COSMIC TROUBLE SHOOTING

Designing Supervised Autonomy for Astronaut - Robot Coaction in Space

Master Thesis | Design for Interaction | Liliane Filthaut | July 2023

Designing Supervised Autonomy for Astronaut – Robot Coaction in Space

Author
Liliane Filthaut

Master Thesis
Delft University of Technology
MSc. Design for Interaction
Faculty of Industrial Design Engineering
German Aerospace Center(DLR)

July 2023

Supervisory team

Chair
Dr. Dave Murray-Rust

Mentor
Maria Luce Lupetti

Company coach
Dr.-Ing. Daniel Leidner

Preface

This master's thesis explores the intersection of design and space exploration, focusing on the development of interactive technologies for enhancing communication and addressing the complex challenges of error handling. The motivation behind this project originates from a lifelong interest in both science and design, which has shaped my perspective as a designer who strives to make communication effortless and bridge the gap between complex contexts and users.

During my bachelor's thesis on designing physics education, where I created an interactive platform for improving the comprehension of the subject, I had the honor of receiving an award for my work. The German Aerospace Center (DLR) was present at the event and took notice of my project, leading to an exciting collaboration that resulted in this master's thesis. This fortuitous connection served as a convergence of my passions for science, space-related topics, and interactive technology design.

Therefore, I would like to express my sincere gratitude to the German Aerospace Center (DLR) for providing me with the opportunity to work on such a fascinating project related to ongoing research in space and Mars exploration. It is a subject that resonates deeply with me, and I am grateful for their support and encouragement. More specifically I would like to thank Dr.-Ing. Daniel Leidner, my company supervisor, for always making time for me and bringing so much enthusiasm to the topic. I am particularly appreciative that he embraced my design-oriented approach for this project, acknowledging the value it brings, despite the traditionally engineering- and robotic-centric nature of the field. Additionally, I want to acknowledge Nesrine Batti for her supportiveness and friendship during my time at the DLR.

Furthermore, I would like to extend my sincere appreciation to my supervisors Dr. Dave Murray-Rust (Chair) and Maria Luce Lupetti (Mentor) for their invaluable support and guidance throughout the entirety of this thesis. Their expertise and insights have been instrumental in shaping the direction of my research and refining my ideas. I am grateful for their open-mindedness and willingness to allow me the freedom to explore and pursue this project in a way that aligned with my vision. Their trust in my abilities and their respect for my independence have been truly empowering. Both their mentorship and encouragement extended beyond the scope of this thesis, and I am grateful for their continued support in my academic and professional journey.

In addition, I am indebted to the robotics, engineers, and researchers at the DLR who actively and generously participated in my research sessions. Their willingness to share their expertise and invest their valuable time in my project was priceless. Their contributions significantly enriched the research process and enhanced the overall quality of this work.

Lastly, I would like to express my appreciation to my friends and family for their unwavering support and understanding. Their encouragement and occasional reminders to take breaks provided much-needed perspective and helped me maintain a fresh outlook on the topic, ultimately enhancing the quality of my work.

This thesis represents the culmination of an enriching and rewarding journey, and I am honored to have had the opportunity to contribute to the field of space exploration through the lens of design.

Method Glossary

UEQ (User Experience Questionnaire): The UEQ scales offer a holistic evaluation of user experience, capturing various dimensions including traditional usability factors such as efficiency, perspicuity, and dependability, as well as experiential elements originality and stimulation [102]

NASA TLX (Task Load Index): The NASA TLX is a subjective tool used to assess mental workload during task performance. It enables the measurement of a participant's mental workload across various dimensions: mental demand, physical demand, temporal demand, effort, performance, and frustration level. By evaluating performance in these dimensions, an overall workload rating can be determined [35].

Premo (Emotion Measurement Instrument): PrEmo is an instrument used for non-verbal self-report that can measure seven positive and seven negative emotions. It has a distinctive advantage as it can measure separate emotions including mixed emotions, and it can be utilized globally since it doesn't require respondents to express their emotions verbally.

Godspeed: The Godspeed questionnaire is a useful tool for measuring the perception of service robots. Five key concepts have been identified in literature to measure robot perception in HRI, namely anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. To create the Godspeed questionnaire, these aspects have been distilled into five consistent questionnaires that use semantic differential scales [12]

Executive Summary

Space exploration has historically driven technological innovation, resulting in significant advancements with applications in various industries on Earth. The exploration of Mars has emerged as a crucial objective, offering opportunities to search for signs of life and gain insights into planetary evolution. However, working on Mars presents numerous challenges, including the lack of a breathable atmosphere, different gravity, extreme temperatures, and the need for infrastructure development. Robots have proven resilient on the Martian surface but still require real-time control and decision-making from ground operators. To address this challenge, a promising solution involves astronauts in orbit around Mars controlling the robots, utilizing high-bandwidth communication techniques and autonomous capabilities.

This work focuses on the Surface Avatar project, led by the DLR and ESA, which involves the humanoid robot, Rollin Justin. The project aims to gain valuable insights into the efficient control of robots in future space missions, particularly through collaborative exploration and construction tasks. Rollin Justin, equipped with autonomous capabilities, features a user interface that allows for both manual controls through various input devices and autonomous operation through interface commands.

While the concept of an astronaut-robot pairing shows promise, several challenges remain. Error handling during teleoperation poses a significant issue, as error messages often lack specificity, leaving astronauts confused and without immediate assistance due to the distance between Earth to Mars and the associated communication delays. Limited situational awareness, unfamiliarity with robot constraints, and a large time gap between training and usage further complicate astronaut interactions with the robot. Addressing these problems is critical for optimizing astronaut-robot cooperation and reducing cognitive workload during Mars missions.

To address the challenges associated with error handling during teleoperation in the context of astronaut-controlled robots, this work adopts a research-through-design approach, with a specific focus on user experience research and design. Extensive initial research including sessions at the DLR and literature review, was conducted to identify key issues impacting error-handling capabilities. Based on the research findings, conceptual solutions were developed to address the identified core issues. These concepts were evaluated for feasibility and desirability, considering expert input. Selected concepts were further developed, drawing inspiration from game cues and elements for user interface design.

High-fidelity prototypes were created to represent the refined concepts accurately: A third-person perspective including game elements to allow for better situational awareness and a debug page that guides the user through potential error reasons in the moment of an occurring planning error. The prototypes underwent evaluation using various methods, including user sessions at the DLR and a comparative study.
The results for both prototypes reveal important enhancements in user experience and a reduction in cognitive workload compared to the existing system. The findings led to informed recommendations for further improvements in the interface design, the robot's camera setup and the communication of errors to enhance error-handling capabilities for astronauts in future missions.

This project showcases the integration of design choices in domains like space exploration, emphasizing their importance for enhancing user experience and system handling. It hopes to highlight the value of collaboration between design and research disciplines, demonstrating the positive outcomes that arise from incorporating design principles in traditionally non-design-focused areas.

Contents

Intro-
duction

This section provides an introduction to the topic of the teleoperation of robots in space, highlighting the problem of dealing with errors and the importance of design in addressing this challenge. It explores the project partner and actor ecosystem involved in this work and outlines the approach taken, offering a high-level overview of the project's process.

Robots for Space

Why We Explore Space
Space exploration has been a historical driver of technological innovation, leading to significant advancements. The need for robust and reliable technology to withstand the harsh conditions of space has pushed scientists and engineers to develop new materials, techniques, and technologies that have found applications in many industries on Earth. For example, satellite technology originally developed for space exploration has revolutionized modern communication, enabling instant global communication, weather forecasting, and GPS navigation [83].

Exploring Mars: A Difficult Task for Humans
One of the primary goals of space exploration today is to study Mars. Exploring Mars is important because it can help mankind search for signs of life beyond Earth and gain insights into planetary evolution, which in turn can contribute to our understanding of our own planet and prepare for future human exploration [30]. However, working on mars is difficult due to the lack of breathable atmosphere, difference in gravity to earth and extreme freezing thermal conditions all adding to the mental and physical load of the astronauts [70, 92]. Consequently, before astronauts can ascend to mars the construction of infrastructure is required [56, 89, 90, 101].

Robots Can Help, But to What Extend?
Robots, on the other hand, have demonstrated remarkable resilience on the Martian surface, enduring the challenging conditions for extended periods of time. Spacecrafts in orbit, landers, and rovers have been conducting extensive investigations of Mars, covering considerable distances, deploying a wide range of instruments and collecting surface samples [84]. Therefore, robots are envisaged to assume the physical tasks involved in establishing crew habitats and providing energy resources [56, 89, 90, 101].

A remaining difficulty with these robots is that their operation still heavily relies on real-time control and decision-making from operators on the ground for day-to-day operations [84]. Operating these robots presents a formidable challenge, including time delays of up to 45 minutes to send and receive an answer [56, 90] and potential communication disruptions along the link between Earth and the robots, making communication unreliable [90].

The Proposal:
A Cooperation Between Astronauts and Robots
A promising solution to address this challenge involves astronauts in orbit around Mars, that control the robots instead of a ground crew. This setup allows the utilization of high-bandwidth communication techniques with less delay, enabling the use of autonomous robots [90].

Autonomy enables the astronaut to control the robot via simple commands, that the robot can execute on its own. For instance, instead of manually using a joystick to navigate to a location one can command the robot via the click of a button like "navigate to base" and the robot will do so autonomously. This, however, is currently limited to simple commands based on the robot reasoning about the world. Moreover, autonomous robots can currently only operate within predetermined domains [88, 90], therefore whenever unknown objects are encountered or the environment changes significantly, teleoperation is still required e.g., manual control via joysticks [88].

In situation where teleoperation is still required, this setup, while having comparably minimal delays of around 800 milliseconds [56], remains challenging. To illustrate, a delay of 10 – 20ms is already noticeable [104]. More delay and it can affect efficient teleoperation, particularly for fine motor or complex tasks negatively [107]. Refer to Figure 1 to see the significant difference in delays.

Therefore, to counter weigh the remaining time-offset as much as possible and reduce cognitive workload the idea of shared autonomy is introduced. Shared autonomy is a hybrid model combining aspects of direct and supervised manipulation, where human operators interact with partially autonomous robots [60]. Therefore, if the conditions allow it, the astronaut assumes a monitoring role for the robot's actions, which reduces their cognitive workload and the need for manual teleoperation. Whenever using the autonomous commands, the remaining time delay can be bypassed [88].

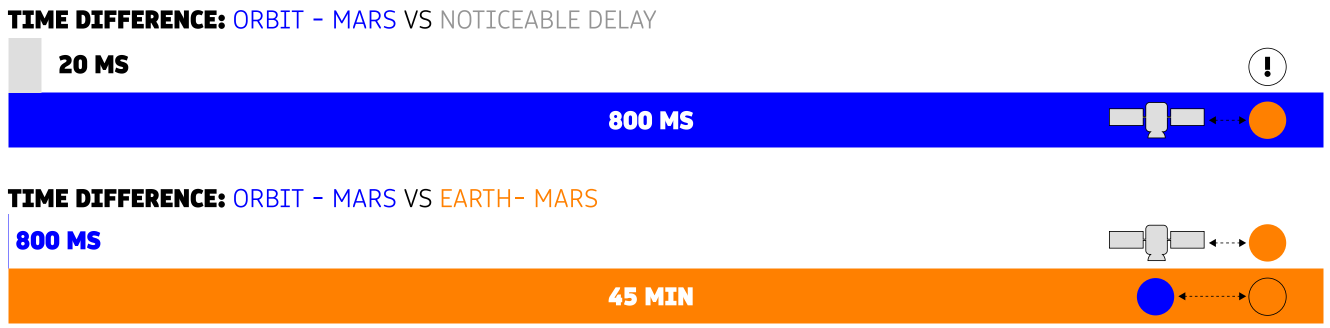


Figure 1: An illustration of time delays in this research context.

Bottom beams: Earth operation vs. astronaut operation in Mars orbit. Top bars: Remaining 800ms delay vs. noticeable 20ms delays.

The Core of This Work:
The Unsolved Matter of Error Handling

A still unresolved aspect of this type of teleoperation is the issue of error handling. In some cases, the robot does not find a solution or fails during execution. Error messages can leave astronauts confused and without instructions, while the communication delay makes seeking immediate assistance nearly impossible, see Figure 2. As of now, Error handling remains a challenge for teleoperation, especially in the field of space exploration [38].

This work focuses on the humanoid robot Rollin Justin from the German aerospace center (DLR). The robot is semi-autonomous, meaning that it can perform certain tasks on its own. The autonomous task can be commanded via a user interface, by clicking on objects that are known to the robot. In addition, teleoperation is available to take over control manually [22, 88]. Rollin Justin is currently being tested for teleoperation on Earth by astronauts on the ISS, to eventually be sent to Mars [101]. It is part of the Surface Avatar project. These experiments aim to reduce the cognitive load of the astronauts as much as possible. However, with situations like error handling, it remains difficult as error causes are as of now not easily deductible. This poses a challenge to astronauts, who must monitor and manage the system, stepping in to correct any errors or issues that arise. This adds on to the high cognitive load of the astronauts, already worn out from emotional and environmental stress of space, making it important to reduce any additional stressors [89]. Therefore, on the example of Rollin Justin, this work focuses on designing user-friendly interventions for planning error scenarios and proposing user experience approaches to the User interface (UI) design of the Surface Avatar system.

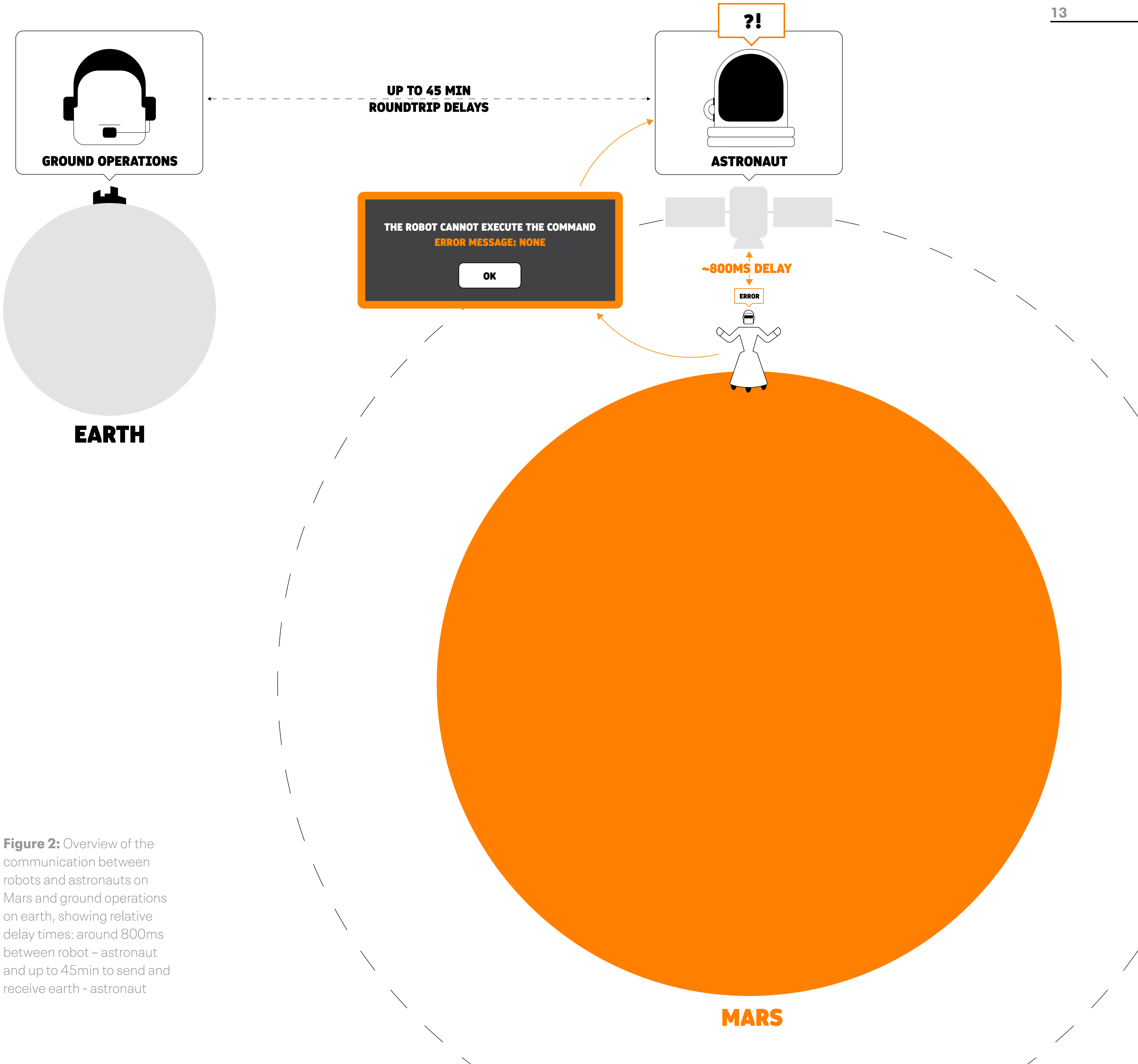


Figure 2: Overview of the communication between robots and astronauts on Mars and ground operations on earth, showing relative delay times: around 800ms between robot – astronaut and up to 45min to send and receive earth - astronaut

The Problem: Dealing with Errors

The scope of this work is to improve interactions between astronauts and robots, particularly around error handling. What seems easy to a human can be extremely difficult for a robot, particularly due to the difference between their world representations and styles of reasoning. This is compounded by communication delays, and a general lack of co-presence.

There are two main categories of errors that Rollin Justin encounters: Planning errors happen if the robot cannot find a plan that will satisfy an operator request. In this case, the robot will stay still for an extended period, and then report that it was unable to comply with the request. The other type are Execution Errors, which happen if the robot has a plan, but encounters difficulties executing it - the robot will begin movement, but stop before the end of the sequence [22, 120].
This work will focus specifically on planning errors, as they are avoidable through better operation and still recoverable via autonomous tasks. On the other hand, execution errors currently necessitate manual teleoperation for resolution, offering limited opportunities for intervention through design modifications.

The main challenges for astronauts when dealing with planning errors are the following:

(1) Lack of feedback and information from the robot: If a planning error occurs, the robot will be unresponsive for several seconds as it attempts to plan, to then provide a generic error such as "Error Message: None". Such messages provide little insight into the problem at hand.

(2) Limited situational awareness: Understanding the robot's perspective can be difficult, as it has a restricted field of view.

(3) Lack of expertise on the robot: The robot has specific constraints related to its kinematics and the environment, such as collision avoidance or reachability [22], that the astronaut may not be aware of. Moreover, the robot can only interact with objects that it accurately recognizes. [56] As a result, it can be challenging for the astronaut to fully comprehend the situation and determine the appropriate course of action.

(4) Large time gap between training sessions and actual robot usage: Astronauts receive initial training on the system and how to operate it, but in many cases, the actual usage may occur six months or more after the training. This time gap can cause the astronauts to forget much of what they have learned, making it crucial for the robot system to be intuitive and easy to use [89].

(5) Restricted external help: While astronauts can request help from ground operations (GO), The substantial time delays make this option highly impracticable [56].

In a nutshell, teleoperating robots is a challenging task that requires specialized training and knowledge [76, 86]. Even in situation where multiple specialized operators for a single robot are present, basic tasks like collision avoidance are still challenging, leading to increased stress levels for the operators [76, 86].

What Can Design Do?

What can design do about these issues? Design choices play a crucial role in our everyday lives, especially in a rapidly advancing technological world. It enables even novice users to navigate and interact with complex systems, such as smartphones, effortlessly, facilitating widespread accessibility and usability. In particular, Human-Centered Design (HCD) helps in developing interactive systems that prioritizes user experience by addressing user needs, requirements, and incorporating human factors and usability principles [43].

The context of teleoperation poses many user experience challenges as most modern teleoperation interfaces are designed for experts and do not consider common usability and learnability standards [109]. In the field of space exploration Interface design is especially restricted. For example, the GUI must follow the ISS Display and Graphics Commonality Standard (DGCS), which includes specific terminology and restrictions on the use of colors designated for ISS subsystems and system conditions. [90]. In general, teleoperation interfaces often have complex, manual controls and high information density that make them difficult to use for non-experts. There simply is a lack of user-centered design [96]. Especially in the context of this report, where astronauts are already confronted by multiple stressors, the need for easier and simpler system interaction becomes apparent. User-focused improvements in usability, learnability, and user experience would benefit any type of user [96], resulting in more accessible teleoperation across a range of applications. Therefore, the inclusion of design principles in the development of teleoperated robots is essential to improve their usability and make them accessible to a wider range of users.

