependent variables. We logged numerous data and also incorporated usability, task load, and acceptance as subjective measures. Arguably, the "neglected time" was the most important efficiency-related measure for the focus on parallel requests. This metric also has high relevance for real implementations and could serve as a KPI (key performance indicator) for overall performance.

## 5.5 Limitations

We studied remote operations with N=23 rather young participants. While this limits generalizability, our data can be seen as a baseline of how remote operation could work in a best-case scenario. Nonetheless, data on prolonged exposure and studies with older participants are necessary. Additionally, due to the focus on high internal validity, we did not have a full setup that could be envisioned for ROs. Our results should, therefore, be replicated in such a setup. Finally, we focused on one scenario. Other scenarios must be evaluated in the future.

The single-monitor setup minimizes distractions but may limit immersion and may not represent an actual operator environment with multiple monitors (e.g., see Schrank et al. [31]). Nonetheless, such a setup was, for example, also used by Andersson et al. [1]. Kettwich et al. [23], who supplied 7 monitors to participants, even reports that participants stated that one monitor would suffice. The major limitation for external validity, which was, however, needed for internal validity, is that the scenario focuses on ideal road conditions with a rather linear street, narrowing the study's applicability and risking steep learning curves. We reduced this by

also mirroring the road works. However, this is still a limitation. Further, there is a limitation to the fidelity of both the visual and physical simulation in the apparatus. We used the Polarith AI asset<sup>5</sup> to simulate the AV and the Mobile Traffic System asset<sup>6</sup> to simulate traffic. While this does not represent the real world, no participant mentioned any confusion about the simulation. Lastly, the lack of a quick "delete and stop" function is a notable omission regarding already started trajectories. Further features would be an overview map of vehicles while no request is active.

In this work, relevant factors such as latency [39] or potential additional feedback for the RO [39] were out of scope. Future work should investigate its effects and potential countermeasures (e.g., for latency, there are possibilities to reduce bandwidth or robust prediction algorithms. See Zhao et al. [39] for an overview).

Furthermore, the driving experience could have affected performance in our task. Future work could consider this in further evaluations.

Finally, some of our assumptions might not hold for realistic implementation (e.g., the maximum 90° internal angle for waypoints might need to be reduced). Implementations using real vehicles will need to evaluate this in the future.

## 6 Conclusion

This work compared three interaction techniques for the remote operation of AVs: *path planning*, *waypoint guidance*, and *trajectory guidance*. We tested these with one to four parallel requests. The designed simulator can be re-used for multiple future use cases and will be provided to interested researchers. In a within-subjects study with N=23 participants, we found that *path planning* was clearly preferred and led to the best overall results. Interestingly, participants were most satisfied with two parallel requests, showing that multiple parallel requests, at least under study conditions, are possible. We discuss these findings in light of necessary AV capabilities and considerations for multiple request handling. Our work helps to safely introduce AVs into general traffic despite potential technical limitations.

## Open Science

The Unity scenario will be available upon request. This includes installation instructions and information on required 3rd party Unity assets. Anonymized data and evaluation scripts are available at <https://github.com/M-Colley/roads-chi25-data>.

## Acknowledgments

We thank all study participants. Mark Colley would also like to thank Johannes Schöning for the invitation to the supportive *CHI-Together* writing retreat.