Overall, this work seeks to introduce more user-centered design principles to the area of tele-robotics and showcase how design choices and principles can improve the overall user experience for operators. By addressing planning errors, it aims to make cooperation between astronauts and robots a more seamless and effective process, with the goal to contribute to the success of future Mars missions.

Project Partner and Actor Ecology

The project partner and actor ecology of the study encompasses several key stakeholders involved in the development and implementation of the research, see Figure 3. These stakeholders include the German Aerospace Center (DLR) and the European Space Agency (ESA) and their astronauts. Each of these entities plays a crucial role in shaping the project and its objectives.

The DLR, as a prominent research institution in the field of aerospace, brings extensive expertise and technical resources to the project. Their involvement provides the robot Rollin Justin and the facilities of GSOC (where the testing in cooperation with ESA and the ISS happen). In the experiments they provide help during the astronaut training and function as ground operations during testing with Rollin Justin on earth with the ISS. The partnership with the DLR enables valuable robot testing with their experienced roboticists and the robot Rollin Justin. Their involvement and contribution of knowledge significantly support the goal of this work, which is to enhance error handling for astronauts during future Mars missions. By incorporating their expertise, the DLR plays a vital role in advancing the research and achieving the objective of this work.

The ESA, as a major space agency, plays a pivotal role in shaping the future of space exploration. ESA's involvement facilitates access to astronaut participants, who serve as critical contributors to the research. They further provide the facilities for the astronauts to train in (e.g., Columbus module mockup) and function as a contact partner during the experiments.

Astronauts, being the end-users and primary beneficiaries of the study, hold a central position in the project. Their firsthand experiences and feedback provide invaluable insights into the challenges and requirements of robot-assisted tasks in space environments.

The partnership and collaboration between these entities create a dynamic and multidisciplinary ecosystem, enabling the project to benefit from diverse perspectives and expertise. Through close cooperation and knowledge exchange, the project aims to advance the understanding and development of robotic systems that can effectively support astronauts in space exploration missions.

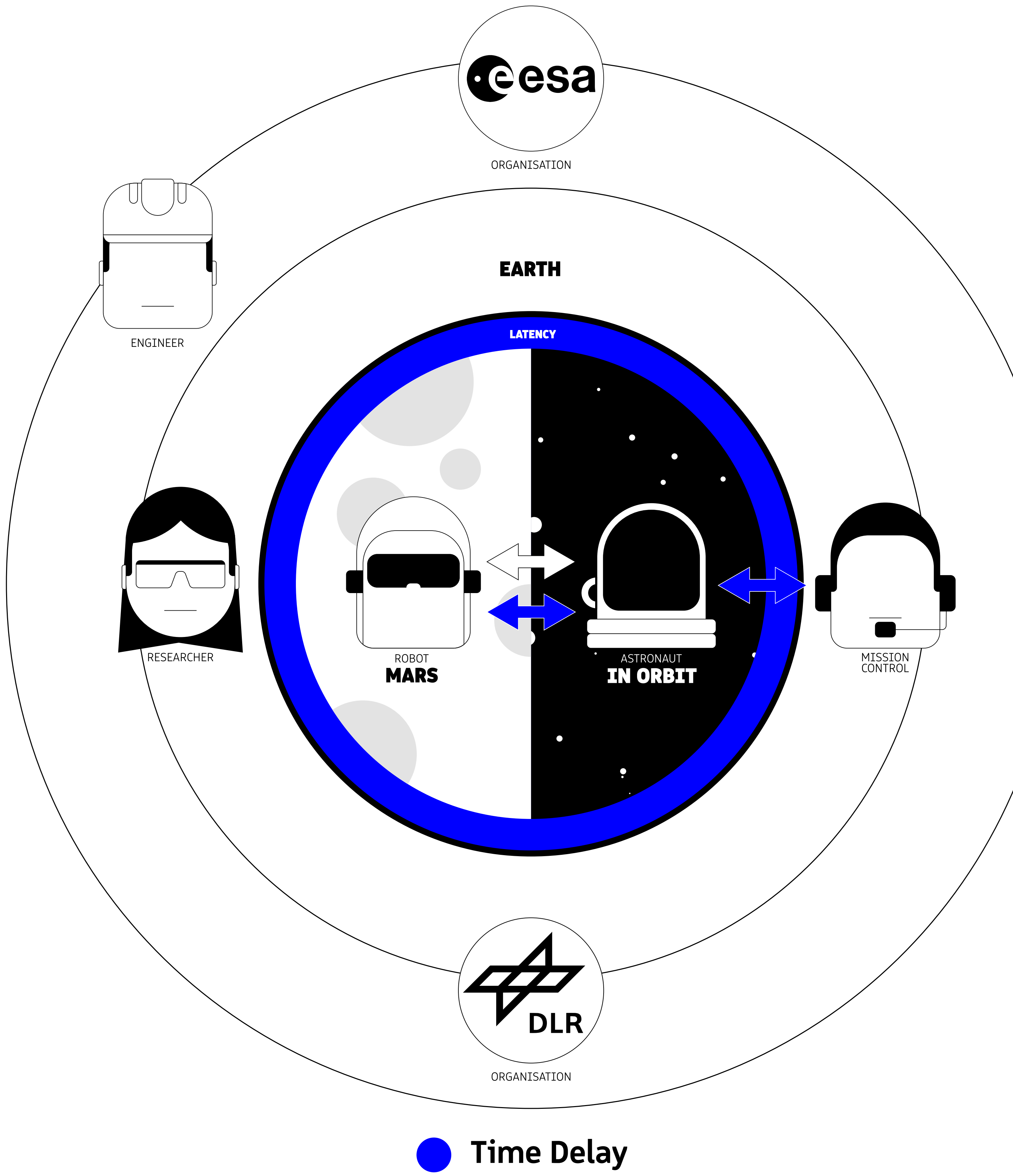


Figure 3: Stakeholder map of involved actors: The robot Rollin Justin, Astronaut operators, ground operations, researchers and engineers working on the robot and involved organizations (ESA and DLR),

Project Approach

This work followed a research through design approach, with a specific focus on user experience research and design. Refer to Figure 4 for an overview of the entire project process.

Initial Research
The initial phase involved extensive research, including sessions at the DLR and a thorough review of relevant literature, to explore the field and identify key issues that could impact participants' error handling capabilities throughout their user journey.

Conception and Selection
Based on the insights gained from the research phase, conceptual solutions were developed to address the identified core issues. These concepts were then evaluated with an expert, considering their feasibility, and expected desirability.

Implementation
Following the concept evaluation, specific research was conducted in the selected concept directions, also drawing inspiration from game cues/elements for user interface design. This domain provided valuable insights as games often involve users who are unfamiliar with the environment, controls, and tasks, similar to the challenges faced in the teleoperation of Mars. Based on that, three high-fidelity prototypes were built to represent the refined concepts accurately.

Evaluation
The prototypes were evaluated through various methods, including user sessions at the DLR and a comparative study. The evaluation aimed to assess the effectiveness and user experience of the prototypes in enhancing error handling capabilities. Based on the comprehensive results obtained from the evaluations, recommendations for further improvements and suggestions were provided. These recommendations are informed by the overall findings of the research and aim to enhance error handling capabilities for astronauts in future missions.

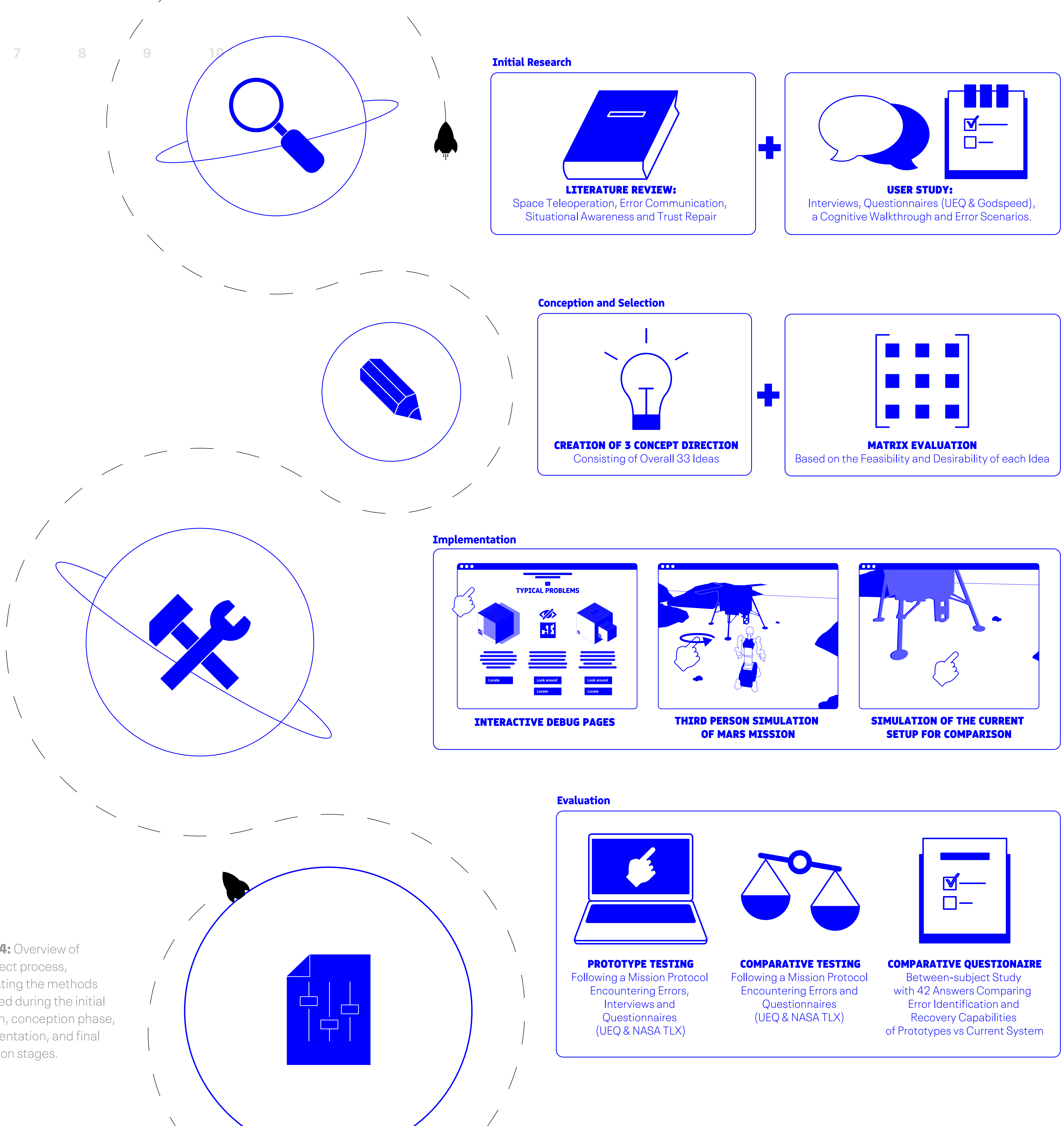


Figure 4: Overview of the project process, highlighting the methods employed during the initial research, conception phase, implementation, and final evaluation stages.

Design Context

This chapter provides a detailed overview of the Survace Avatar Project, including the testing ground, training, experiments protocols, available controls, the robot, the GUI and the target user.

Surface Avatar

This work centers around the humanoid robot Rollin Justin and its role in the Surface Avatar experiments [81]. The Surface Avatar project, scheduled for 2022-2024, is an ISS-to-Earth telerobotic experiment led by DLR and ESA. It builds upon previous telerobotic experiments such as Kontur-2, Haptics, Interact, SUPVIS Justin, and Analog-1, see Related Work, p. 36. The project involves a team of four diverse robots situated in a multi-site analog environment at DLR, controlled by an astronaut on the ISS [81].

In this context multi-site analog environment refers to a simulated setting that replicates certain aspects of the intended operational environment on Mars. The term "multi-site" indicates that there are multiple locations within the analog environment where the robots are deployed, and tasks are performed. This environment aims to mimic the conditions, challenges, and constraints that astronauts and robots would encounter during actual space missions, allowing for realistic testing and evaluation of the robot's capabilities and the interaction between astronauts and robots [81].

The robot team consists of a humanoid robot (Rollin Justin), a lander with a robotic arm for delivering components and storing samples, a quadrupedal robot for exploring challenging terrains and a rover for traversing and collecting samples [81].

By conducting collaborative tasks in exploration and construction scenarios, these experiments aim to gather insights on how to effectively control robots in future space missions [81].

During the writing of this work, the ongoing Surface Avatar experiments have primarily focused on individual robots rather than a multirobot team. As a result, the development of a user interface enabling control of multiple robots is still in progress [81]. Therefore, this work specifically concentrates on the control of the humanoid robot Rollin Justin and does not encompass the control of other robots. However, all other aspects of the experiments remain unchanged.

The following sections provides a detailed overview of the system, including the test ground, training, experiments protocols, available controls, the robot, the GUI and the user.

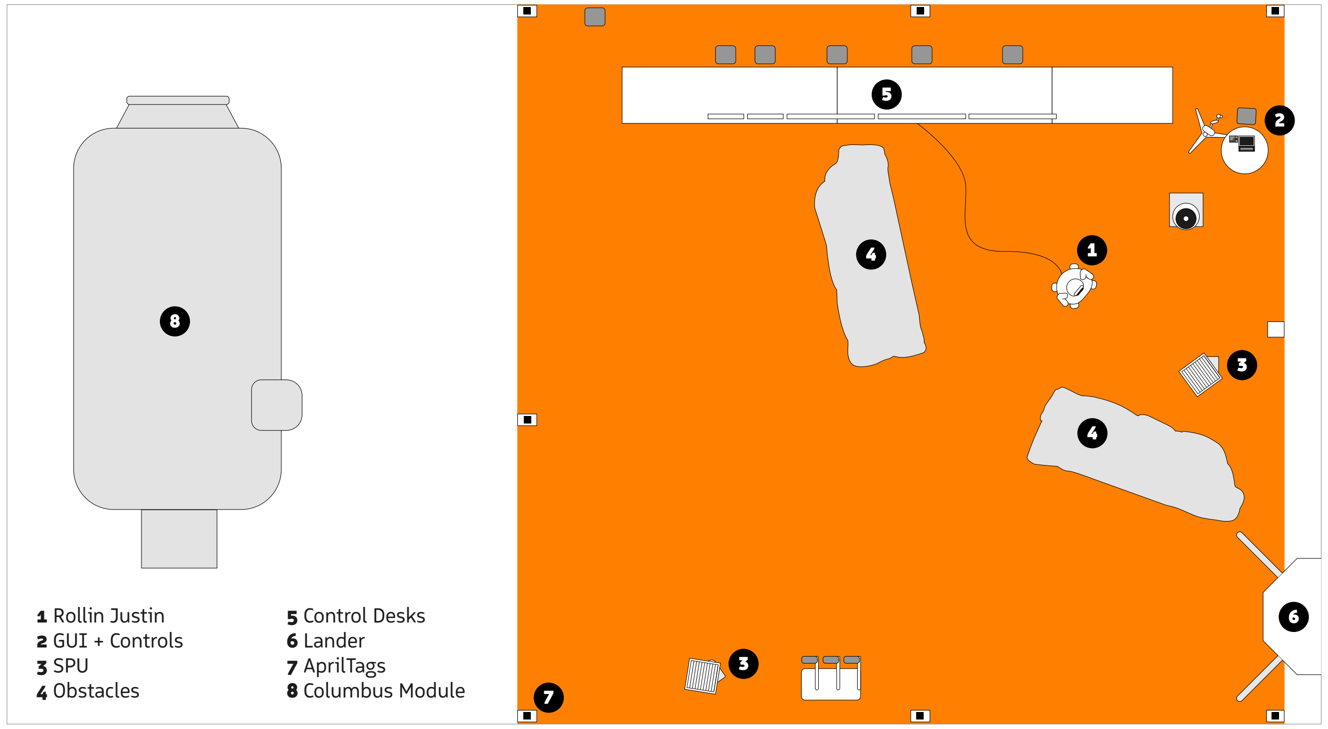
Figure 5: Test ground used for experiments and dry runs, showing: (1) Rollin Justin, (2) Setup of GUI and Controls, (3) Solar Panel Unit (SPU), (4) Obstacles like Stone structures, (5) Control desks uses by Ground Operations during testing, (6) Lander mockup, (7) April Tags on operation floor to indicate bounds, and (8) Columbus Module Mockup

Test Ground

As of now, the current setup of the system is on Earth at the German Space Operations Center (GSOC), at Oberpfaffenhofen near Munich, where it is operated by the German Aerospace Center (DLR). Meanwhile astronauts from the ISS test it in orbit so that it can be further improved and evaluated. These testing try to imitate the real scenario as much as possible [81]. Therefore, the testing environment on earth is made to fit as much as possible mars. Working and non-working replicas of modules that would be on mars are present, the robot Rollin Justin is fully working, the astronauts have to follow typical protocols that would happen on mars, and they have to build up and set up everything from scratch as they would need in the real context [101].
The setup and testing ground can be seen in Figure 5. The big orange area represents a printed mars surface that is to scale of an actual area on mars. It covers roughly 12x11m and therefore around 132m2 that the robot can freely operate in. Outside of the mars surface, a Columbus module replica is located (8).

This replica is where the astronauts teleoperate the robot on the ISS. Since in the testing's astronauts are in the actual Columbus Module in space, in the context of the experiments this module will be used as a habitat mockup.

On the mars surface, the robot is placed (1) together with obstacles (4) and modules for the robot to interact with (3, 6). The modules that the robot can interact with are Solar panel units (see Figure 5 (3), short SPUs) and a part of a lander (6). The obstacles can also be interacted with, allowing the robot to place items on them. Moreover, a version of the system setup (2) is located on the test ground so that research can run dry runs and experiment in smaller scale. The system setup will be explained in more detail under Control System and Robot, p. 26. Additionally, ground operations (5) is located on the test ground so that they can intervene immediately. Ground control on the other hand is located at ESA in Cologne [51].



The Training

Astronauts receive a training before every experiment roughly six months prior. In the training, the astronauts are at ESA in Cologne while the robot is in Munich at the DLR [89]. At ESA, they also have a detailed replica of the ISS's Columbus module. In the training, they receive an introduction to the robot, assets and the experiment scenario presented via slides by another astronaut. They get informed on navigation, the map overview, the need for localization and object representations. After that they move to a Columbus mockup and start building up the hardware: Laptop, joystick and Sigma7. Once everything is set up, they get to do small hands-on experiments directly form the Columbus mockup, operating Justin at the DLR [51]. These exercises are different form the experiments. This is to ensure that the system allows the astronauts to handle new unknown situations, which would not be the case if they got trained on the exact experiment scenario. After the exercises follows an open discussion where the astronauts get to ask questions and can get feedback [51].

Experiment Protocol

At the day of the experiment, astronauts receive a mission protocol that they must follow and try to complete. NASA requires missions that are not too complex, as they need to ensure that the missions are solvable and do not pose unnecessary stress to the astronauts [82]. At the beginning, astronauts have to set up the system and install all controls and the laptop with the interface on the ISS. They have a manual available to them that explains this process step by step and during experiments a constant Voice Communication Subsystem (VoCS) [88] is held with the astronauts so that they can ask questions at any time. Once everything is set up, they have to do a quick check out where they make sure that they have the correct version of the system and verify that all controls work [51]. Then they receive a protocol like the one in Figure 6. In this task they must complete a data inspection with a data interface probe (DIP see Figure 8, p. 28), check the status of the SPU and read out data, install new software, reboot the SPU, and take snapshots of anything they deem noteworthy [51, 56]. During the entire experiment they are in contact with GC and GO. However, due to time delays contact might have higher response times. Moreover, due to the delay, the robot's reaction time is not immediate [51].

You can use this to ask the ground robotics team any questions and receive feedback. The ground robotics team will also use this text messenger to provide task descriptions, so keep an eye on it.

Please command Justin to perform:
-Data inspection using the data interface probe (DIP)
-SPU status and data readout
-Software installation
-SPU reboot
-Take snapshots of any object or situation worth noting

Figure 6: Standard chat text and typical protocol.

Control System and Robot

The system that the astronauts can use to control the robot consists of the GUI, a Joystick with three degrees of freedom including mode Buttons (ABC) and a Sigma 7, see Figure 7. Additionally, they have a fixed handle with a yellow enabling button.

The yellow button needs to be pressed whenever other controls than the GUI are used. Only while the yellow button is pressed actions from the controls will affect the robot [101]. Refer to Figure 8, on the next page, to see an overview of how each control affects the user.