## References

- [1] Jonas Andersson, Daban Rizgari, Mikael Söderman, and Johanna Vännström. 2023. Exploring remote operation of heavy vehicles—findings from a simulator study. *Intelligent Human Systems Integration 2023* 69 (2023), 46.
- [2] Jonas Andersson and Mikael Söderman. 2024. Navigating the challenges of remote operations of automated road vehicles: a socio-technical perspective. *Intelligent Human Systems Integration (IHSI 2024)* 119 (2024).
- [3] Daniel Bastos, Paulo P. Monteiro, Arnaldo S. R. Oliveira, and Miguel V. Drummond. 2021. An Overview of LiDAR Requirements and Techniques for Autonomous Driving. In *2021 Telecoms Conference (ConfTELE)*. IEEE, New York, NY, USA, 1–6. doi:10.1109/ConfTELE50222.2021.9435580
- [4] Daniel Bogdoll, Stefan Orf, Lars Tötzel, and J. Marius Zöllner. 2022. Taxonomy and Survey on Remote Human Input Systems for Driving Automation Systems. doi:10.1007/978-3-030-98015-3\_6 arXiv:2109.08599 [cs].
- [5] David Brecht, Nils Gehrke, Tobias Kerbl, Niklas Krauss, Domagoj Majstorović, Florian Pfab, Maria-Magdalena Wolf, and Frank Diermeyer. 2024. Evaluation of teleoperation concepts to solve automated vehicle disengagements. *IEEE Open Journal of Intelligent Transportation Systems* (2024).
- [6] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [7] Mark Colley, Lukas Gruler, Marcel Woide, and Enrico Rukzio. 2021. Investigating the Design of Information Presentation in Take-Over Requests in Automated Vehicles. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (Toulouse & Virtual, France) (*MobileHCI '21*). ACM, New York, NY, USA, Article 22, 15 pages. doi:10.1145/3447526.3472025
- [8] Mark Colley, Bastian Wankmüller, and Enrico Rukzio. 2022. A Systematic Evaluation of Solutions for the Final 100m Challenge of Highly Automated Vehicles. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (Sept. 2022), 178:1–178:19. doi:10.1145/3546713
- [9] Mary Cummings, Songpo Li, Dev Seth, Matthew Seong, et al. 2021. *Concepts of operations for autonomous vehicle dispatch operations*. Technical Report. Collaborative Sciences Center for Road Safety.
- [10] Matthew D Dean and Kara Kockelman. 2022. Our self-driving future will be shaped by policies of today. *Nature Electronics* 5, 1 (2022), 2–4.
- [11] Daniel J Fagnant and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77 (2015), 167–181.
- [12] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.
- [13] Michael Gafert, Alexander G. Mirmig, Peter Fröhlich, Vanessa Kraut, Zoja Anzur, and Manfred Tscheligi. 2023. Effective remote automated vehicle operation: a mixed reality contextual comparison study. *Personal and Ubiquitous Computing* 27, 6 (01 Dec 2023), 2321–2338. doi:10.1007/s00779-023-01782-5
- [14] Michael Gafert, Alexander G. Mirmig, Peter Fröhlich, and Manfred Tscheligi. 2022. TeleOperationStation: XR-Exploration of User Interfaces for Remote Automated Vehicle Operation. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. ACM, New York, NY, USA, 1–4. doi:10.1145/3491101.3519882
- [15] Gaetano Graf and Heinrich Hussmann. 2020. User Requirements for Remote Teleoperation-based Interfaces. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '20)*. ACM, New York, NY, USA, 85–88. doi:10.1145/3409251.3411730
- [16] Gaetano Graf, Hao Xu, Dmitriy Schitz, and Xiao Xu. 2020. Improving the Prediction Accuracy of Predictive Displays for Teleoperated Autonomous Vehicles. In *2020 6th International Conference on Control, Automation and Robotics (ICCAR)*. 440–445. doi:10.1109/ICCAR49639.2020.9108011 ISSN: 2251-2446.
- [17] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, Amsterdam, The Netherlands, 139–183.
- [18] Naohisa Hashimoto, Yanbin Wu, and Toshihisa Sato. 2022. Human Factor Issues in Remote Operator of Automated Driving System for Services - One Operator to N Units of Automated Vehicles in Automated Vehicle Services -. In *HCI International 2022 - Late Breaking Posters (Communications in Computer and Information Science)*, Constantine Stephanidis, Margherita Antona, Stavroula Ntoa, and Gavriel Salvendy (Eds.). Springer Nature Switzerland, Cham, 368–374. doi:10.1007/978-3-031-19682-9\_47
- [19] Satish Jeyachandran. 2020. Introducing the 5th-generation Waymo Driver: Informed by experience, designed for scale, engineered to tackle more environments. <https://blog.waymo.com/2020/03/introducing-5th-generation-waymo-driver.html> [Online; accessed 03-DECEMBER-2023].
- [20] Lei Kang, Wei Zhao, Bozhao Qi, and Suman Banerjee. 2018. Augmenting Self-Driving with Remote Control: Challenges and Directions. In *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications (HotMobile '18)*. ACM, New York, NY, USA, 19–24. doi:10.1145/3177102.3177104
- [21] Raheleh Kari and Martin Steinert. 2021. Human Factor Issues in Remote Ship Operations: Lesson Learned by Studying Different Domains. *Journal of Marine Science and Engineering* 9, 4 (April 2021), 385. doi:10.3390/jmse9040385 Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- [22] Carmen Kettwich, Andreas Schrank, Hüseyin Avsar, and Michael Oehl. 2022. A Helping Human Hand: Relevant Scenarios for the Remote Operation of Highly Automated Vehicles in Public Transport. *Applied Sciences* 12, 9 (Jan. 2022), 4350.