The Joystick allows the user to move the head or the base of the robot. The Sigma 7 allows the user to control the right arm of the robot with exact precision including the opening and closing of the hand. The Sigma7 has a limited range of motion and therefore needs to be reset by disabling and enabling the yellow button [101]. Similar to lifting a computer mouse to the other side of a mousepad. Both controls can be either enabled over the interface by selecting a teleoperation command or over the ABC mode buttons. With (A) for operating the robotic arm via Sigma 7, (B) to drive around via the joystick and (C) to look around via the joystick.
The GUI allows the user to operate the robot via task-level commands. This is also where the autonomy of the robot takes place since it executes commanded actions by itself [101]. Refer to the next paragraph, p. 30 for more details on the GUI. Figure 8 (next page), moreover, showcases Rollin Justin, where the camera for the GUI view is placed and tools that the robot has on himself. The DIP can be used to inspect other objects and the wiper can be used to clean for example solar panel surfaces.

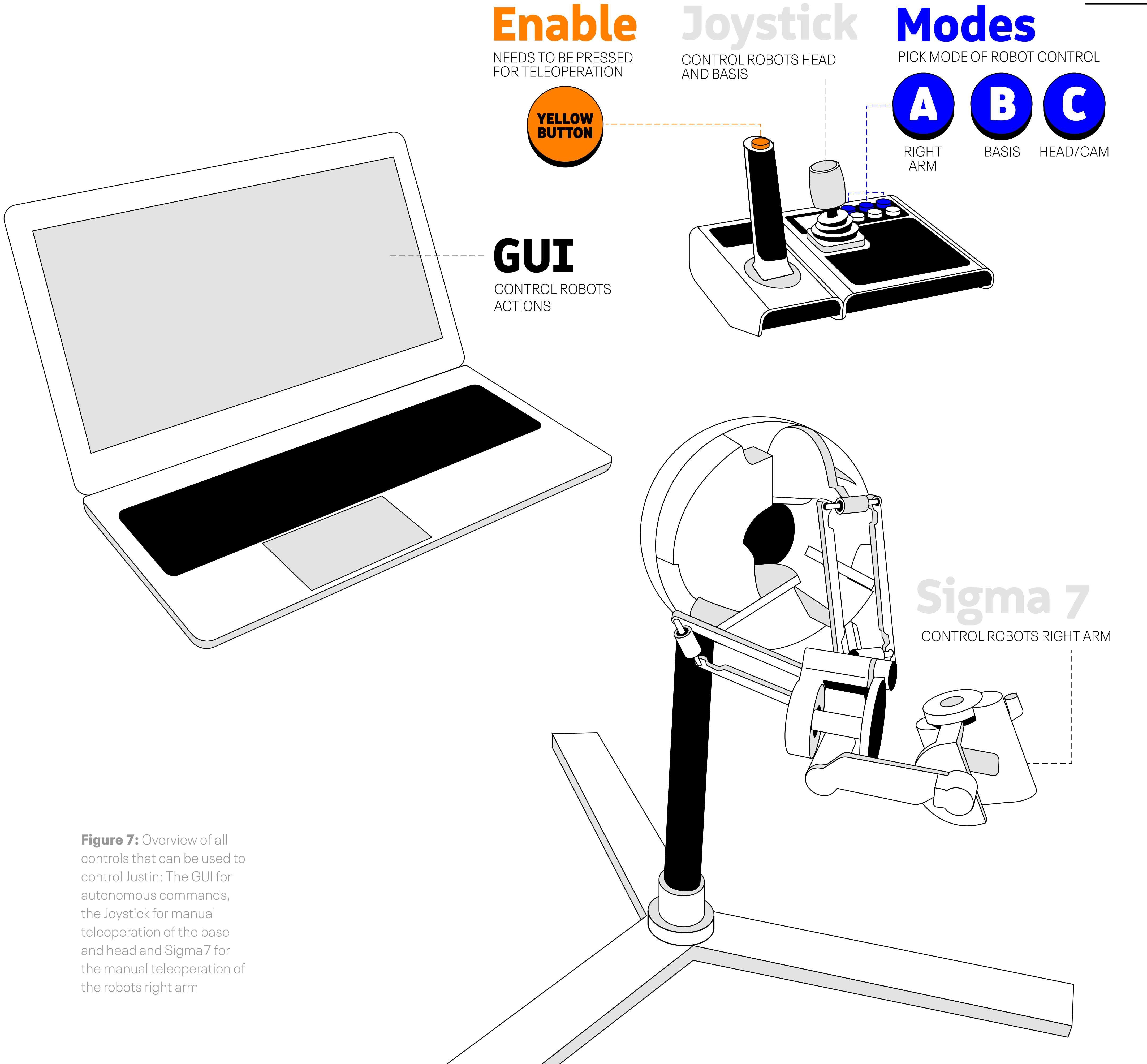
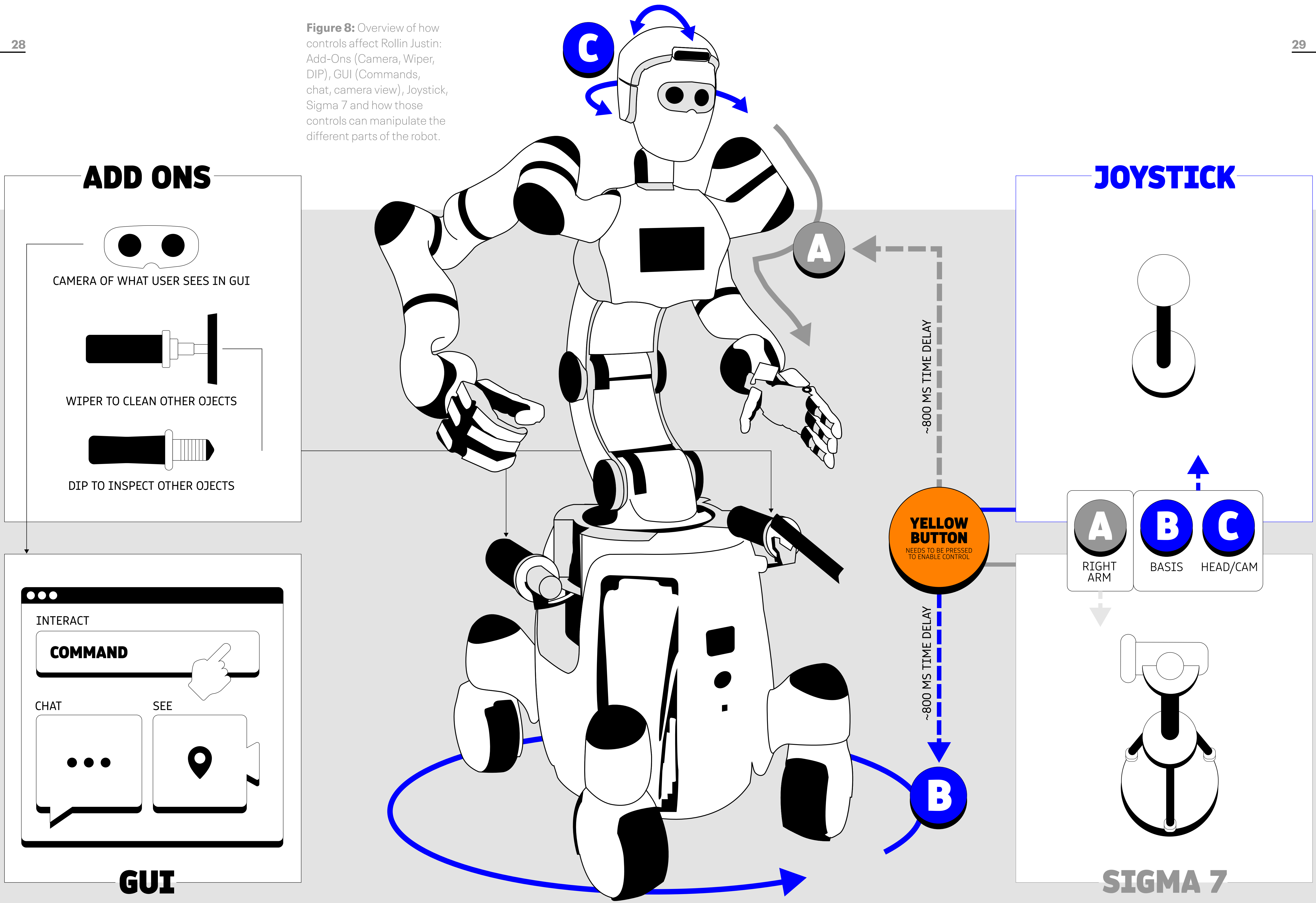


Figure 7: Overview of all controls that can be used to control Justin: The GUI for autonomous commands, the Joystick for manual teleoperation of the base and head and Sigma7 for the manual teleoperation of the robots right arm

Figure 8: Overview of how controls affect Rollin Justin: Add-Ons (Camera, Wiper, DIP), GUI (Commands, chat, camera view), Joystick, Sigma 7 and how those controls can manipulate the different parts of the robot.



The Interface

Figure 9 shows an abstracted version of all the functions of the Interface. This GUI is installed on a laptop that has a touch pad. Generally, the interface shows the robots camera view in the middle of the screen from an ego-centric perspective and has interaction possibilities on the sides. On the left side of the GUI one can find a Chat, a top-view map, system status, help manual, and going back to the main screen.

Elements for Experiment Purposes
Going back to the main screen is just there for the purpose of the experiment as the main screen also contains the questionnaire that the astronaut must fill out after the testing. The help manual is simply a PDF with additional information and the system status gives an overview of their current system but is not used during any of the experiments. These three aspects are not particularly important for this research. On the other hand, the chat and map are used and important.

Chat and Map
The Chat provides the astronaut with the protocol and further functions as the communication tool to ground control. Considering that during experiments a constant voice-loop is held, this chat does not get used much but becomes crucial for ground communication in the real scenario as the time delay would make a voice connection impossible. The map gives an overview of objects in the close environment that are not in the direct camera view of the robot. It updates after every interaction and the robot's representation also contains rotation of the body, hands and head including what it is holding in its hands. If one clicks in the map it will enlarge over the interface. Then one can click into the map and command the robot to navigate there, use a slider to rotate the base of the robot or simply get feedback on the robot's position in the environment. Additionally, the map contains information on the camera angle of the robot's view.

Camera View and Overlays
In the middle of the interface, one can see the camera view of Justin. The camera view is augmented with overlays for all the objects that the robot can interact with. However, for the robot to be able to produce the overlay, AprilTags on the objects need to be visible. The overlays represent what the robot perceives from its surroundings. Therefore, they do not perfectly always match and require a command called localization. This command makes sure that how the robot perceives object placement in the environment and actual placement match up. Localized objects have a blue overlay and can me interacted with. If a user clicks on them in the interface, they become selected, and the overlay turns orange.

Messages
At the top of the middle of the interface the user further receives pop-up messages that contain information on four topics:

1. Information that the robot is currently executing, and the user must wait.

2. Information on execution and planning failure, see Figure 9.

3. Feedback from GO like standby messages.

4. Data readout information. For example, when investigating modules with the DIP.

Robot Stats
On the right side of the GUI one can find the command panels and information on the selected robot. In the top right corner one can click on the overview to select another robot. This interface can be used to control multiple robots, but the scope of this research focuses on the control of Justin. Additionally, there is a little icon that if clicked shows a 3D model of the robot representing the current state of Justin. Therefore, if Justin's arm would point forward, this would also be the case in the 3D representation.

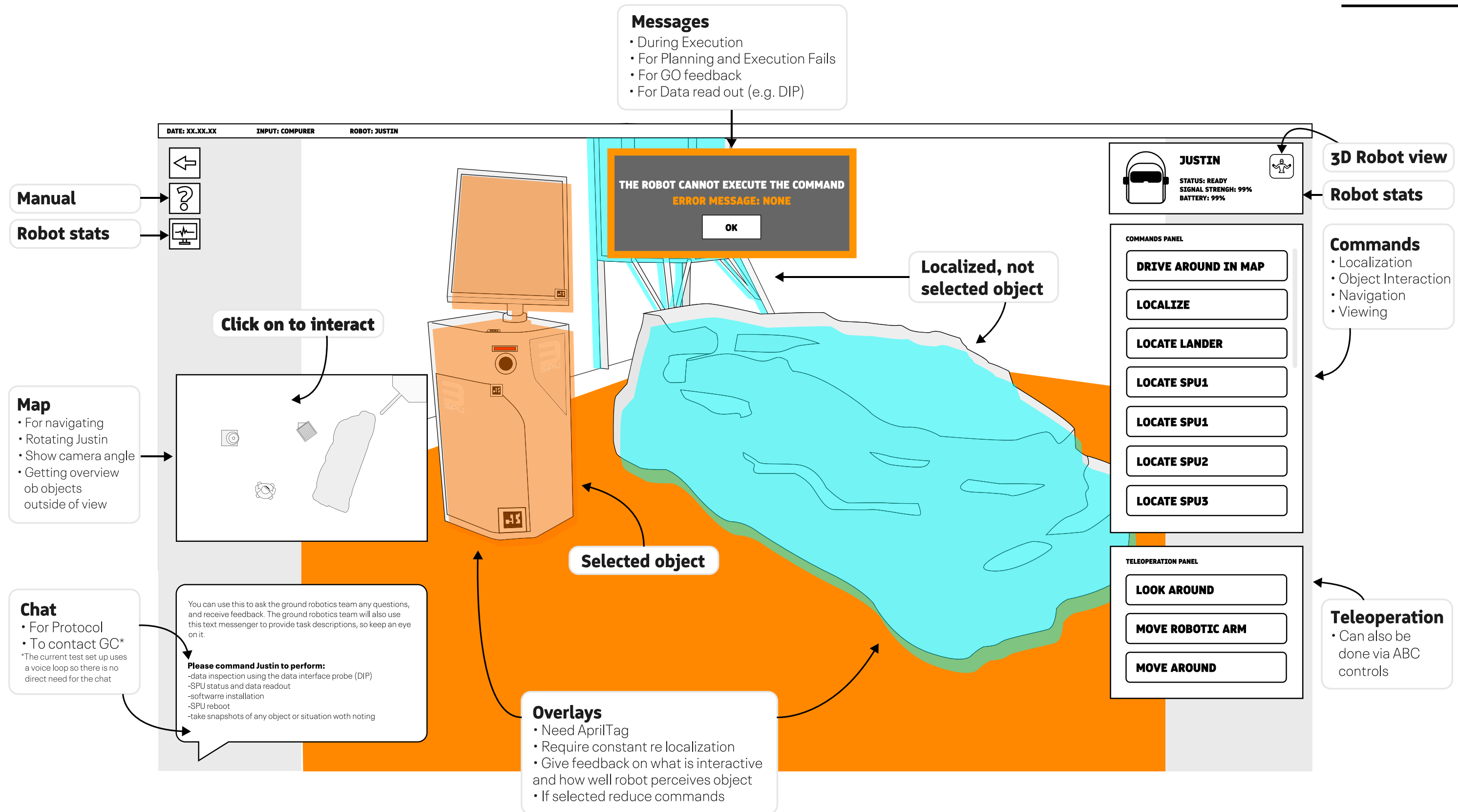


Figure 9: Interface of GUI explaining each feature, going clockwise: Message, 3D Robot View, Robot statistics, Commands, Teleoperation, Overlays, Chat, Map, Robot statistics (again), and the Manual.

Commands: Teleoperation and Autonomous
One can moreover find the command panel on the right side. Here the user can select from all the actions that the robot can execute autonomously. They contain 4 categories of actions: Navigation, Look around, Localization and Object interaction specific commands. If an object in the camera view is selected, only actions that are related to that object will be displayed, otherwise all actions currently possible to the robot will be visible. Below the command panel one can further find the teleoperation panel. Here the user can to the same as with the ABC buttons previously explained and choose between the three modes of teleoperation. [101]

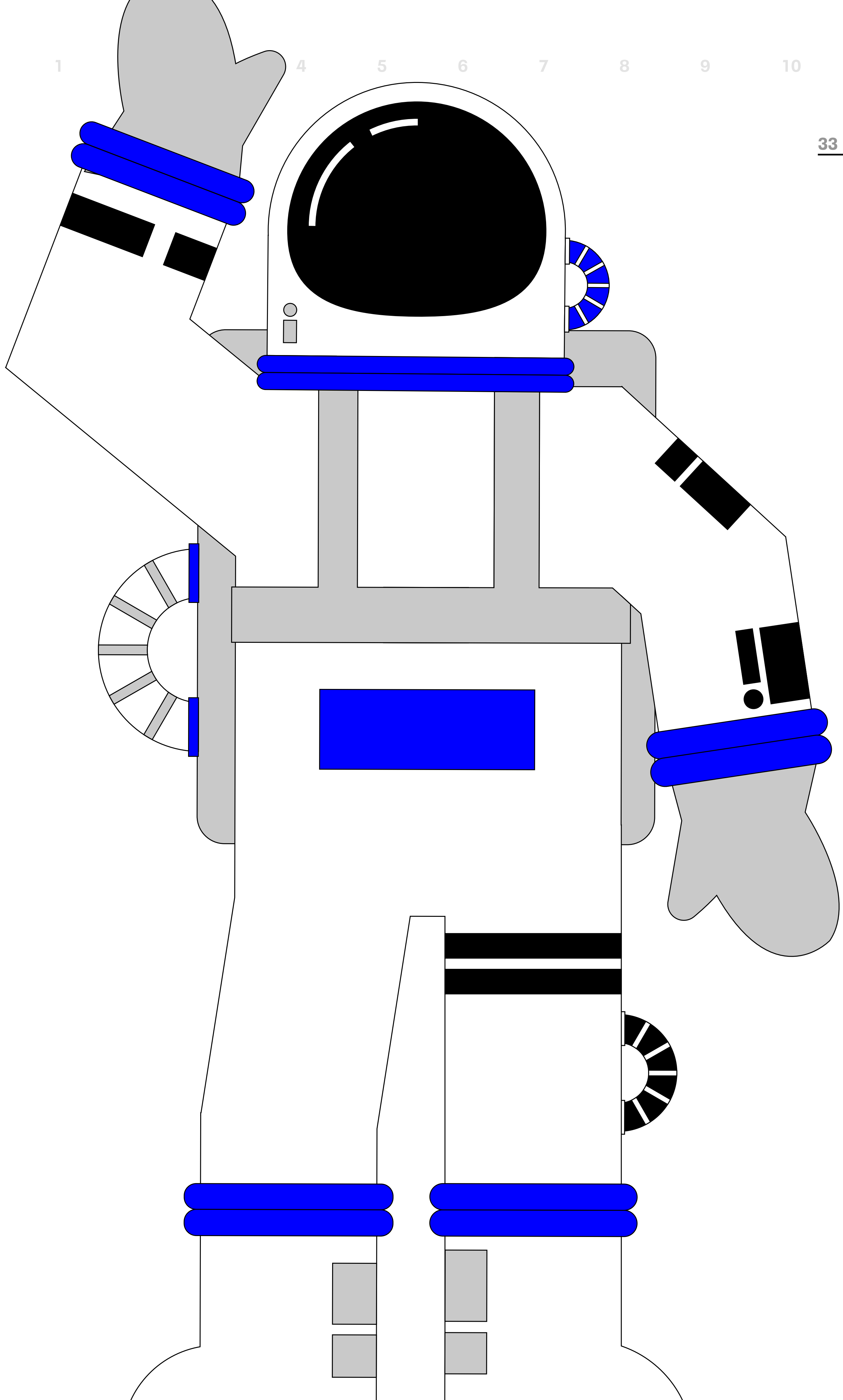
Other
The top bar of the interface contains information on the robot's teleoperation state (what is currently selected) and whether the robot is online or not.

Interface Design Limitations
Lastly, there are additional regulations that must be considered. The GUI must adhere to the ISS Display and Graphics Commonality Standard (DGCS) which also includes terminology, with the goal to create an overarching design language that runs through all interfaces [90]. This further reduces the need for astronauts to learn a new design language and the switch between different interfaces is easier.
For example, the use of certain colors, as they are designated for ISS subsystems and system conditions, respectively, is restricted. The current interface only uses gray-scales and an addition of blues and oranges [90]. Colors like green, yellow, or red have predetermined purposes. The DGCS further addresses button shapes requiring navigation buttons to be rectangular and command buttons to be rounded, they may include symbols [26]. Relevant examples for the use of specific terminology are "Yes" instead of "Confirm" or "No" / "Inhibit" / "Off" instead of "Cancel". Regarding symbols, the DGCCs also includes a list of predetermined icons [26]. While the idea of keeping a coherent design language is good practice, it should be noted that the DGCS was created in 2001. Since then, there has been a shift in design language and styles that even astronauts will stumble upon, at least outside their professional realms.

The Target User: Astronauts

The surface avatar system is designed for astronauts in orbit to remotely control a robot on a planetary surface. Understanding the target user is crucial to tailor the system to their specific needs. [38] Astronauts are not typical users, but highly trained specialists with unique qualifications and experience. To gain insight into their needs, it is helpful to consider the recruitment requirements that the European Space Agency (ESA) has for selecting astronauts, as well as their expectations for fulfilling specific roles and responsibilities during space missions. In order to become an astronaut with the European Space Agency (ESA), candidates must meet a rigorous set of criteria, including physical fitness, academic qualifications, and relevant professional experience. They must undergo multiple tests and trainings, medical screenings, and interviews. Selecting an astronaut takes over a year. ESA requires astronaut applicants to be highly skilled professionals that can work under high pressure and stress. Additionally, adapting to changing situations, with mature and quick judgment is on the list of expectations [9]. They must fit certain physical standards, psychological standards (like self-control) and show excessive education in technical or scientific areas including an outstanding professional background. Moreover, experience with aircrafts or being a test pilot or flight engineer in the past helps [9]. Astronauts spend long periods away from their home or earth needing to be able to leave their loved ones behind for an extended time.

Aside from these qualifications, an additional recruitment consideration concerns public relations. They need to enjoy the spotlight and communicate the importance of their job and space exploration. Active presence in social media is mandatory [9]. Moreover, fluency in English is required and speaking Russian desired. Astronauts often work in multicultural teams and openness to at least American, Russian, and Japanese culture is additionally preferred [41].



Problem Analysis

This chapter is an exploration of related work in teleoperation for space missions, highlighting the existing gap in research on error communication. It emphasizes the importance of preventing errors through the incorporation of situational awareness and teleoperation design elements.

Moreover, it contains the setup for a user study conducted to gain insights into the teleoperation system, with a specific focus on error handling.

Related Work

Exploring Space – A Difficult Task

The Global Exploration Roadmap suggests that robots would be utilized on the Moon and Mars for various purposes such as preparing landing sites, exploring the environment, collecting, and analyzing samples, and setting up infrastructure for future missions involving crew [42]. To make the most of these surface assets, researchers have recommended that astronauts operate these robots from an orbiting spacecraft in the future. This idea has been proposed in several studies [23, 45, 99, 115] as well.

Considering the limitations of space exploration such as lag and latency, low communication bandwidths, degraded operator performance due to spacecraft conditions such as microgravity, and extraterrestrial surface conditions increasing teleoperation difficulty [74, 92], have led to a series of research aimed at facilitating robotic control in space.

Advancements in Teleoperation for Space Exploration

In the context of space telerobotics, previous experiments include the Kontur-2 project, where a force-feedback joystick was used to teleoperate a robot from the ISS [44, 67]. Another significant project is METERON, a collaborative effort led by ESA, DLR, NASA, and ROSCOSMOS to explore teleoperation technology in space. Within METERON [115], experiments such as HAPTICS [3, 4], SUPVIS-E, SUPVIS-M [63, 66], and ANALOG-1 [64] have been conducted to study force-feedback perception, intuitive GUI interfaces, and task-level commands for robot autonomy. Additionally, the Avatar-EXPLORE experiment by the CSA [29] and the Surface telerobotics experiments by NASA [61, 62] have focused on autonomous navigation and interactive commanding of rovers in space.
These recent studies in telerobotics consider factors of mental workload and learnability for the users. For instance, integrating force feedback to improve the telepresence for the user [3, 4, 44] also under consideration of the users' environmental stressors of microgravity [64]. Another attempt at improving user experience shows in the SUPVIS experiments [63, 66, 79, 80] that investigated the use of an intuitive GUI, with experiments even considering learnability and situational awareness [88].

Bridging the Knowledge Gap: The Almost Untouched Domain of Error Handling

However, despite extensive research on teleoperation control, there is limited investigation specifically focused on error handling in the context of teleoperation in challenging environments like space exploration. To the best knowledge of this author, this is not just the case for space exploration but other sectors too. For instance, a study by Honig et al. [38] reviewed 52 studies on resolving failures in human-robot interaction. They additionally drew insights from related fields such as human-computer interaction, human factors engineering, cognitive engineering, and experimental psychology, and found only limited research on human errors, communicating failures, and the cognitive, psychological, and social factors that influence the development of strategies to mitigate them [38]. More specifically, the study did not find any literature dealing with human errors, the main part of the research is focused on technical failures with some considerations of social norm violations [38].

Know your System: Anticipating and Preventing Errors

Nonetheless, research form other areas may offer insights applicable to the teleoperation of robots on Mars. While it is crucial to have direct feedback and effective resolution strategies for handling errors, there are other factors that influence the user's ability to perceive and utilize this feedback [96]. For instance, situational awareness of the user, as their understanding of the system and environment greatly impacts their ability to identify and resolve issues. Additionally, the design of the robot itself plays a role, as a robot that can anticipate next actions or guide the user can help prevent errors. In essence, the user's knowledge base and the system's ability to guide and facilitate usage determine how effectively the feedback will assist in error situations [38, 96]. Therefore, before delving into specific communication strategies and error handling concepts, it is essential to address these factors as well.

The Importance of Situational Awareness

Situational awareness (SA) plays a crucial role in human-robot interaction (HRI) when it comes to supervisory control of remote applications. SA empowers users to make informed decisions, while its absence can result in speculative actions and consequently errors [17, 33, 38]. In the context of Computer-Supported Cooperative Work (CSCW) and Human-Robot Interaction (HRI), situational awareness lacks a standardized definition [46]. However, there are three definitions that cover important aspects for the error handling skills of a robot operator: (1) Concept/task awareness (the participants' comprehension of the execution process of their tasks) [15], (2) Situation awareness (The recognition and interpretation of environmental elements in a specific time and space, as well as the anticipation of their future state) [65] and (3) Informal awareness (The overall awareness of actors presence and activities in the surrounding environment.) [16]. Generally, the importance of situational awareness for teleoperation has been proposed in many teleoperation domains such as Industrial Welding [17], Vehicle teleoperation [59], Agriculture [5] and Urban Search And Rescue (USAR) [33].

Research on the teleoperation of vehicles showed that SA can be positively affected by reducing the need for multitasking [95], better quality of transmission [74], more available viewpoints [73], information cues such as earcons (auditory icons) or icons [46], limiting information in the Interface [91], adjusting cameras to light conditions [97], active communication from the device to the user [53] and added position information and previewed movement of other actors [118]. Moreover, when piloting remotely controlled drones, previous experience with video games showed to improve spatial orientation and SA [21]. In general, concrete solutions propose refining the operator viewpoint, such as transitioning from an egocentric to an exocentric perspective. Moreover, utilizing vibro-tactile feedback, automated viewpoint adjustment, and combining maps [74]. Augmented reality (AR) techniques have shown promise in improving situational awareness [60, 74]. For instance, projected turn direction arrows, different overlayed colors on objects to indicate reachability or generally superimposing of multi-modal interfaces (e.g., combining data from different sensors or sources, such as cameras, lidar, or other sensors, and overlaying the resulting information to create a more comprehensive understanding or representation of the environment.) [60, 74]. Visual representations of force and sound have some benefits in this regard but limited impact on teleoperation performance [74]. Multimodal/multisensory interfaces offer context-specific actions and have shown to improved situational awareness in complex situations [32].

Improving situational awareness enhances users' abilities and reduces the occurrence of errors. In the event of an error, users with a good understanding of the system and heightened situational awareness are better prepared to address the issue [17, 33, 38]. While improving the user's awareness is one way to go, other approaches try to let the robot adapt to the user. For instance, the use of social signals such as the users' subconscious facial expressions and gestures. The signals are used to trigger an error detection response from the robot, allowing users to implicitly assist in issue resolution [69]. Other methods explored for vehicle teleoperation consider eye-tracking to monitor operator SA and adjust the system based on that [37]. In the context of mental model building for human-robot collaboration first order mental models (adjusting the robots' behaviors based on user actions) or second order mental models (recursive reasoning, where a robot forms a belief about a human's model of the robot) are common methodologies [105]. However, in the context of space teleoperation, there is no direct contact to the robot. Time delays, restrictions in seeing the robot and limited bandwidth make the use of social signals or user action difficult to capture.

The Error Happened: What Now?
If errors cannot be eliminated entirely, other methods have been explored. For example, rebuilding trust, appropriate error communication, and applying gestalt principles to make error messages easier to understand.

If a robot throws an error, it can negatively impact the users trust in the system. Looking at Human robot interaction (HRI), several studies looked at different approaches to reduce the negative impact of errors in human robot pairing. Specifically, four repair strategies have been explored to reduce the undesirable influence of errors [20]: (1) Denial (based on the "Theory of Misinforming", it aims to deflect blame or responsibility and portraits itself as entirely blameless. This involves shifting the fault away from the trustee to another entity.) [10], (2) Apology (based on the "Theory of Forgiving", which serve as manifestations of regret or remorse. They may convey that the trustee sincerely values the trustors well-being and is motivated by genuine concern rather than self-interest.) [117], (3) Explanation (base on the "Theory of Informing" to provide a clear and straightforward explanation of the underlying reasons for a breach of trust.) [27], and (4) Promise (based on the "Theory of Forgetting", it seeks to establish expectations for the future, by shifting focus towards future positive outcomes and encourage a sense of forgetfulness towards past trust violations.) [103]. According to Fratczak et al. robots in an industrial context that communicate and apologize, are perceived as more predictable, less frightening, and easier to work with [94]. However, the studies by Esterwood et al. and Zhang et al. found that none of the repair strategies fully restored trustworthiness if trust violations happen repeatedly. Apologies, explanations, and promises showed similar effectiveness, while denials were consistently the least effective [20, 120]. Considering that Rollin Justin is semi-autonomous and capable of performing certain tasks independently, it is also worthwhile to explore the field of eXplainable AI (XAI). Here, providing contextual information, historical data, and proper references behind decision making enhances trust [6]. Moreover, Visual explanations, augmented reality-based explanations, and interactive user interfaces have shown to be beneficial for trust enhancement [6].

In terms of error communication, one study about eXplainable AI for robotic failure investigated the content of such messages. They found that explanations that incorporate both the history of recently accomplished actions and reasoning about the environment have shown the highest improvement in failure identification and solution identification scores [24].
Another aspect that has been examined is the use of multiple modalities, such as audio and visuals, to facilitate better comprehension and attention. For example, audio cues can grab users' attention and convey simple information, while visuals and contextual information provide a more comprehensive understanding of the error. Additionally, it is suggested to limit the provided information to not overwhelm the user [38]. In terms of Gestalt, combining visuals with written or spoken text has been found to enhance attention to information compared to text alone [40]. Visual warnings that are presented with organized information groupings and ample white space are more effective in capturing and maintaining attention compared to a single block of text [119].

While existing research in teleoperation control and human-robot interaction provides valuable insights, gaps and limitations remain in understanding error handling in challenging environments and the factors influencing users' perception and utilization of error feedback. Building on what is known from Human Computer interaction (HCI), eXplainable AI (XAI) and general Human Robot Interaction (HRI) this work aims to address these gaps. Specifically, by exploring error handling strategies in the context of teleoperation in space, with a focus on enhancing situational awareness and developing effective error communication techniques.

A User Experience Approach to Robotic Teleoperation

There has been numerous research on robotic systems for space exploration, particularly about the technical aspects of robot operation. However, little attention has been paid to the user experience and the human factors that influence the interaction between humans and robots. Especially the effect of dealing with errors. This gap in research inspired to take a more user-centered approach to examining the current user experience of Rollin Justin. Specifically, focusing on design methods to gain insights into the user experience of interacting with the system, with a distinct focus on error perception, awareness and implications.

Research Objectives

The primary objectives of the initial research with the robot were to evaluate the current user experience, find the main pain-points of error handling, comprehend the robot's perception by the user, analyse the level of situational awareness, understand the target group and map the environment. The reasons for each of these objectives can be seen in Figure 10.

The Questions for these objectives and the objectives themselves were extracted by using the WWWWWH Method [35]. Using this approach, most known factors and unknown factors were revealed and showed where more understanding is required.

Figure 10:
Objectives and their respective reasons.

MAP CURRENT EXPERIENCE
Map current experience with the system and the emotions of the user

Reason: To be able to design a better experience, compare it to the existing one in evaluation and get a better understanding of where the issues are.

FIND PAIN-POINT AREAS
Find areas that elicit the negative experience towards the planning errors. Find origin of negative emotions. Find reasons for Planning error (human fault or system etc.)

Reason:
- To find opportunities for improvement
- Understand where one could intervene in the system to fix the issue

UNDERSTAND ROBOT'S RELATION/PERCEPTION
Learn about the robot human relationship

Reason: A system that the user trust will be used more often, with More confidence and have a positive impact on the overall experience

UNDERSTAND SITUATIONAL AWARENESS
How well is the current system in terms of situational awareness?
How much do participants understand of the robot and the system?

Reason:
- It is an important aspect of teleoperation
- Without it, errors are harder to tackle

UNDERSTAND TARGET GROUP
Understand who the users are, their goals, motivations, problems

Reason:
- To make it more human tailored and fit the design choices to their needs (Understand what the users wish for in the interface, understand how they work and how a typical day looks like)
- Find opportunities where one could intervene in the system to fix the issue
- To design with preconditions in mind, potentially be able to make the transition smoother, keep knowledge alive and reduce need to learn system over

MAP THE ENVIRONMENT
Understand conditions (mentally, physically, context, workplace)

Reason:
- To design an intervention that works under the given extreme conditions and be aware of limitations.
- To ensure feasibility of the final design
- Get a better understanding of the target user (certain behaviours are context dependent and need to be understood to design a fitting intervention)

TRANSFER KNOWLEDGE
Understand what is out there already?
Find opportunities/context with similar problems
Find Inspiration (Exhibition, online, other contexts (like gaming), speculative design)

Reason:
- To increase the design space
- To improve the ideation phase
- To improve understanding of problem from another perspective

1 2 3 4 5 6 7 8 9 10

Research Process

A research plan was developed based on the listed objectives to be used for dedicated sessions of context and user research with individuals who had previously used the Surface Avatar system. Over the course of 8 days, sessions were conducted at DLR to gather insights. A total of nine participants were involved in the research, participating in individual sessions lasting each one hour. The sessions focused on exploring the general usage of the Surface Avatar system, with a specific emphasis on error handling scenarios.
The setup contained interview questions, questionnaires (UEQ which offer a holistic evaluation of user experience [102] and Godspeed for measuring the perception of service robots [12]), Usability tests [28, 111] using Storyboards [113] based on scenarios in combination with PrEmos (an instrument used for non-verbal self-report that can measure seven positive and seven negative emotions) [25] and an investigation into error awareness, see Figure 11. The interview questions were spread throughout the session. So were the questionnaires and scenarios, to ensure that the tasks would not get too repetitive and keep the participants attention. In addition to the sessions, footage of two astronauts using Surface Avatar from the ISS were analyzed with a task analysis method [112].

The sessions started with interview questions on the person themselves so that the researcher could later map their demographics based on the user's prior knowledge with the system, see Figure 14, p. 49. [14]. The questions then focused on the users experience with the system and in case they obtained or observed any of the trainings how they perceived that. These questions further functioned as a small sensitizer for the usability test that followed [100]. Sensitizing is helpful in getting participants to reflect on their prior experience with the robot and bring back memories [100].

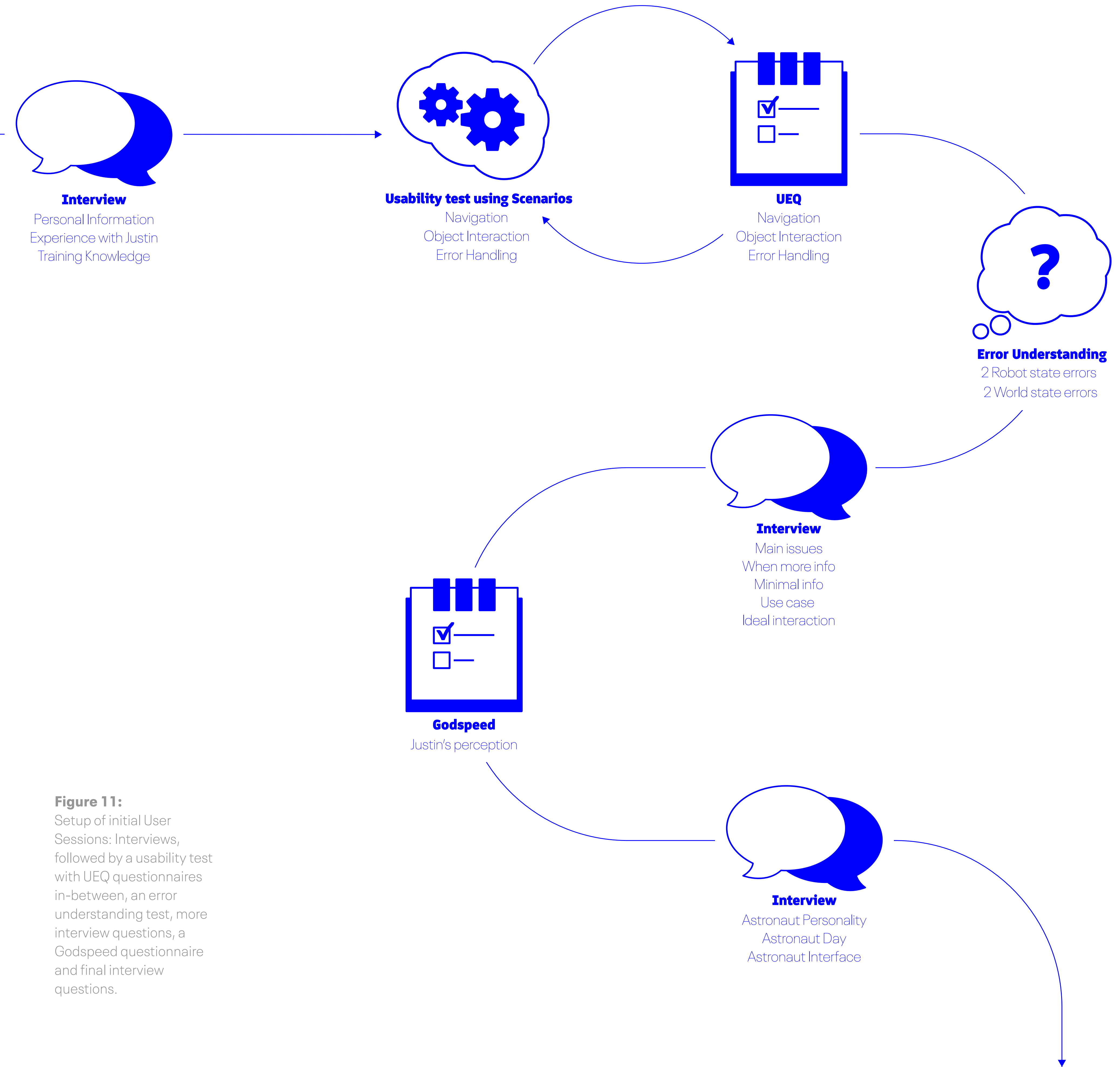


Figure 11: Setup of initial User Sessions: Interviews, followed by a usability test with UEQ questionnaires in-between, an error understanding test, more interview questions, a Godspeed questionnaire and final interview questions.

Research Methods and Setup

Usage Scenarios
Since the robot setup was not always available to be used during the sessions a workaround needed to be made. Visual aids in the form of a storyboard were preprepared to guide the user through a typical scenario with the robot [18, 113]. Considering that the robot is in constant development and can therefore also fail, be unresponsive or simply be unavailable, it was to some degree more realistic to use the scenario images as for the real experiments with astronauts these issues usually get fixed before usage. The experiment scenarios were divided into three categories: navigation, object interaction, and error handling. In the scenario, participants were asked to find a mug, navigate the robot to it, pick it up, and deal with the robot's failure. This task is similar to those performed by astronauts on the ISS during experiment testing [51, 56].
In one such experiment, astronauts had to find and fix a broken SPU, requiring them to navigate, use GUI commands, and handle robot errors. This experiment involves similar subtasks as the one used in the study, enabling comparison with the available video footage. However, it was decided not to use the exact same scenario as all participants had already seen or completed it. To obtain more general results, the structure and basis of the experiment was kept but the task were changed to one not previously encountered by the participants. This decision aimed to simulate the experience of astronauts, who encounter new tasks during their missions.
The storyboards illustrated the interface and controls available for subtasks in the scenario and showed the changing GUI and camera view throughout the process. Consequently, the user did not have to remember all actions available to them and could pick how they would usually decide to interact with the system step by step [18, 28, 111, 113].