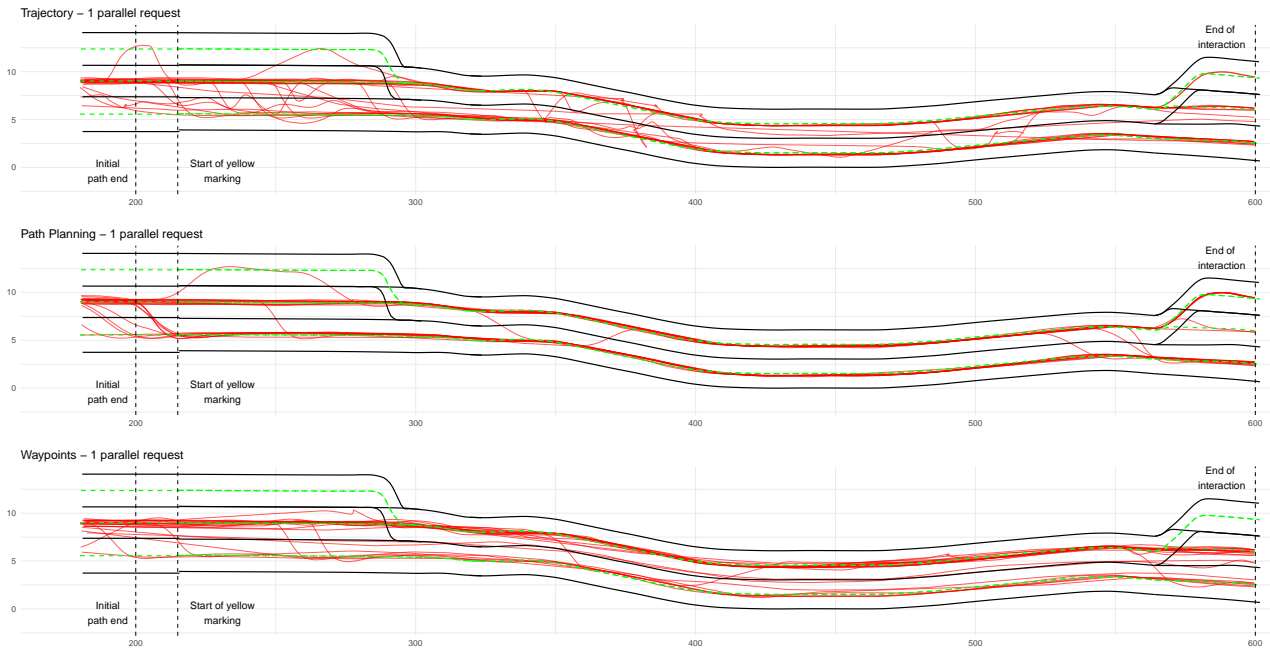
<sup>5</sup><https://assetstore.unity.com/packages/tools/behavior-ai/polarith-ai-pro-movement-with-3d-sensors-71465>; Accessed on 08.12.2024

<sup>6</sup><https://assetstore.unity.com/packages/tools/behavior-ai/mobile-traffic-system-194888>; Accessed on 14.08.2024