Addressing User Experience with PrEmos and UEQ
While going through the scenarios the participants were further presented with PrEmos [25] and were asked to indicate how they feel for every step they took. PrEmo is an instrument used for non-verbal self-report that can measure seven positive and seven negative emotions, refer to Method Glossary p. 4 for a detailed explanation. The collected data can be valuable for assessing the emotional effect of the current Surface Avatar system on the user [25].
Since the scenario could be split into three categories users received visual aids only up to that part. After each part the users received a UEQ [102] that they could fill out based on their experience, focusing specifically only on that part of the interaction they just went through. The UEQ scales offer a holistic evaluation of user experience, capturing various dimensions including traditional usability factors such as efficiency, perspicuity, and dependability, as well as experiential elements originality and stimulation [102], refer to Method Glossary p. 4 for a detailed explanation. Then they continued with the next part and repeated the process until error handling.

Setup for Investigating Planning Errors
After carrying out an analysis with PrEmo and UEQ, the main point of the research started: Understanding the user's robot awareness and fixing capabilities when encountering errors. However, a usability test with the robot itself was not possible, as forcing the robot into planning errors was not achievable. To simulate the context, User testing with prototypes was conducted instead (using simple interface mockups and interaction scenarios) [28, 111]. In addition to the user testing, participants were asked to follow a Talk aloud protocol (to get a better understanding of user behavior and thought process) [106] and interactive fiction (to enhance the simulated situation, by asking the user to believe that the test setup is the real robot setup) [34].

In this part for the study participants were presented with an image of the robot's interface in a specific context, see Figure 12 for an example of the first interface visual.

Participants were told that they had attempt to command the robot to perform a specific action highlighted with a mouse-over icon on the interface (see Figure 12) but encountered an error message when trying to do so. Subsequently, the participants were then asked to vocalize their assumptions as to what went wrong and how they would attempt to resolve the issue. The researcher would explain the effect of each action proposed by the participant. If the action resolved the problem, the participant was presented with the next image, otherwise the user could tell the researcher their next approach at solving the problem until they solved it. Overall, four images were shown to the user. Every time the participant solved a problem the story proceeded, and they encountered a new error of a different type that was connected to the image shown, see Figure 13, p. 47. The methodology used in this study is reminiscent of a text-based adventure game or interactive fiction, where users are required to imagine the effects of their actions through text-based feedback [34]. However, in this study, participants were provided with feedback on the effects of their actions via speech not text. They were presented with static images of the interface that contained clues regarding the reason for the failure, they did not have to imagine everything, just the explained consequences of their actions.

The method makes use of Talk aloud protocol [106] in the sense that it asks the participants to vocalize their assumptions and next steps. Additionally, aspects of usability testing are utilized as the user is put into a realistic, even if error-prone scenario. Participants can interact freely with the system, however limited over speech, and the errors are based on existing problems [111]. The use of this methodology allowed for an investigation into users' assumptions about the robots reasoning and usage of the interface and controls to solve errors.

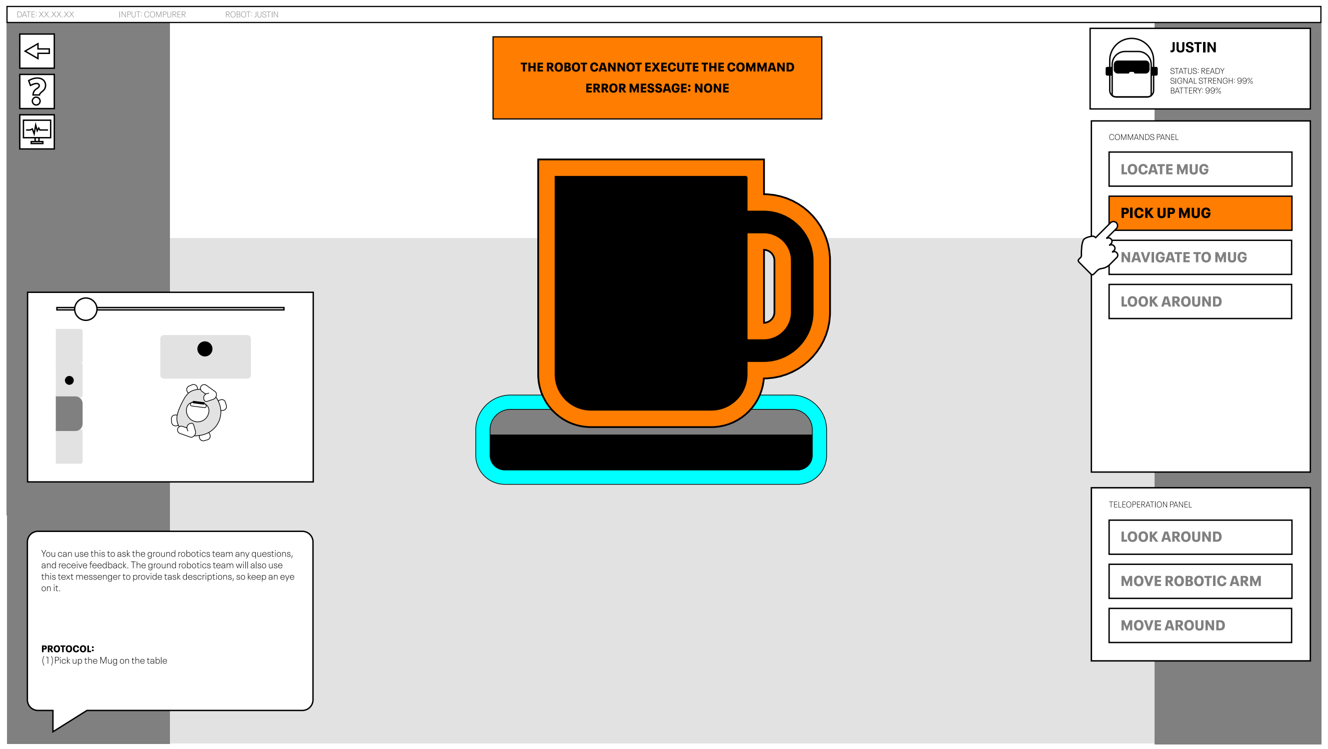


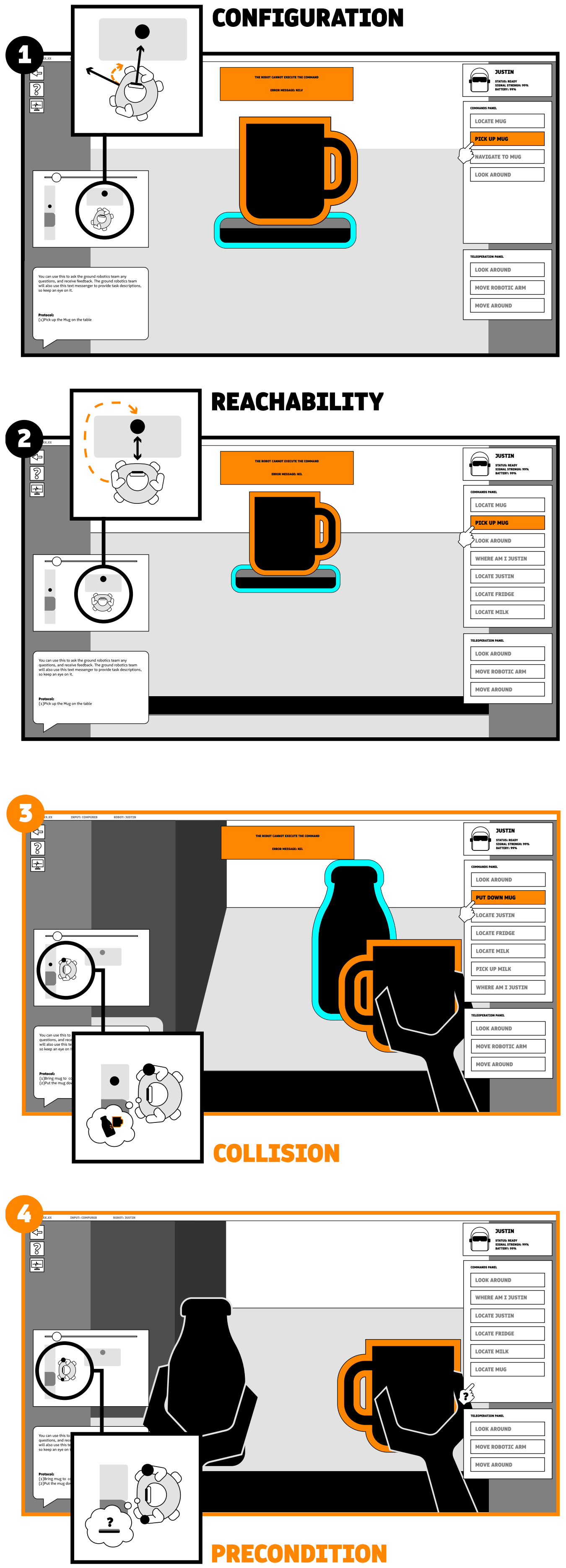
Figure 12: The interface for testing: The selected objects are highlighted in orange, and the located objects in cyan.

Investigated Error Types
Overall, this test included four errors, two of which were related to the robot's state and two to the world state, see Figure 13.

The scenarios depicted in the visuals were not typical of a Mars mission but rather unfamiliar to the participants to ensure that they had not encountered the exact error before. While the errors were unfamiliar to the user, they were not unlikely to happen in the presented scenario and the validity of the scenarios was ensured with a pilot test asking the head of the Surface Avatar team for feedback. Overall, the four error scenarios on the right were tested with the participants, see Figure 13.

More interview questions followed this, focusing on what the user thinks the main pain point with errors is, when a better understanding of the robot is needed to handle errors (training, during interaction, when encountering errors etc.) and what the minimal information is for them to be able to deal with errors. Moreover, the interview dove into the use case and then asked how the ideal interaction looked like for them.

Figure 13: Overview of all four tested error types: (1) configuration, (2) reachability, (3) Collision and (4) Preconditions.



(1)Orientation error: An error occurs during the attempt to pick up a mug in the first scenario due to the robot's misalignment. The user's camera view is directly facing the mug, while the robot's body is oriented in a different direction.

(2)Reachability error: The second error arises from a reachability issue. Despite the robot being properly oriented (as participants were asked to fix this in (1)), the mug is positioned too far on the other side of the table, making it difficult for the robot to access it.

(3)Collision error: Participants were then asked to brin the cup to the counter. The milk on the counter is in the way and the robot does not want to place it. While there seems to be a lot of space on the counter, the robot is very careful with collision trying to protect itself. Moreover, within its code, it is only trying 5 possible positions before giving up. It could therefore be that trying to execute this action with the exact same setup multiple times would sometimes throw a planning error and sometimes not.

(4)Precondition: The final error is related to a precondition. In this case, the operator is unable to place the mug down because the system restricts mugs to be placed only on coasters. This situation is analogous to a Mars context where a user cannot rotate a SPU without first unlocking it, which may not be visually apparent if the user is unaware of the locking mechanism. In the first scenario, the user observed the mug being placed on a coaster, but there are no explicit indications that a precondition is causing the issue. Preconditions are typically inherent to the robot's programming and may not have visible cues.

Godspeed: Measuring Robot Appearance
Additionally the perception of the robot was investigated using a Godspeed questionnaire [12]. The Godspeed questionnaire is a useful tool for measuring the perception of service robots on the aspects of anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety [12], refer to Method Glossary p. 4 for a detailed explanation. However, this aspect was not a main part of the research and mainly there to ensure that the robot's perception is not a cause of negative emotions in the user experience.

Diving deeper: Interviews
Lastly, the interview focused on the astronaut. Since it was not clear when setting up the research plan whether astronauts would be available, the plan was flexible and had a different set up and rephrasing of questions depending on the participant. For non-astronauts the questions focused on what they observed when interacting with astronauts generally, when they receive the training and when they are interacting with the system (Surface Avatar). For astronauts these questions were directed at them. Since astronauts are the main users of this system, the questions were focused on learning their perception of surface avatar, their motivations, goals and wishes. They further tried to understand the typical day of an ISS astronaut.

Astronaut Video Footage: Task Analysis
Aside from the sessions the astronaut video footage was analyzed, and interesting observations noted with timestamps. For one of the experiments, where astronauts must find and fix a defect SPU, a task analysis [112] was conducted, allowing to find patters, potential usage preferences and problems.

Participants

Tests were carried out with employees of the DLR who had prior experience with the robot and system due to being part of the development team or astronaut training or had volunteered for a Dry run before. Selecting the participants happened based on their knowledge with Justin and overall, nine participants took part. Figure 14 shows the participant demographics. Three of the users were not part of the development team of Justin and had only little usage experience with the robot. This is very similar to the situation that astronauts are in, considering that they get a training six months prior to their experiments sessions and have no further interaction with the system in-between. The rest of the participants were in one way or another part of the research team working on Justin and Surface Avatar. While an ideal setup would have consisted of research sessions with astronauts that had already interacted with the system, this was not possible. Not only were no astronauts available at any given time, as most of the astronauts that had interacted with the system were stationed either at the ISS or ESA in Cologne, it later turned out that astronauts are extremely busy and inaccessible. Setting up experiments with astronauts simply requires a multitude of organizational steps, outside the scope of this work [51].

The research conducted in this study received ethical approval from the Human Research Ethics Committee (HREC) at TU Delft. All research activities followed the approved data management plan, and participants were asked to provide informed consent by signing a consent form. It is important to note that participants volunteered for the study without receiving any form of compensation.

Figure 14:
Participant demographics of first user study

Name	Part of Team	Usage Experience	Training Experience	Astronaut Contacts
01	No	Once, 5 Months ago	Check out once	None
02	No	Tester once, couple days ago	Check out once	None
03	No	Months ago once	1 x Checkout	None
04	Yes	Many times Tester	Few Checkouts	A few: Spoken to and observed
05	Yes	Once and single parts multiple times	Observed once and helped once	A few: Spoken to and observed
06	Yes	Several Years: Many times tester and test observer	Observed multiple	Many Years: Spoken to, observed, trained
07	Yes	Multiple times	Observed multiple	A few: Spoken to and observed
08	Yes	5 Years with Surface Avatar	Holds Checkouts and observed many Trainings	Many Years: Spoken to, observed, trained
09	Yes	Multiple times	Observed once	A few: Spoken to and observed

Analysis Results and Problem Review

Through a combination of interviews, questionnaires, scenario-based tasks, and an error investigation, the study explored various aspects of participants' interactions with the robotic system. The results of each method are discussed in this section.

Navigation, Object Interaction and Error Handling

The results of the scenarios are presented in Figure 15, they depict the answers from all 9 participants from the UEQs split up into the scenarios. Additionally, the PrEmo answers of all participants, are positioned below each UEQ result. As expected, having to handle errors fares the worst. Navigation performs well on average everywhere, except for novelty and object interaction is generally positive.

Looking at the PrEmo results for navigation, the user experience is split into three different positive emotions joy, pride, and fascination. While there are a few negative emotions such as anger or boredom, they are not notably high. Moreover, looking at the explanation's users gave, negative reactions are mainly due to the constant need to locate an object and not the navigation process itself.

"I will localize again. I find this a bit annoying [...] it's kind of unnecessary." – P06 (translated from German)

"So, the localization action that I have to do manually, I don't think that's a good thing" – P04

Otherwise, the scenarios showed that users have an interaction preference for moving the head with 6/9 approaching this via the joystick. There was no prominent difference between the participants that are part of Surface Avatar and those that are not (2/3 non team members and 4/6 team members). Moreover, 6/9 preferred to use the interface's command option to navigate. One participant (P06) that had a higher knowledge of the system also mention that it would ensure more accurate positioning of the robot.

"As a user, I know that if I just navigate with the joystick, I'm not that accurate. And if I do that with the "navigate to", I'm pretty sure that I'm in a position where he (the robot) can handle the object." – P06 (translated from German)

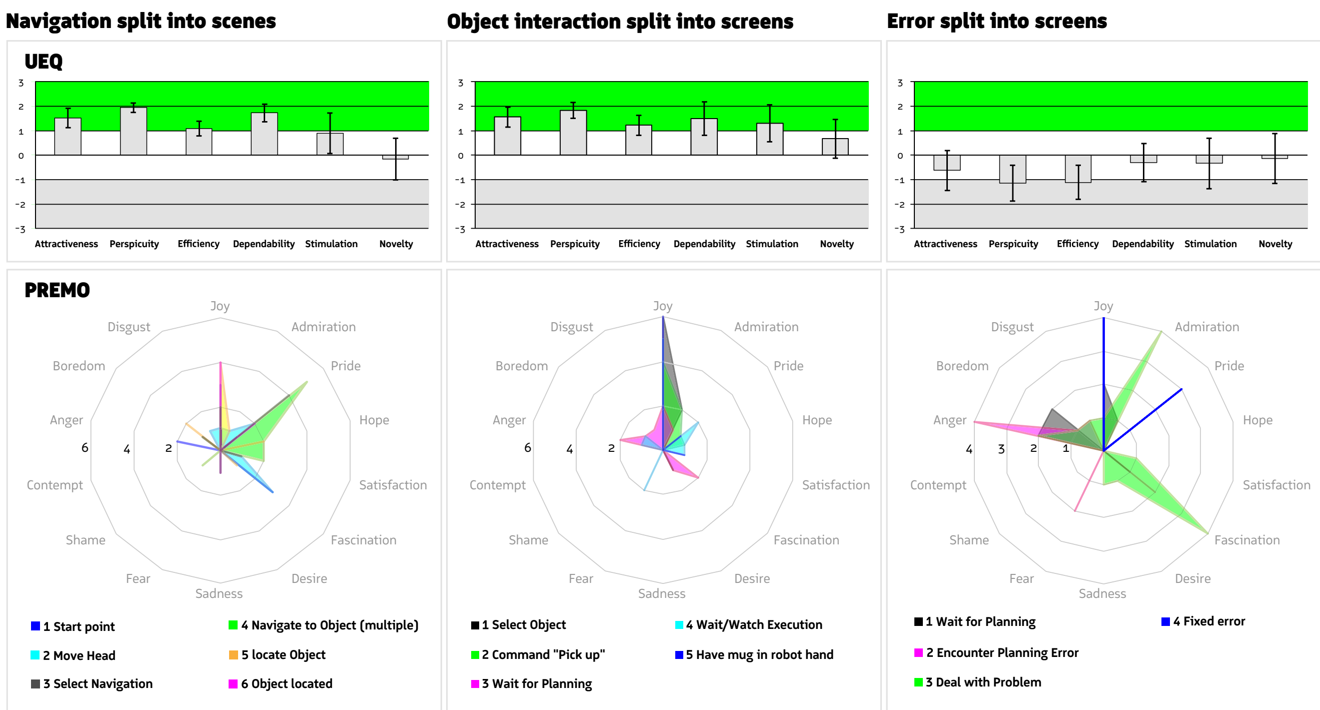


Figure 15: Overview of UEQ and PrEmo results for all three scenarios. From left to right: Navigation, object interaction and error handling.

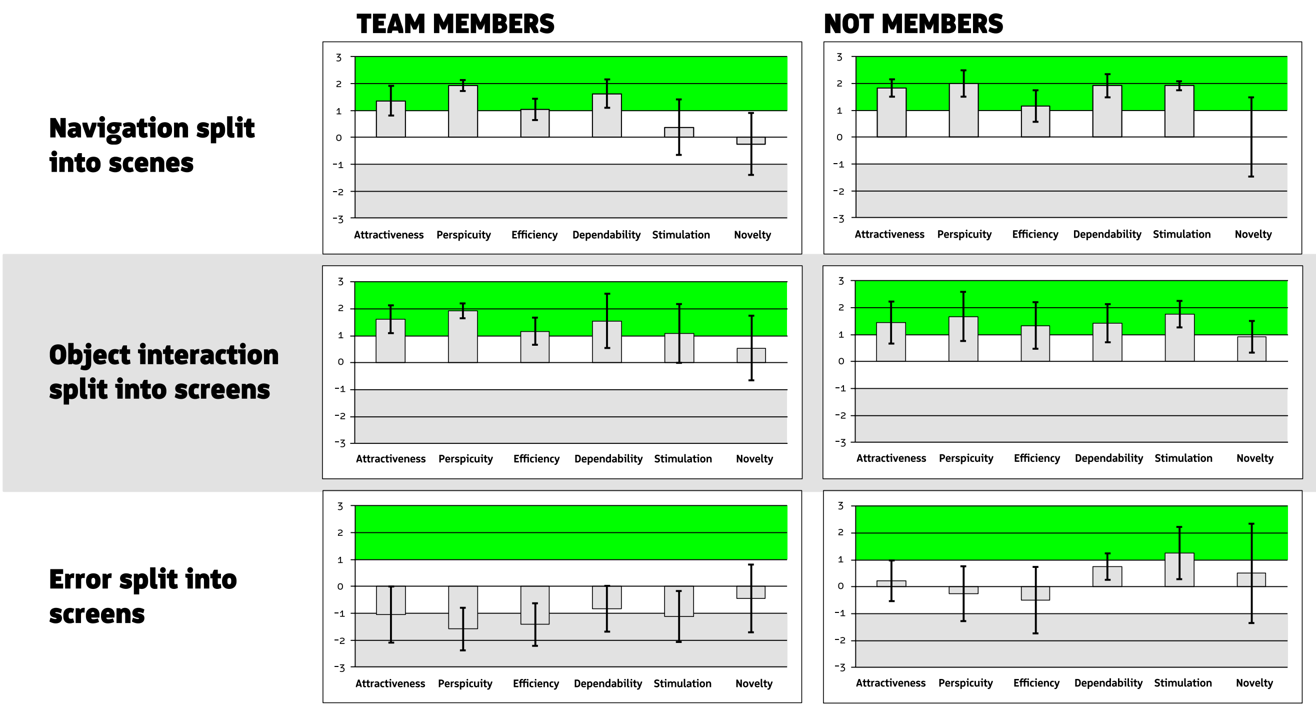


Figure 16: Comparison of UEQ results between the three non-team members and the six team members.

Besides, only 1 user decided to use the map for navigation and one to move the head, every other user ignored this feature for navigating.

Interacting with Objects shows the most positive results for the UEQ and the PrEmo. The PrEmo is mainly Joy throughout the interaction. However, there is a bit of anger and fear when waiting for the robot to plan and execute. It should be noted that 4/9 participants mentioned that they are fine with the waiting as they are aware of it and 4/9 mentioned a dislike and being annoyed for the waiting, therefore showing very split opinions. Waiting for the execution on the other hand depends on whether the execution itself is visible in the camera view or not.

"If I see it, I feel 1 (Joy), If I don't see it, then I am 9 (Fear). Unfortunately, I am scared. I don't know what exactly is happening" – P06 (translated from German)

"Assuming I see it, I am happy. [...] If I don't see it, I would be 6/7 (Fascination/Desire) [inaudible] probably even 12 (Anger). It depends on how much confidence I have in in the robot doing this task" – P08

Lastly handling errors is the scenario that shows the worst UEQ. Encountering an error is especially bad when it comes to efficiency and perspicuity. However, it should be noted that this part of the user's journey was rated a lot more negatively by the members of the surface avatar team compared to the other users, see Figure 16.

Looking at the PrEmo results further shows most of the negative emotions (specifically anger) happens when encountering the error, while dealing with the error shows a more complex mix of emotions split into anger, admiration, and fascination. This split of emotions has also been observed by other literature on HRI. While some studies suggest that people prefer predictable behavior in robots [75] others indicate that people find unpredictable behavior more engaging [31, 38, 52, 57]. Moreover, several research studies, as listed by Shanee Honig and Tal Oron-Gilad in their paper review, suggest that failures can create enjoyable interactions with robots [38] Looking at the reasons that participants give one could summarize that it is on the one hand annoying, especially if the error message is bad, but on the other hand intruding because it feels a bit like a challenge.

"It would just say fail to execute. And it will not give you a reason. So as a user, I would feel like, I don't know what is up with my system [...] it is not a good feeling." – P04

"I think there grows an interest in me. How do I solve this problem now?" – P02 (translated from German)

"First, I am at 12 (fear) [...] but then 6 (fascination). It is interesting, what do I do next, that is the question." – P09 (translated from German)

Error Scenarios

Moving on from the scenarios to the investigation of error handling. As mentioned earlier, the error scenarios focused on two types of causes: problems stemming from the robot state and problems from the world state. The robot's state is the position the robot has in the world which includes his orientation, kinematics, etc. The world state represents what the robot's internal representation of the world is. This for example includes object localization or that certain actions are not provided to the user before a precondition is fulfilled [56]. Note that the robot's world state might not always align with the user's world state, which become very visible when object's overlay's do not align well. The overlay is so to speak where the robot thinks the object is. .

Error Scenario: Orientation

The first one causes an error when trying to pick up a mug because the robots body orientation is bad: The head and therefore the camera view of the user is facing directly at the mug while the body is facing a different direction. This issue is not unlikely to happen as when navigating the base of the robot the head does not move along, vice versa. A direct hint that the orientation is the cause lays in the map.

Looking at the results, 6/9 users assumed the right category of error "Motion planning" although it ranged from reachability issues to collisions and arm choice, see Figure 17. Only one of the 6 people immediately understood that the robot's base orientation in relation to the head was bad. The solution to the problem was to orient the base to be in line with the head which required a movement of the body. Figure 17, moreover, shows the actions that participant took in trying to solve the problem. It should be noted that actions are counted once a user performs it at least once, the graph does not take into consideration how often the same user repeated the same action. In the end 7/9 users solved this, although 3/7 received guidance from GC after contacting them. On average the participants needed 3 attempts before resolving the error.

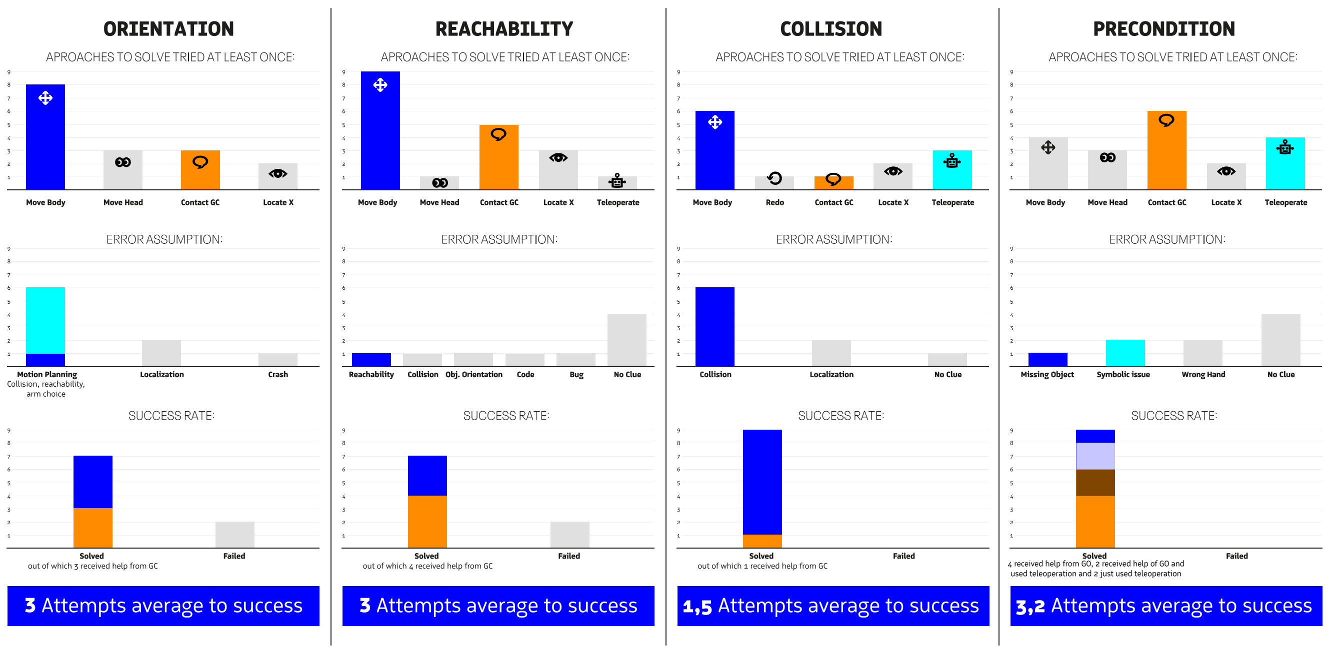


Figure 17: Results of all four error investigations, from left to right: Orientation, reachability, collision and precondition

Error Scenario: Reachability

The second error is caused due to a reachability problem. Even though the robot is oriented well now, the mug is too far on the other side of the table. Only 1/9 participants assumed reachability to be the root of the error while the rest of the participants had completely different assumptions and 4/9 stated to have to no clue what the cause was, see Figure 17. This is a clear difference to the first error. While there is a hint in the map that the mug is further on the other end of the table, it is hard to tell from the camera view how far the mug really is. This is also a problem in the real setup as depth perception from only the camera view of the robot is very bad. The actions that the users took are also more spread out than in the first scenario with more contacts to GC. Again, 7/9 participants were able to resolve the issue with 4/7 receiving help from GC. The average attempts were also 3, however in this case there was a substantial difference between users, some taking many steps others very few.

Error Scenario: Collision

The third error looked at collision. The milk on the counter is in the way and the robot does not want to place it. While there seems to be a lot of space on the counter, the robot is very careful with collision trying to protect itself. In this scenario, 6/9 participants assumed correctly that collision was the problem, see Figure 17. There are multiple solutions to this problem like moving a bit to the side where more space is, removing the obstacle first, teleoperating the mug placement or even redoing the action multiple times. In the end, all participants solved this problem with only 1/9 contacting GC for help. Most of the people moved the body but 3/9 decided to place the mug via teleoperation. Overall, this only took 1,55 attempts on average which is the lowest number for all four situations.

Error Scenario: Precondition

The last error is based on a precondition. In this scenario the operator is not presented with the option to put down the mug because mugs can only be placed on coasters. Looking at Figure 17, the approaches to solving this issue are very spread out and only one participant assumed the right causation after a long time of going through all possible problems known to them. This participant is one of the two users with the highest experience with Justin. Another 2 users went in the right direction assuming that it was a problem in the symbolic state of the robot and 4/9 participants had no idea what the problem was. Either teleoperation or retrieving the coaster from the table first were solutions to this issue. In the end, every participant resolved the problem, but 6/9 received help from GC and 2/6 that received help decided to teleoperate still. Another 2 users decided to teleoperate on their own. Therefore 4/9 resolved the issue with teleoperation and 5/9 using the coaster. Overall, it took participants 3,2 steps to find a solution, the highest number of attempts of the four error scenarios.

Godspeed

Assessing the perception of the robot itself is not the focus of this work. Nevertheless, the Godspeed questionnaire was used to reduce the factors that could potentially impact the results of the study. If the questionnaire resulted very badly one could assume that other negative emotion during the user experience might not directly connected to the interaction and system but the robot itself. Nevertheless, the questionnaire showed no striking results, see Figure 18. While anthropomorphism and animacy are worse than the rest, they are not too low. Moreover, both are likely a bit lower because the robot is very slow, and the user does not see a lot of the robot's bodily movement as they are limited to the camera view.

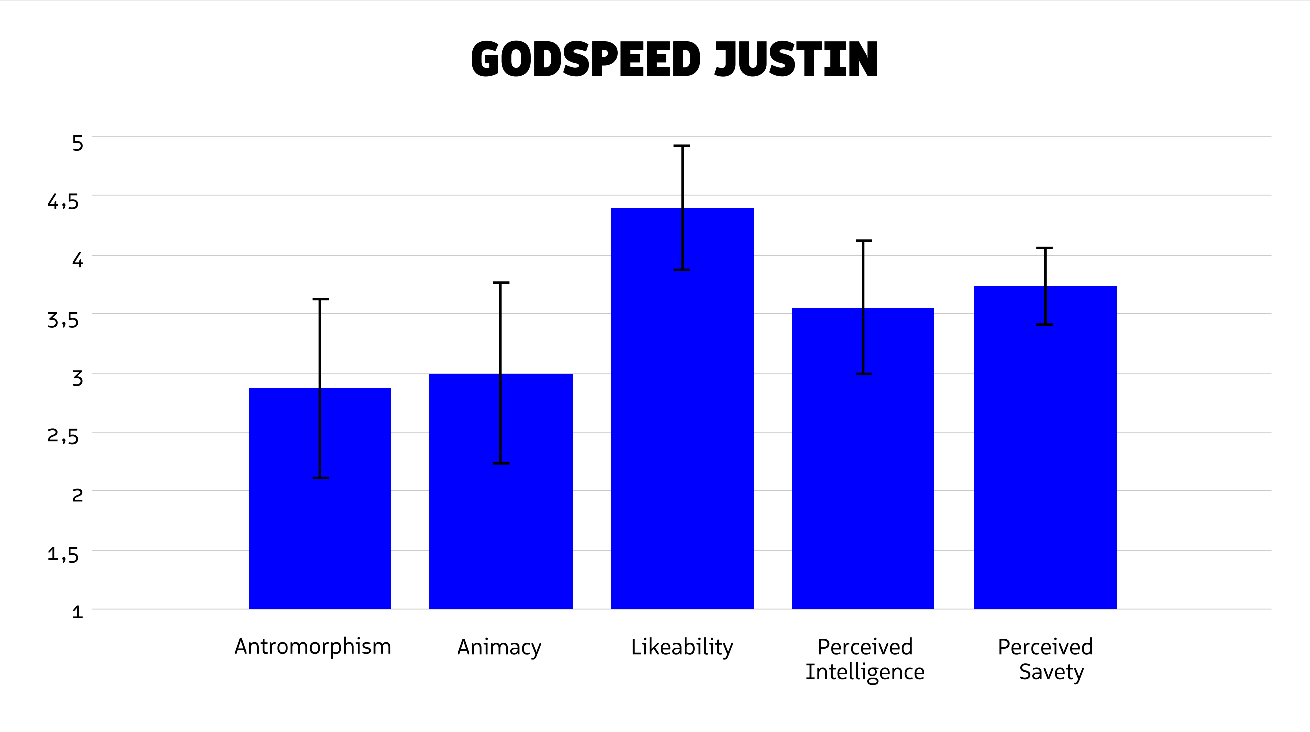


Figure 18: Godspeed questionnaire results for the five scales: Automorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety.

Interviews

This section showcases the interview results from questions that focused on errors, the use case of the system and the participants ideal interaction with the system.

Errors
One part of the interviews focused on the errors. Some participants mentioned more than one reason of what they think the main pain point of error handling is for them. Five out of nine (5/9) participants reported that the current error message does not provide sufficient information.

"If the description for an error is bad, then I have no way of finding out how to tackle the problem." – P06 (translated from German)

"I think, it's very frustrating that you don't get some sort of feedback, at least something, right" – P04

"The robot cannot do it due to a planning error does not seem enough for me" – P03

"At least one error message that is understandable" – P05 (translated from German)

Moreover, 7/9 think it is due to the lack of understanding of the robot's state.

"I think […] humans do not understand the robots view" – P02

"The user does not understand on what basis these actions are displayed. These actions are practically displayed on the basis of free conditions. This is not visible to the user"-- P06 (translated from German)

"I think its intimidating […] that you don't know what's wrong with the robot. It's comparable to not knowing a language. You cannot communicate" – P04 (translated from German)

On the other hand, participants had spread opinions on where more information is required so that they can handle errors. 3/9 Would like to have it during the training, another 3/9 when interacting with the system (like more information on the robot's state, e.g., a reachability map). Another 2/9 only need more information when the error message pops up and the remaining 2/9 would like to receive input when waiting for the planning.
When looking at what this information should be, specifically what the minimal information is they would like there is a more noticeable difference. However, some participants mention multiple things. 5/9 would like instructions what to do and a clear reason of the error. Another 5/9 mentioned that more information on the robot state/perception is required. Additionally, 2/9 think more feedback on what objects are interactable is needed and one person noted that a better camera view for depth perception would help.

Use Case and Ideal Interaction
Participants mentioned different use cases when asked about the application of the Surface Avatar system. On the one hand, the general safety of the astronaut was mentioned, as mars is a dangerous place, and it would otherwise not be possible. On the other hand, participant noted that astronauts think it makes their missions more efficient. Doing extravehicular activities (EVA) takes a lot of preparation time and the suits have limited dexterity in their hands. Moreover, it was mentioned that the system can be used for any kind of teleoperation task, not just mars.
When it comes to the ideal interaction all users either mentioned higher immersion, better higher-level autonomy, or both: Being able to delegate tasks easily with minimal effort, while jumping into an experience similar to what is displayed in the movie "Ready Player One" if teleoperation is needed.

Astronaut Observations and Footage Analysis

In the end, astronauts are the ones supposed to handle Surface Avatar. Therefore, considering their needs and understanding how they interact with the system is vital. Astronauts are very inaccessible, and it was not possible to interview them directly. However, participants (6/9) that observed, talked to, or trained them, shared their observations and impressions during the sessions. Moreover, video footage of the astronaut testing from 2022, with two astronauts, was shared with the researcher and analyzed on their usage patterns, see Figure 20. The two astronauts followed the typical protocol described under Experiment Protocol (see p. 25) controlling the robot on earth directly from the ISS.

Many of the qualities mentioned under The Target User: Astronauts (see p. 32) have also been observed by the interviewed participants. Especially, that they are very good at following commands (6/6 participant mentioned), coping with stressful situations and unexpected (5/6 participant mentioned) and that they are distinct in their field (4/6 participant mentioned).
Moreover, participants shared their observations of how they observed the astronauts use the interface.

It was mentioned by Participant 07 for example, that the mini map was an astronaut's suggestions for the GUI. Additionally, 3 of the 6 participants stated that the astronauts lack the understanding of the robot's state and limitations. Which is understandable considering that they are not the once who implemented the system nor do they have experience with surface avatar except for the training that lays six months in the past. Lastly 3/6 mentioned that the astronauts grasp the control use faster than the command use of the interface. There were additional observations, but since they were different among the participants or only mentioned once, they were deemed irrelevant.

Video Analysis
Figure 19 provides a guide on how to interpret the task analysis conducted based on the astronaut videos, offering an overview of the key elements to consider. Figure 20 presents the comprehensive task analysis, capturing the complete sequence of actions and durations and can be observed on the next page.

The two astronaut videos showed that both consulted GO multiple times (GO talking is highlighted in blue, see Figure 19). This is something that is not that easily possible in the real scenario but currently easy due to the voice loop and therefore there was no usage of the chat.
Both astronauts further had to get used to the concept of localization and understand its need. One of the astronauts (1) received an explanation on the overlays and localization and where then told to navigate to SPU3. The astronaut then commanded "locate SPU3" as they confused it and though it would navigate to the module, since they did not notice any chance and the robot did not act, they contacted GO for explanation. The other astronaut (2) specifically mentioned localization as one of the things that they need to get used to.

"What did it do, when I selected SPU3, it didn't do anything" – (astronaut 1)

"I think I need to get used to kind of which commands needs to proceed to which commands, in terms of locate, localizing myself within the map and localizing myself to different objects." – (astronaut 2)

These results also align with previous research of Rollin Justin in the METERON SUPVIS experiments [88], where "Localize" was one of the main commands where participants showed difficulties grasping the concept. Moreover, their spatial awareness lacked. One astronaut (2) was told to back away from the lander even though the astronaut felt that they were already far away enough. They also struggled with the teleoperation as the depth perception from just the camera view is limited. The other astronaut (1) was told to make use of the 3D representation of Justin to get a better awareness of how they are oriented, but they still showed confusion and kept asking questions. Moreover, when navigating to SPU3 they felt too close to the module and GO assured them it was the perfect position.

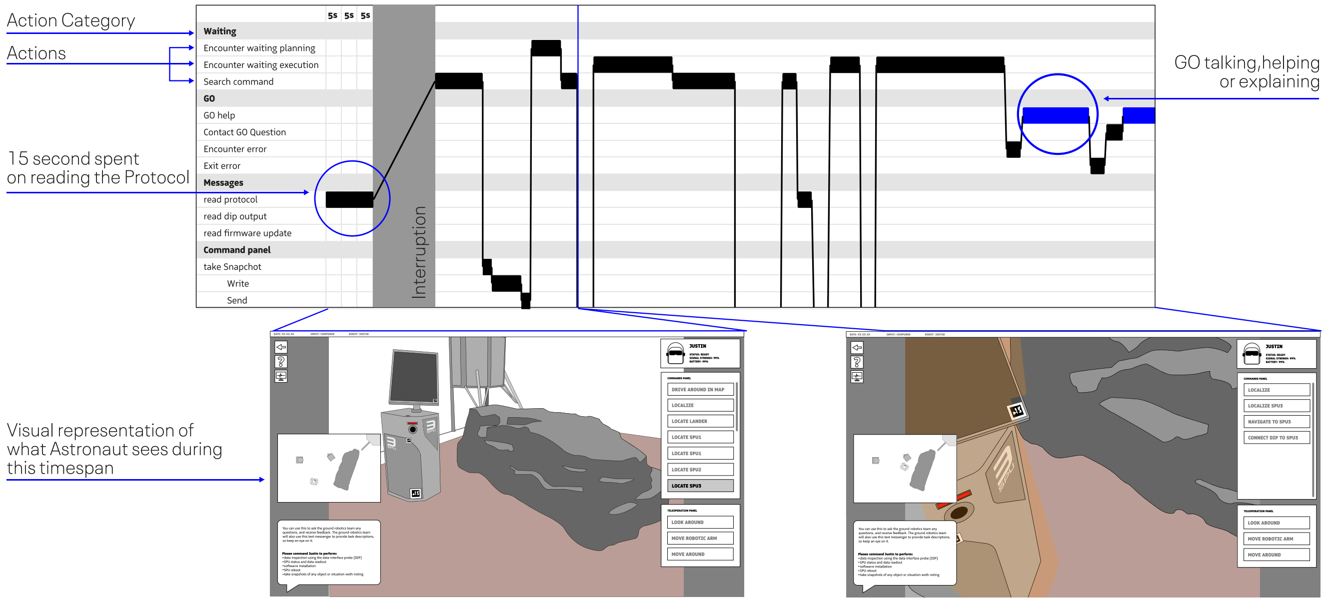


Figure 19: How to read a task analysis: A task analysis provides a visual representation of the sequence of actions and durations, offering insights into how users interact with a system or perform specific tasks.