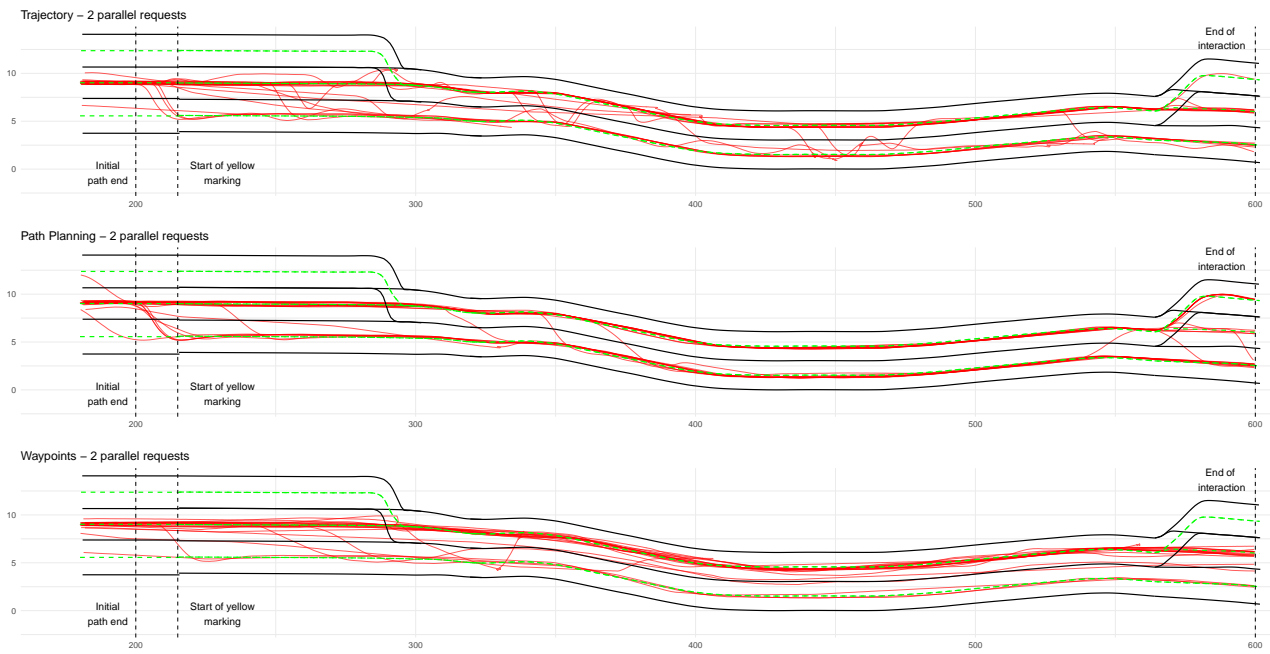
- doi:10.3390/app12094350 Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.
- [23] Carmen Kettwich, Andreas Schrank, and Michael Oehl. 2021. Teleoperation of Highly Automated Vehicles in Public Transport: User-Centered Design of a Human-Machine Interface for Remote-Operation and Its Expert Usability Evaluation. *Multimodal Technologies and Interaction* 5, 5 (May 2021), 26. doi:10.3390/mti5050026 Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- [24] Maximilian Kneissl, Sebastian vom Dorff, Adam Molin, Maxime Denniel, Tong Duy Son, Nicolas Ochoa Lleras, Hasan Esen, and Sandra Hirche. 2020. Mixed-Reality Testing of Multi-Vehicle Coordination in an Automated Valet Parking Environment. *IFAC-PapersOnLine* 53, 2 (2020), 17564–17571. doi:10.1016/j.ifacol.2020.12.2669 21st IFAC World Congress.
- [25] Liangkai Liu, Sidi Lu, Ren Zhong, Baofu Wu, Yongtao Yao, Qingyang Zhang, and Weisong Shi. 2021. Computing Systems for Autonomous Driving: State of the Art and Challenges. *IEEE Internet of Things Journal* 8, 8 (2021), 6469–6486. doi:10.1109/JIOT.2020.3043716
- [26] Domagoj Majstorović, Simon Hoffmann, Florian Pfab, Andreas Schimpe, Maria-Magdalena Wolf, and Frank Diermeyer. 2022. Survey on Teleoperation Concepts for Automated Vehicles. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, New York, NY, USA, 1290–1296. doi:10.1109/SMC53654.2022.9945267
- [27] Aarian Marshall. 2023. Robotaxis Can Now Work the Streets of San Francisco 24/7. <https://www.wired.com/story/robotaxis-cruise-waymo-san-francisco/>. [Online; accessed 12-AUGUST-2023].
- [28] Santhanakrishnan Narayanan, Emmanouil Chaniotakis, and Constantinos Antoniou. 2020. Shared autonomous vehicle services: A comprehensive review. *Transportation Research Part C: Emerging Technologies* 111 (2020), 255–293. doi:10.1016/j.trc.2019.12.008
- [29] SAE. [n.d.]. J3016\_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International. [https://www.sae.org/standards/content/j3016\\_202104/](https://www.sae.org/standards/content/j3016_202104/)