"Looking at the map view, it does not look like we are that close to the lander, so I did not realize" – (astronaut 2)

"Am I looking straight now, or did I turn all the way to the right?" – (astronaut 1)

"Oh, that feels very close to SPU3" – (astronaut 1)

These results too align with the previous METERON SUPVIS research, where the factor that was found to slow down the users the most was the limited field of view. However, the situational awareness was rated with 8.62/10 on a SUS score by the 14 participants which is within the range of "Best imaginable" for the standardized SUS questions [89].

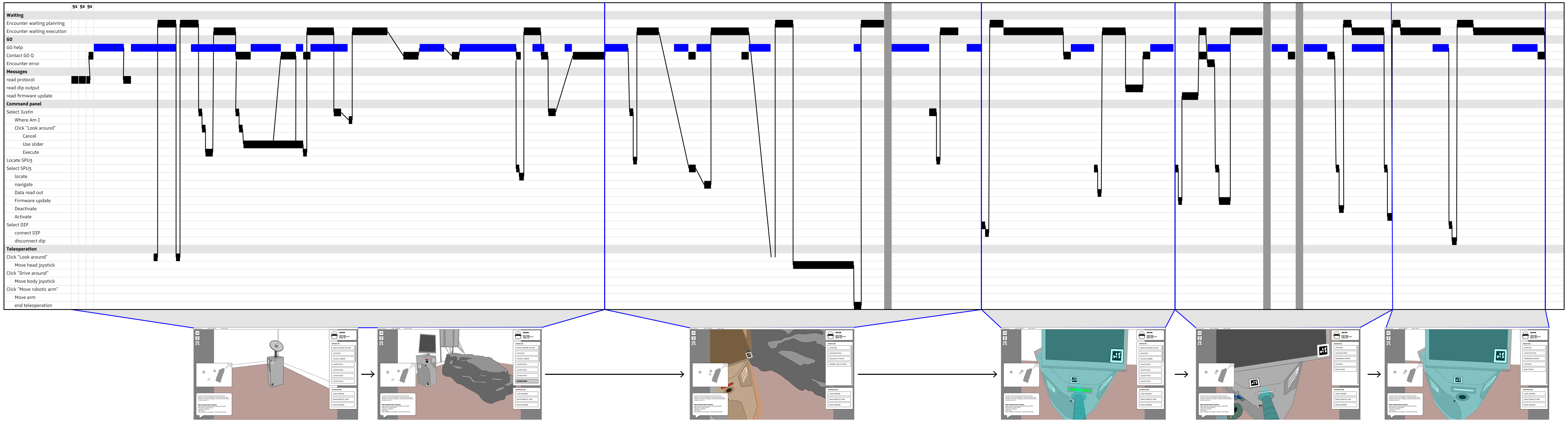
Aside from that, one of the astronauts (1) had initial questions on terminology like "DIP" and difficulties finding the right look around since there are two (one in teleoperation and one in the command panel). Additionally, they reached the end of motion for the robot's head. The other astronaut (2) had to handle an execution error and teleoperate the DIP insertion which took up the entire time of the experiment leading to an unfinished protocol. They had to end after 17 minutes excluding time of interruptions, of solely teleoperating the right arm to fix and execution error during DIP insertion. At one point while teleoperating the robot, the astronaut checked the interface, specifically the command panel, likely to find an alternative option to teleoperation. This is probably because the teleoperation becomes tiring over time. Studies have shown a general higher fatigue and less precision over time when using a system with higher telepresence vs a joystick teleoperation mode [116]. Moreover, it was noted by astronauts in a similar experiment to surface avatar with the same robot Justin, that telepresence and teleoperation becomes tiering after 20-30 minutes. [58]. This, and the fact that the second astronaut did not finish the protocol after they got stuck with teleoperating, further highlights the need to reduce teleoperation via higher level task commands over the interface.

The current system presents certain challenges related to situational awareness and the astronauts' understanding of the robot, resulting in a high reliance on seeking external assistance. These identified issues provide valuable insights for this work by highlighting areas for improvement and guiding the focus of the research towards addressing these concerns.

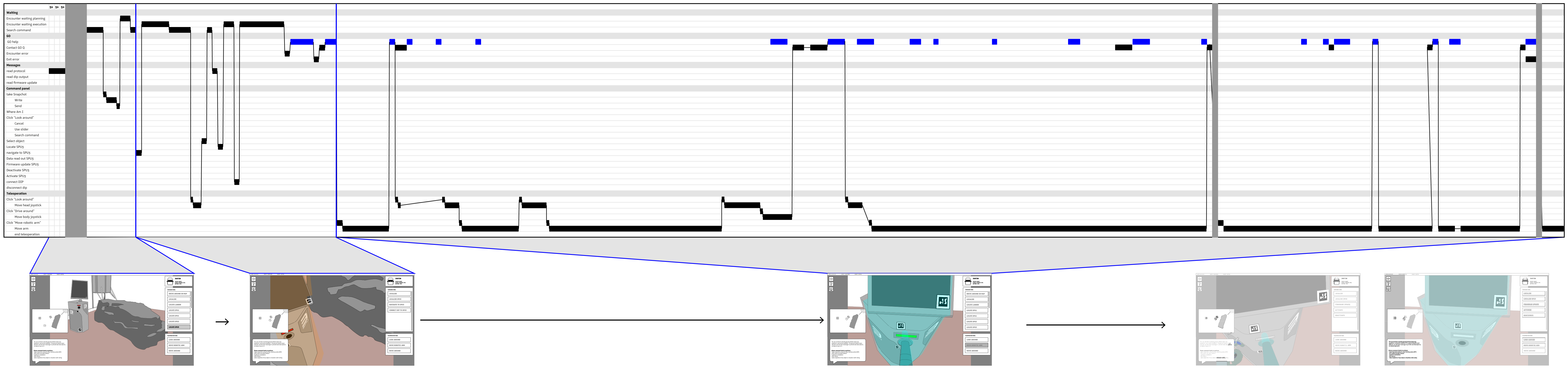
However, to help build upon the success of the Surface Avatar UI, this analysis focused specifically how to further improve and add new capabilities to the system. Previous research [89] has shown that the overall system is generally intuitive, easy to learn, and has been very well-received by astronauts. The Surface Avatar project provides a significant advancement for robotic teleoperation in space and the found difficulties are only a small excerpt of an otherwise well designed setup. By recognizing the strengths of the current system and building upon them, this work strives to contribute to further improvements that can effectively address the identified challenges and enhance the system's performance, particularly in error handling.

Figure 20: Task Analysis of Astronauts' System Usage and Visual View
This figure presents a task analysis of astronauts' system usage based on video footage. The left column displays action categories, including "Waiting" for planning or execution, contact with GO (Ground Operatotions), messages received by the interface, usage of autonomous commands, and manual control in teleoperation. Each column represents a time span of 5 seconds. The figure highlights the duration of each action, such as the sections of teleoperation for the right arm or the waiting time for execution. The blue sections indicate timespans where GO talked and tried to help. Dark grey beams represent interruptions in the experiment and time gaps that have been omitted. Below each task analysis, visual views are provided to illustrate what the astronauts were roughly seeing during each step of the mission. Overall, the figure provides insights into the amount of time the astronauts spent on each action while interacting with the system and how frequently they switched between different actions.

Astronaut 1



Astronaut 2



Result Overview

In the context of improving the error handling for astronauts, the research highlights several factors. The scenarios showed that the overall journey for the user works well with the current system with minor problems like the constant need for localization or waiting times. As expected, error handling performed worst, however, it is not a purely negative experience. While encountering an error is a negative experience, dealing with said error showed a mix of curiosity and anger.

The error investigation further highlighted how errors differ from each other. Overall, the awareness of the participants "why a certain error occurred", except for collision, was low. Even though there were hints in the interface, especially the map, participants took for 3/4 errors 3 or more attempts to solve the problem. This indicates that the users had difficulty understanding the robot. Moreover, 7/9 participants further mentioned during the interview that their lack of understanding the robot is a main pain point when handling errors. Looking at the astronaut videos and the astronauts observed behavior by the participants further indicate that there is a general lack of understanding for the robot's world state and configuration. Additionally, methods such as the map, which would make it easier to understand the current configuration of the robot, are not considered directly. However, this could also be since any additional information is located at the edge of the interface and users tend to focus more on the center or top left corner of a screen [93].

However, multiple factors influence the gaze of a user, and the attention could also be due to the general layout, attention grasping elements or the user's current intentions/goals [93]. The experiment further showed that when a planning error happens and the less obvious the error is, participants tend to get help form GO. In addition to that, the astronauts' videos showed an increased need for GO help when encountering problems. Nonetheless, either with or without the help of GO, almost all participants managed to recover and solve the problems.

Overall, one can conclude that the current system does not provide enough direction for participants to recover easily without additional help and that some form of further information is required. The overall insights were combined into a user Journey that is used in the next step to review the problems origin and find opportunities for improvement.

Scoping and Problem Review

The combined insights can be seen in Figure 21, next page. The user journey was created based on a typical experiment of finding and fixing a defect SPU. It includes all the necessary steps to solve the mission, as well as a typical error that can occur. Moreover, it is divided into nine subtasks, with each subtask further broken down into smaller tasks.

The nine subtasks are analyzed across five categories: control options with indicators for the best solution, emotional journey with notes on why emotions may rise or fall, available solutions for emotional dips, conflicts that can occur despite the available solutions, and opportunities to resolve those conflicts.

The choice of control options is based on input from Dr.-Ing. Daniel Leidner, leader of the group for Fault-Tolerant Autonomy Architectures at DLR [55], who has extensive experience developing Justin for teleoperation in space. The emotional data was gathered using the PrEmo method and interview questions in participant sessions. Astronaut video footage was also observed to identify where emotional spikes and drops occurred. The information in the "current solutions" category reflects the interface's status as of February 2023, which is important to note as the interface is constantly changing. The "conflicts" category is based on data from the participant sessions, interviews, and astronaut video observations, as well as NASA's requirements on cognitive workload [82]. Lastly, opportunities for resolving conflicts were identified based on all the aforementioned factors.
It's worth noting that while the participant sessions did not include this specific task, the tasks they did perform required similar or identical subtasks, so they were included in the journey.

Based on the user journey and the listed opportunities, three main concerns have been identified. Firstly, the lack of general robot awareness among users, secondly, the underutilization of existing tools designed to help users understand and troubleshoot robot errors and lastly astronauts start of as novices with the system, and they struggle to navigate the GUI in the beginning. These concerns also connect to the constant contact to GO.
From these three challenges the following directions are proposed:

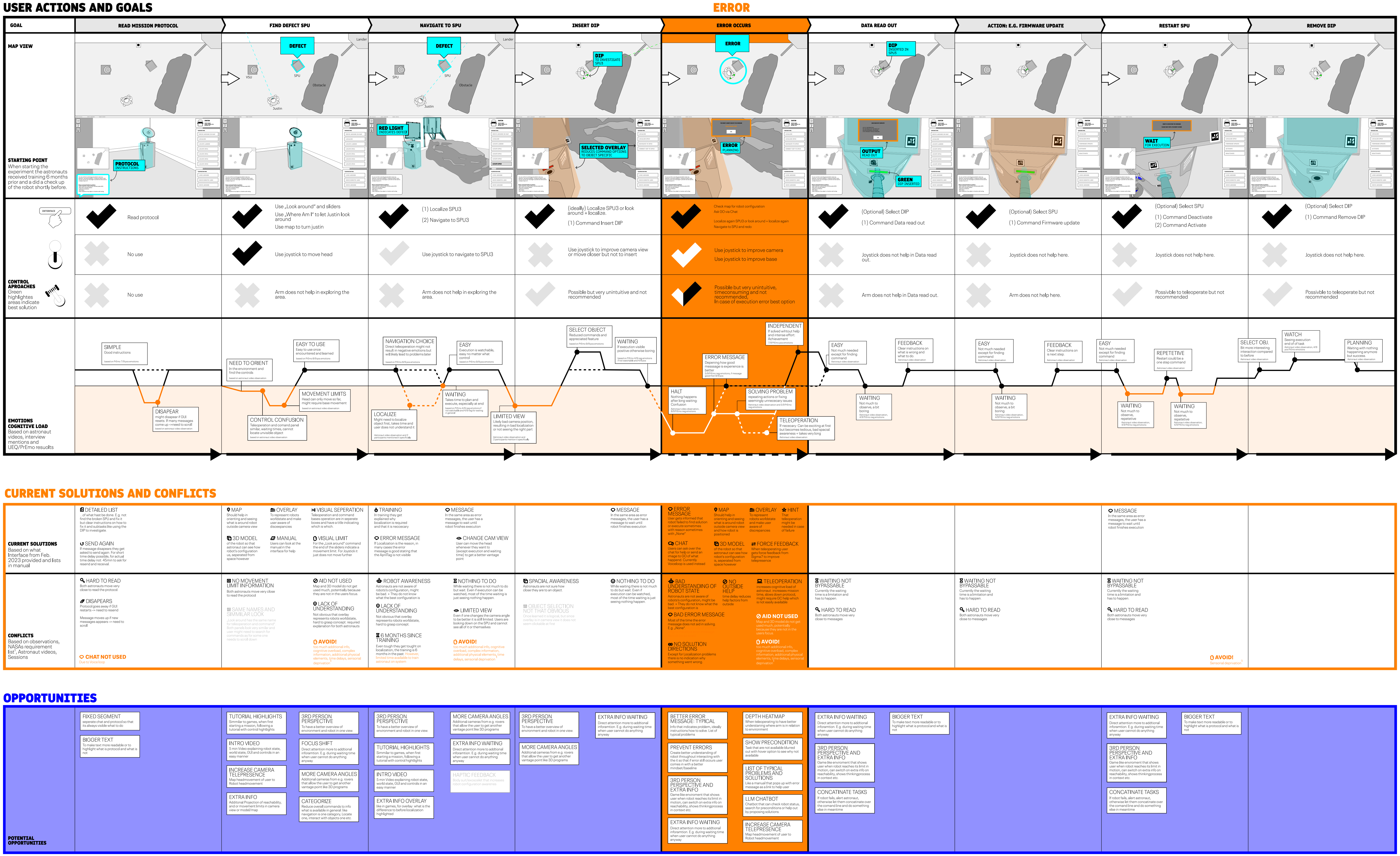
1.Increase robot awareness: creating more awareness for the robot throughout the usage of the system, so that errors occur less frequently and if an error still occurs, the user has a better baseline of knowledge to handle it.

2. Improve utilization of existing tools in the moment of encountering errors: providing information or pointing out existing tools that could be helpful in the case of an error and making the user aware of them.

3. Assist novice users: developing tools that can help novice users get started, as astronauts are typically thrown into the system with minimal information and may struggle with the initial steps of using the system.

Concrete solutions have been developed based on the three directions, which will be explained in the following section.

Figure 21: User Journey showing the process of a typical experiment with an error encounter (left to right: Receiving Protocol, finding the SPU3, navigating to it, inserting the DIP, error and error recovery, reading out the data and removing the DIP), The Journey is split up into seven sections, from top to bottom: Goals, visual representation of process, control approaches, emotions, current solution for problems, conflicts with current solutions and potential opportunities for improvement



Concept Directions

This chapter presents an overview of three concept directions developed to tackle the identified main issues, providing insights into their rationale and showcasing visual examples of their integration within the system. Feasibility and desirability considerations are discussed to refine and select ideas for prototyping.

Concept Overview

To address the challenges of increasing robot awareness, improving tool utilization during error encounters, and assisting novice users three concept directions were identified:

First, the "Third Person Perspective" that allows the user to gain an overview of their surroundings by seeing the robot in its environment, with the aim of improving situational awareness.
Second, the "Debugging Dashboard" concepts that aim to provide more detailed information and guide the user in the moment of a planning error.
Lastly, general "Usability Add-Ons" are proposed to refine the interface and provide more interaction cues to the user.

In this section, we describe the three main themes along with their motivation, and develop the concepts with each theme, refer to Figure 22, for an overview of visualizations for each concept. Finally, we carry out a selection procedure to identify concepts to take forward for further development.

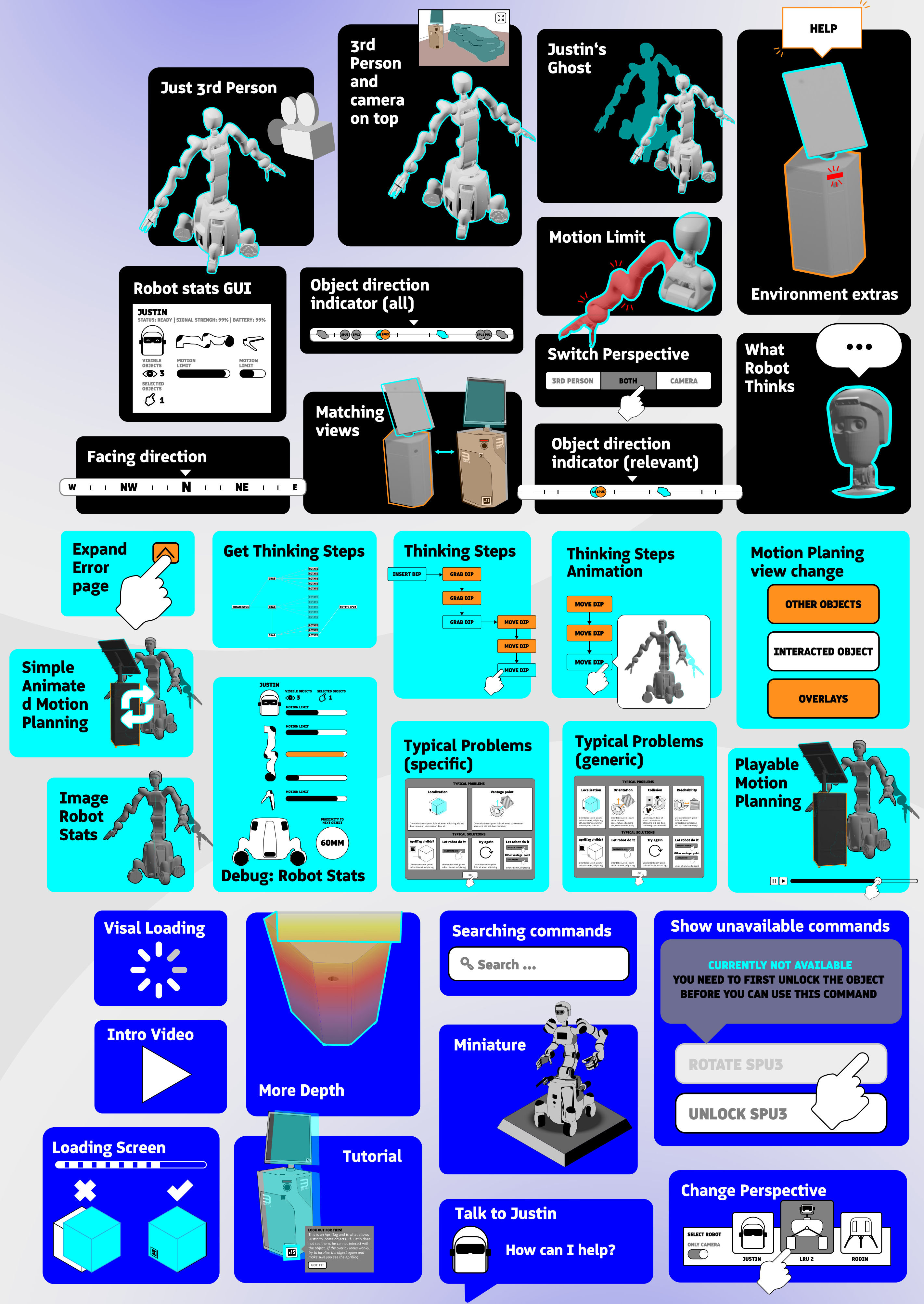


Figure 22:
Overview of Concepts with their respective visualizations and feature directions, top to bottom: 3rd Person Perspective, Debugging Dashboard and Usability Add-ons.