- [30] Andreas Schrank and Carmen Kettwich. 2021. Roles in the Teleoperation of Highly Automated Vehicles in Public Transport.
- [31] Andreas Schrank, Fabian Walocha, Stefan Brandenburg, and Michael Oehl. 2024. Human-centered design and evaluation of a workplace for the remote assistance of highly automated vehicles. *Cognition, Technology & Work* (2024), 1–24.
- [32] Hanbyul Seo, Ki-Dong Lee, Shinpei Yasukawa, Ying Peng, and Philippe Sartori. 2016. LTE evolution for vehicle-to-everything services. *IEEE Communications Magazine* 54, 6 (June 2016), 22–28. doi:10.1109/MCOM.2016.7497762 Conference Name: IEEE Communications Magazine.
- [33] Zhenyu Song, Longfei Shangguan, and Kyle Jamieson. 2017. Wi-Fi Goes to Town: Rapid Picocell Switching for Wireless Transit Networks. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM '17)*. ACM, New York, NY, USA, 322–334. doi:10.1145/3098822.3098846
- [34] Felix Tener and Joel Lanir. 2022. 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 (CHI '22)*. ACM, New York, NY, USA, 1–13. doi:10.1145/3491102.3501827
- [35] Jinke D. Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1 (1997), 1–10. doi:10.1016/S0968-090X(96)00025-3
- [36] VINNOVA. 2024. HAVOC - Heavy Vehicle Operation Centre. <https://www.vinnova.se/globalassets/mikrosajter/ffi/dokument/slutrappporter-ffi/trafiksakerhet-och-automatiserade-fordon-rapporter/2020-02953eng.pdf?cb=20220705105235>
- [37] Jaap Vreeswijk, Azra Habibovic, Olav Madland, and Floris Hooft. 2023. Remote Support for Automated Vehicle Operations. In *Road Vehicle Automation 9 (Lecture Notes in Mobility)*, Gereon Meyer and Sven Beiker (Eds.). Springer International Publishing, Cham, 149–164. doi:10.1007/978-3-031-11112-9\_12
- [38] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (CHI '11). ACM, New York, NY, USA, 143–146. doi:10.1145/1978942.1978963
- [39] Lin Zhao, Mikael Nybacka, Maytheewat Aramrattana, Malte Rothhämel, Azra Habibovic, Lars Drugge, and Frank Jiang. 2024. Remote Driving of Road Vehicles: A Survey of Driving Feedback, Latency, Support Control, and Real Applications. *IEEE Transactions on Intelligent Vehicles* (2024).

## A Appendix - Movement Patterns

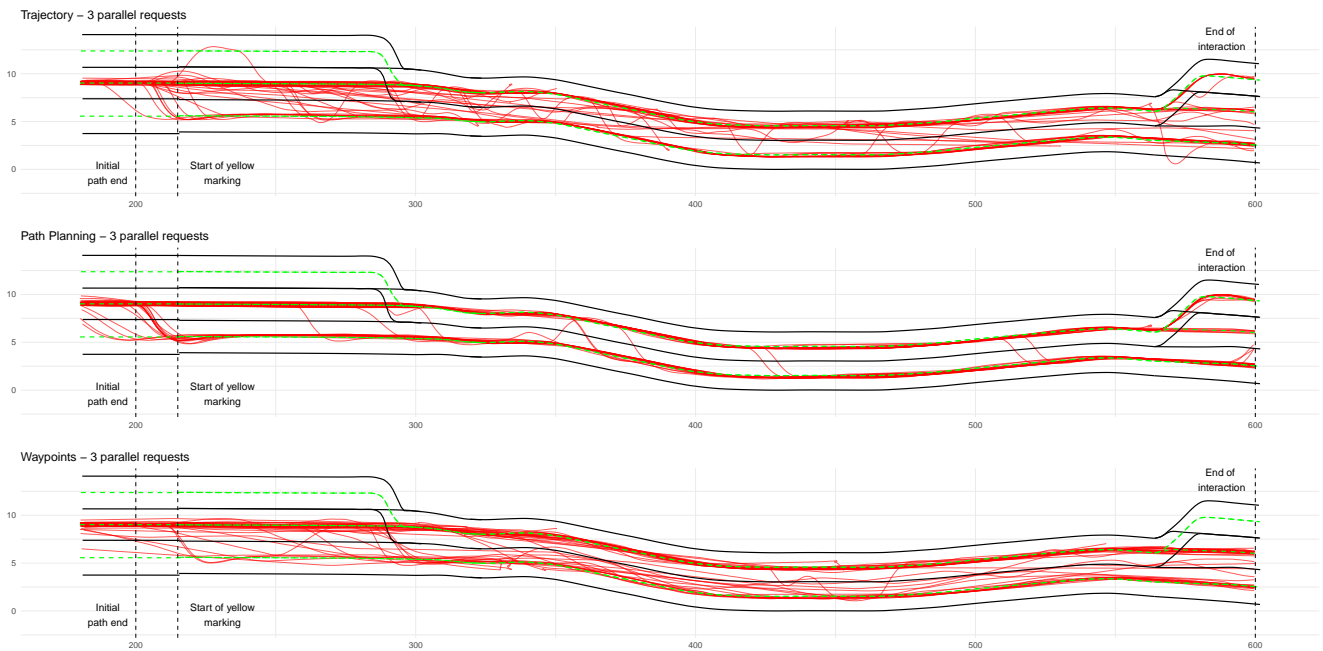




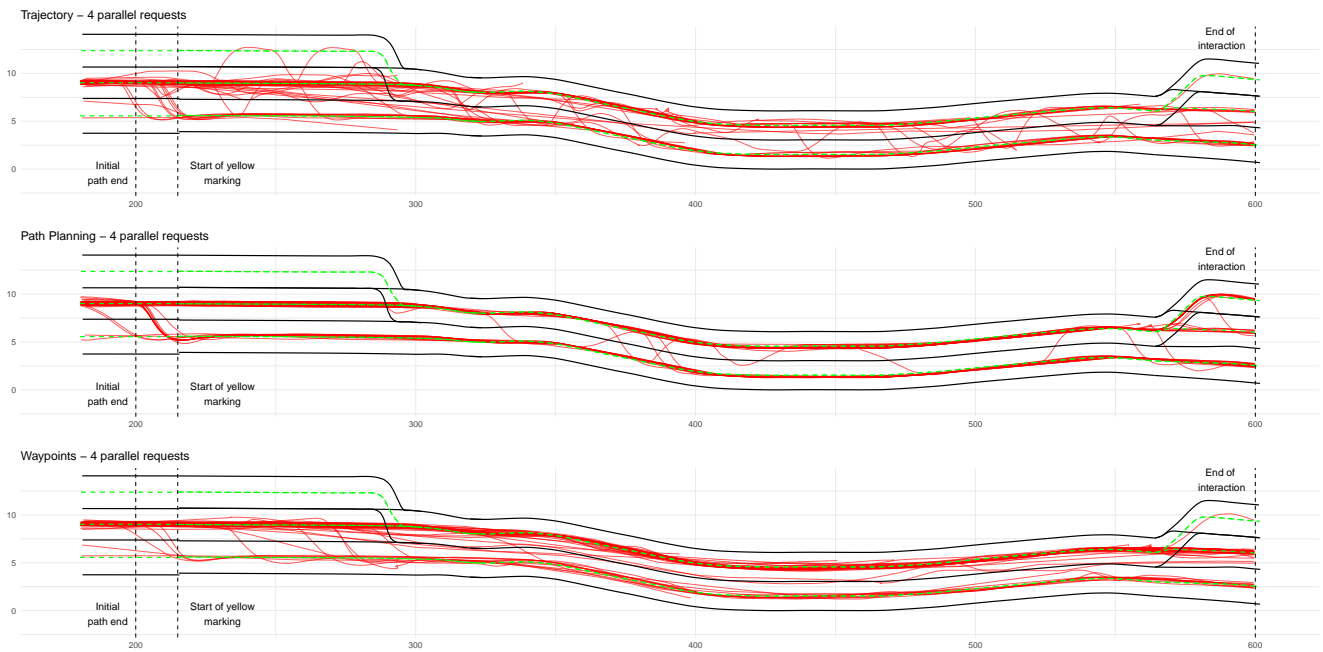
(a) One request.



(b) Two parallel requests.

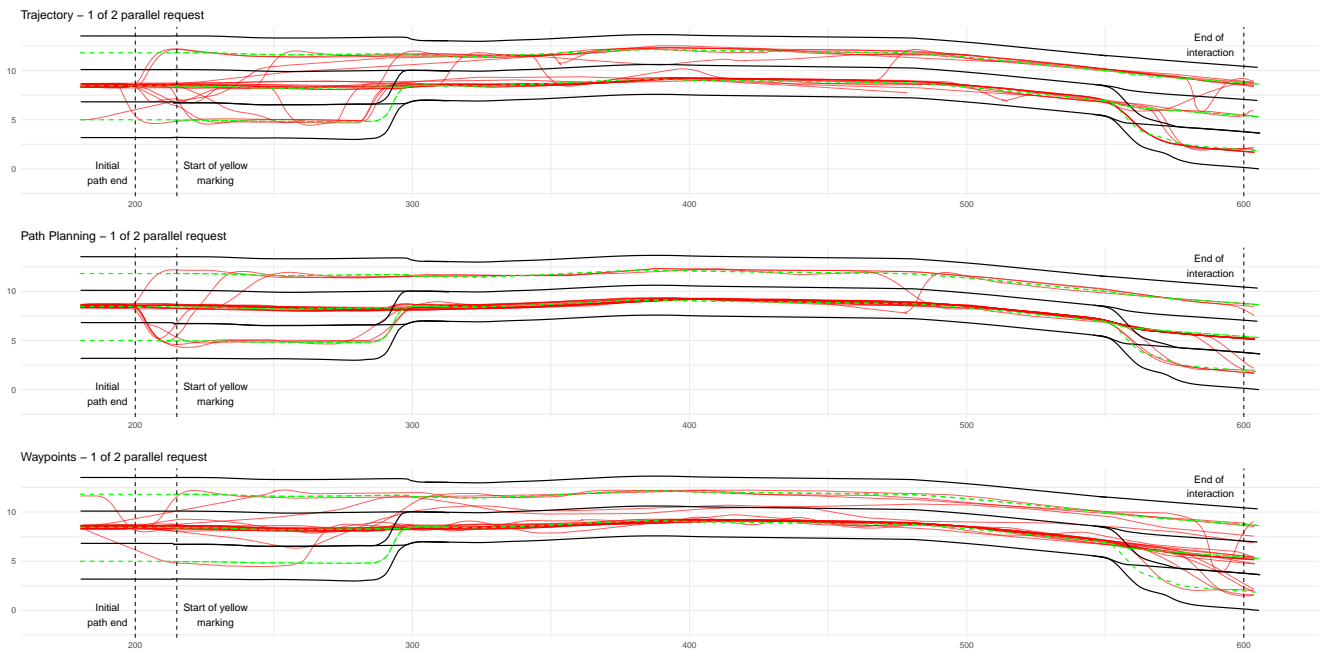


(c) Three parallel requests.

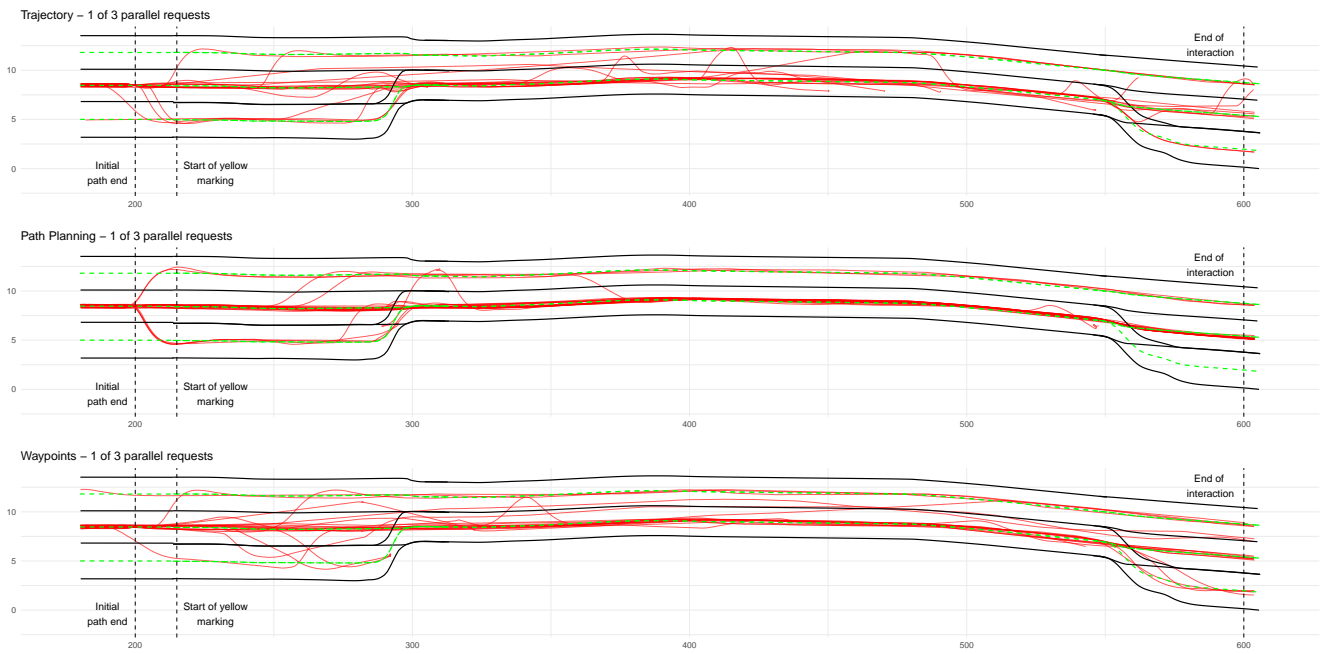


(d) Four parallel requests.

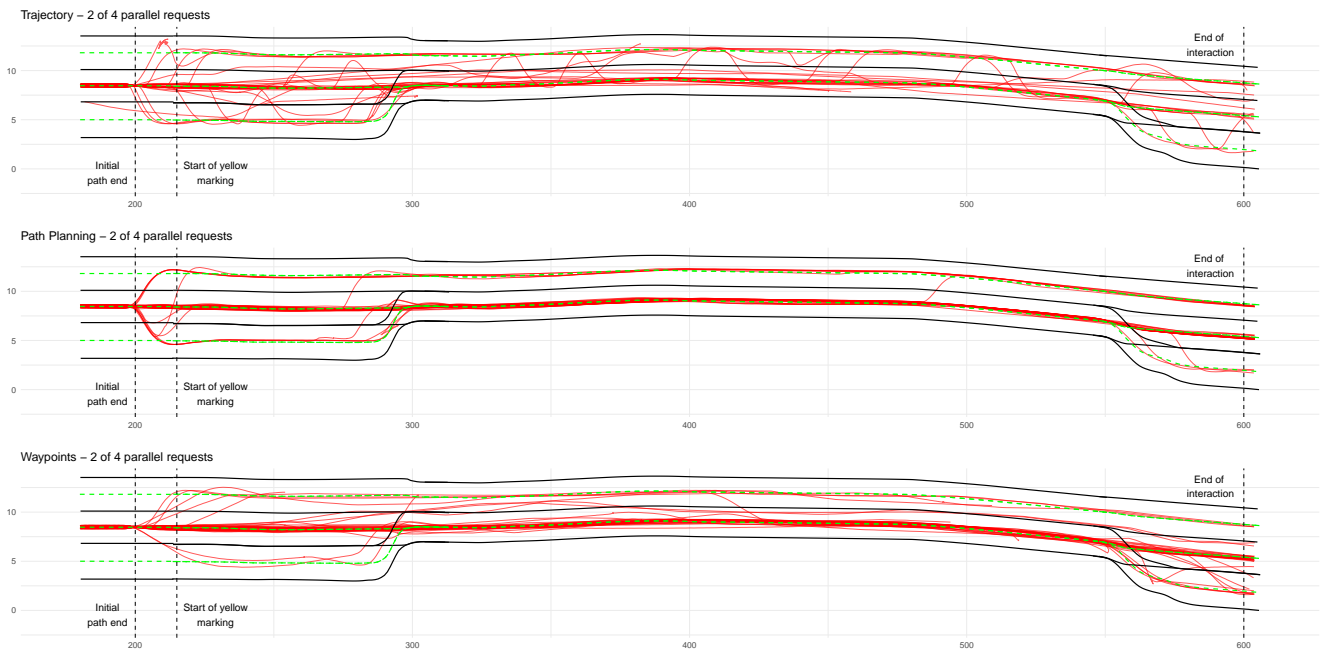
Figure 16: Overview of movement patterns for the left-sided road works.



(a) Two parallel requests.



(b) Three parallel requests.



(c) Four parallel requests.

Figure 17: Overview of movement patterns for the right-sided (mirrored) road works.