3rd Person Perspective

This concept direction aims to improve robot and situational awareness by providing a view of the robot and its surroundings from a perspective centered on the robot. This approach can offer valuable information like a ghost preview of the robot's thinking process, augmented information about the environment, and an easier way to access helpful features from the map. By adopting a robot-centric perspective this concept aims for users to gain a better understanding of the robot's actions, potentially preventing errors and helping them handle issues that do arise more effectively.

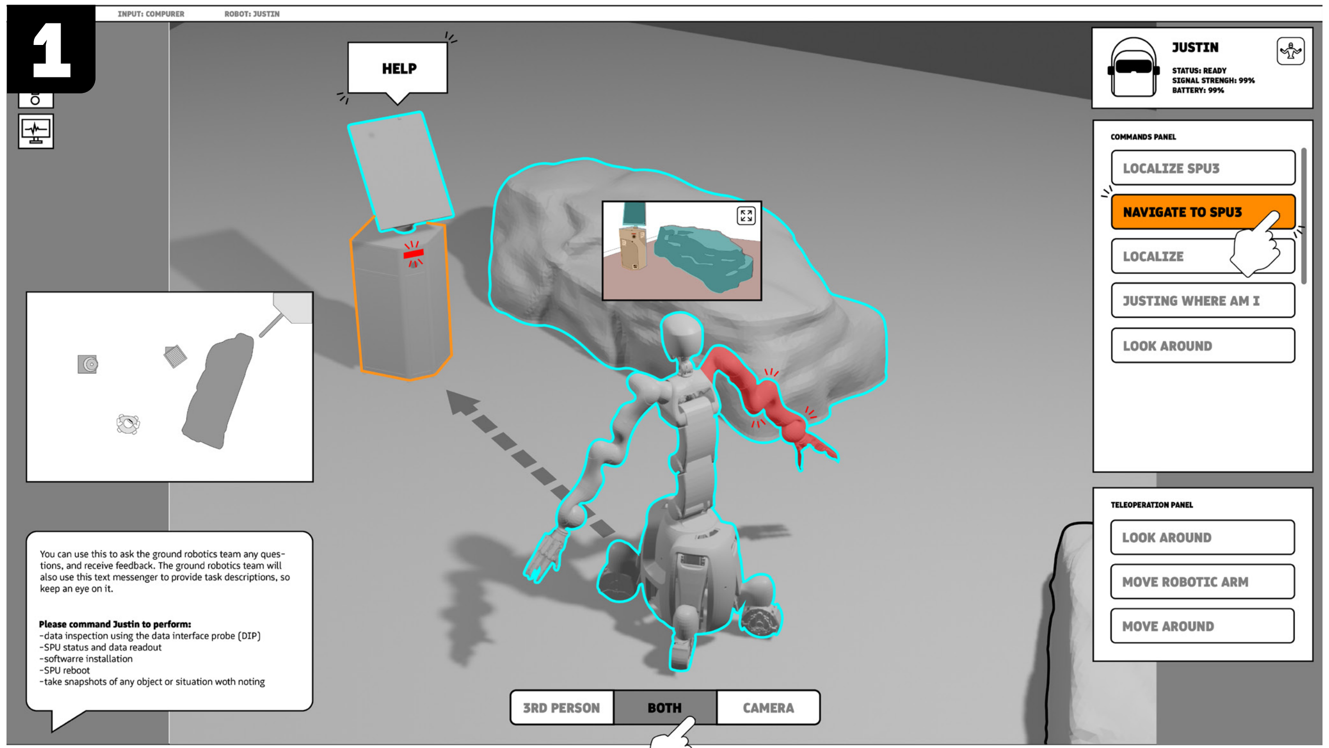
Why a 3rd Person Perspective?
Multiple ways of implementing third person views have been explored already, including cameras mounted on a robotic arm [78] or simply attaching it on a long stick [87], using a pair of ground robots with one working as the capturing camera [87], Micro Aerial Vehicle (MAV)-based systems that use flying camera which follow an Unmanned Ground Vehicle (UGV) on their own creating a third-person view for the user [1], creating a third-person-view interface that looks like a real photo by accurately matching 2D images to 3D shapes, Virtual Environment Vehicle Interface (VEVI) that allows for a view that can be easily customized as the operator has the ability to freely adjust the viewpoint and direction through the use of an HMD and a head tracker [87] and general navigation in a 3D environment with 3D free view [108].
However, most of the research in this area focuses on implementation strategies and underlying algorithms. Little attention is given to the user's perception and experience. Moreover, research that explicitly compares the impact of different viewing modes for the user, is limited.

Nevertheless, research has shown that using a third-person perspective for robot teleoperation and control is safer and more effective than a first-person view. First-person views make it difficult for the operator to be aware of surrounding physical objects, especially those that are occluded or out of view. [108] In contrast, a third-person view obtained from for example a camera mounted on a robotic arm is easier to control and results in fewer collisions [78].

Looking at general orientation, a user study involving wayfinding tasks found that a mobile 3-D map with third-person perspective leads to better clarity ratings, lower workload, mental demand and effort scores, and higher preference scores compared to a first-person perspective. Moreover, it leads to better pleasantness ratings, lower mental demand scores, and higher preference score compared to a 2-D map [98].
For robot operation, another study on rescue tasks in remote areas investigated the effectiveness of different camera images for mobile robot teleoperation. It comprises a true third-person view through a video feed and found that an image where the robot is positioned at the center of the camera view with a clear survey of the surroundings shows high efficiency in remote control of a mobile robot [1].
In terms of using virtual reality to increase telepresence, a study with an aerial robot found that augmenting the environment by adding the robot's camera view improved task performance in terms of accuracy and reduced number of crashes, while minimizing distractions in terms of the number of gaze shifts and total time distracted [39]. The tested augmentations included adding the robot's camera view as an outline, augmenting the robot by adding first-person view on top of the robot, or augmenting the user interface by adding first-person view in the corner of the GUI.
However, according to the paper, first-person perspective alone, such as through video display glasses, may degrade overall situational awareness that can aid in understanding operating context, such as identifying obstacles and other surrounding objects that are not in direct view of the robot [39].

While more research is required, this suggest that a third-person perspective is a safer, and more effective way for robot teleoperation and control. Adding first-person perspective information through augmented reality can improve task performance, while first-person perspective alone may degrade overall situational awareness.

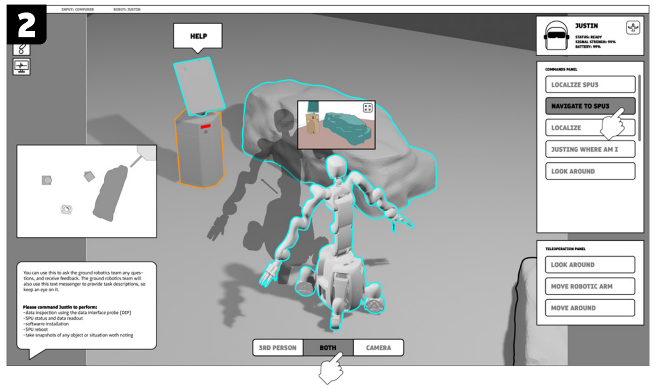
Figure 23: Visuals of third person perspective concept showing different approaches to the third person perspective and its effects on the UI: (1) augmented indicators and camera position (2) Ghost previews and camera position, (3) Camera position middle bottom, (4) and (5) both show view from behind and compass indicators



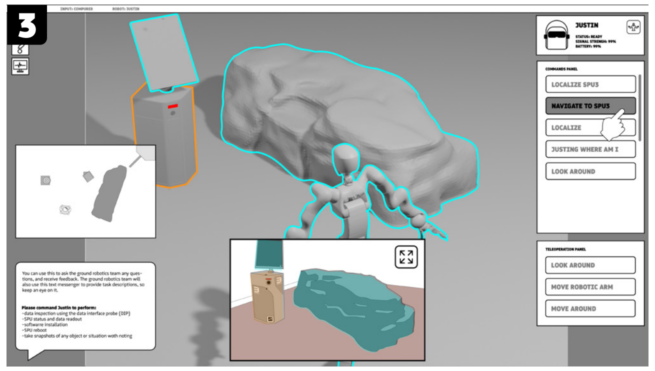
Concept Opportunities

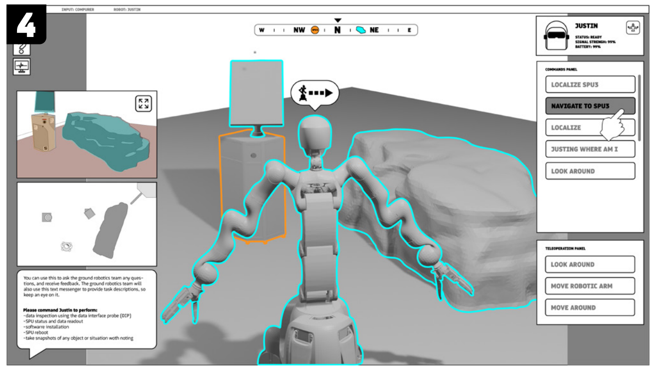
Generally, there are different ways of viewing the robot in its environment. Looking at video games, there are typically four modes of camera perspective: top-down/ isometric, first person, second person and third person[49]. Moreover, the camera can be static restricting the view to one angle or flexible allowing the user to have a free viewpoint [108]. As mentioned earlier, there are multiple options of creating the third person perspective for telerobotic however, considering that there is already a 3D environment at hand in the existing GUI that is used to create the map via a top-down view, it makes sense to use this option. This reduces the complexity of the concept and allows to focus on additional benefits of the third person perspective while keeping the flexibility of different vantage points. Moreover, keeping some form of first-person perspective in the GUI to enhance the user's awareness, as suggested by [39] and to keep the user alert of localization, should be considered.

Under the mentioned conditions, the Figures 23 show different visualizations on how the 3rd Person Perspective could look like in the GUI of Surface Avatar including features like motion limitations, augmented information, planning previews, navigational aids, and robot statistics. The visuals show different combinations of possible features. For example, (1) captures the idea of showing the robots motion limitations by highlighting the robotic manipulators that get close to singularity (like the red arm). It also covers the idea of allowing the user to choose their perspective and switch between third and first person or both. Moreover, it encapsulates the idea of showing augmented information like highlighting which SPU in the environment needs help or indicating the motion direction with an error.



Additionally, (2) shows how the third perspective could be used to show the robots planning in the form of a ghost preview. This allows the user to watch the robots planning process compared to watching a static screen which is currently the case.



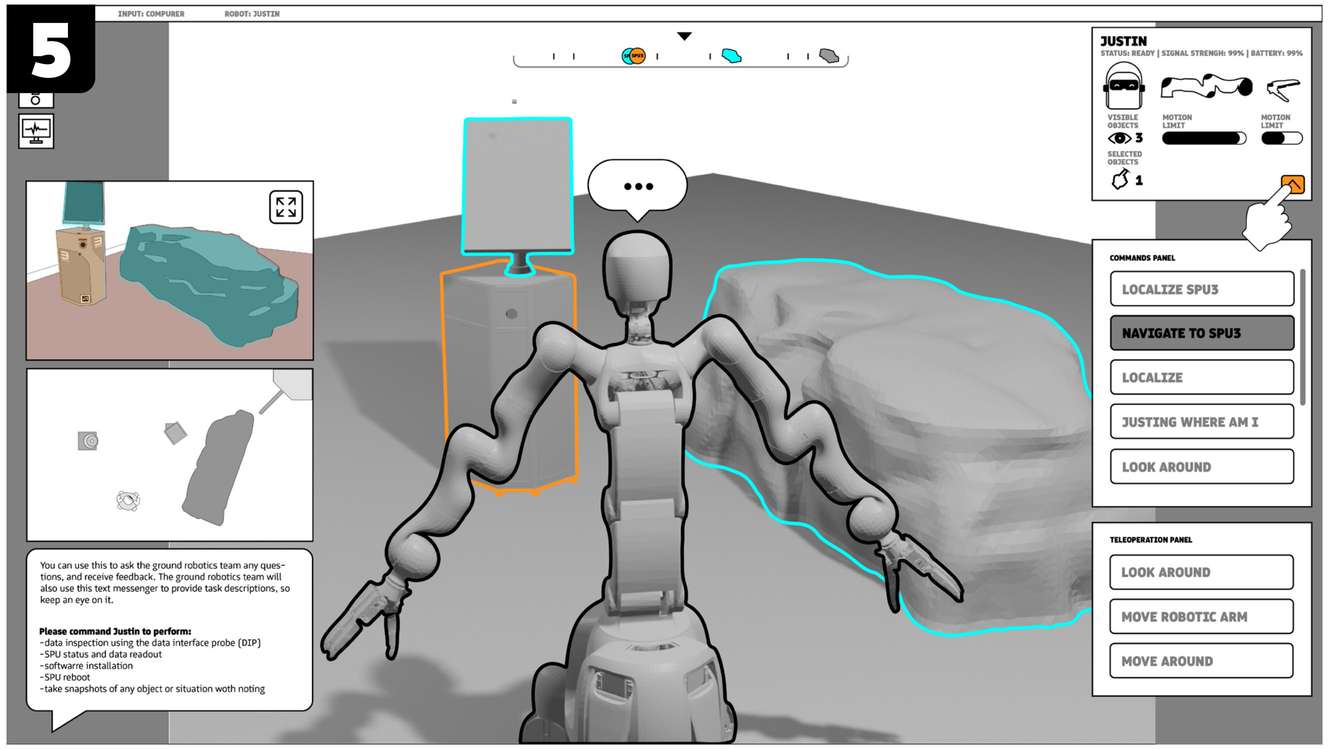


On the other hand, (3) displays a different way of combining first and third person viewing and similar to (1) and (2) tries to highlight located and selected objects by adding a colored outline with the same tint as localized object from the first-person view.

The images (4) and (5) both take a different vantage point for the third person perspective and explore how more orientation in the environment could be added with compass directions or waypoints/quest markers. Waypoints are navigational aids that indicate a specific location or direction, while quest markers indicate the location of objectives or targets for completing a task or mission [19]. Both are typical video game elements and could be used to direct the astronaut to the mission location and add information on the robot's facing direction. Moreover, both visuals (4) and (5) showcase how the robots planning could be augmented as a thinking bubble, so that the user is made aware of that step of the process more directly.

Image (5) moreover has a section in the upper right corner with additional robot statistics like motion limits or number of localized objects.

Which of these ideas and features should be considered for testing, will be investigated further under Reducing the Options, see p. 82.



Debugging Dashboard

This concept direction is designed to help users address errors as they occur by offering additional information and resources to aid in the error diagnosis process. Possible solutions in this direction include showing the robot's motion planning to identify where it became stuck, displaying real-time robot statistics like object detection and proximity, and providing typical problems and solutions. By providing more information and support during error resolution, this concept strives to help users improve their robot awareness and understanding of typical issues, ideally resulting in a more effective and efficient problem-solving process.

Why a Debugging Dashboard?
Given the wide range of possible interactions and unstructured, constantly changing environments in which mobile robots operate, it is highly unlikely to identify every possible type of robotic failure. Even trained roboticists may struggle to identify the cause of failures. A large study on failure handling in HRI reviewed 52 papers and looked at several factor influencing capabilities of a human to deal with errors when interacting with robots [38].

According to the paper, having an additional screen to display errors has become a popular feature in today's commercial robots [38]. A well-designed debugging page can help users understand the meaning of the symptoms or warnings provided by the failing robot, recall, and resolve the failure, and increase their positive evaluations and trust in the robot.

As such, it is essential to provide informative cues to help users recall and resolve a failure. According to research, the use of visual warnings with organized information groupings and generous white space [119] is more likely to hold attention than a single block of text. In more general terms, the review of the 52 papers suggests that information that is visual, concrete, repeated, specific, personal, novel, typical, humorous, and self-generated is more likely to be remembered by users [38].

It is also important to ensure that users can understand failure indicators the robot provides, as this will help them react appropriately. Background knowledge, wording, typographic design, felt involvement, motivation, expectations, training, experience, interface design, workload, and stress level can all impact users' comprehension levels, highlighting the importance of the design of the provided information [38]. There is limited research on what information should be communicated to help users cope with robotic failure situations. However, expressing physical limitations through motions, demonstrating appropriate emotions and awareness of errors, and having robots request help from a human partner have all been proposed as effective strategies [38].

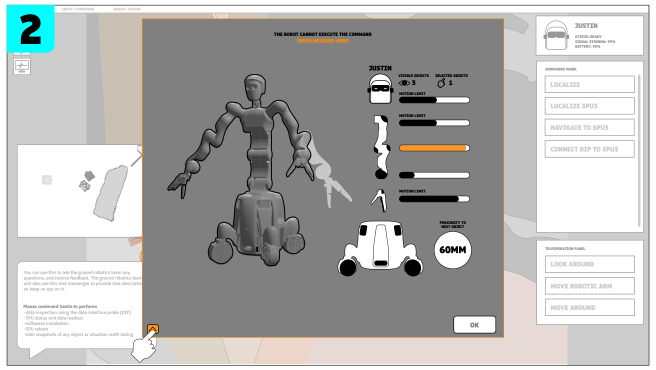
In conclusion, the complexity of robots and the environments in which they operate make it difficult to identify every possible type of failure. However, choosing the right information and way of presenting it may help in making a user more aware of their options and help them solve problems more efficiently.

Figure 24: Overview of the Debugging Dashboard concept; (1) Motion planning, (2) Robot statistics, (3) Typical problems and solutions, (4) Motion plan flowchart with animation steps and (5) Motion plan flowcharts as graph

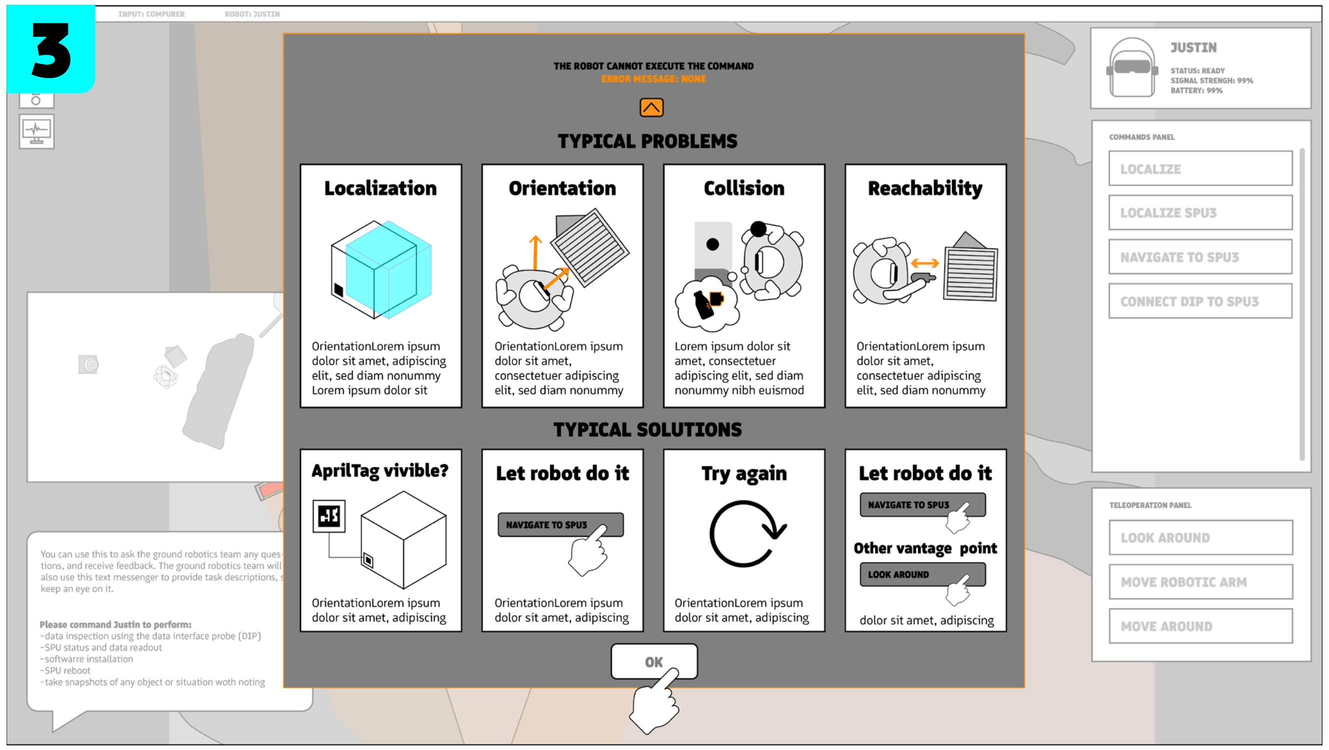
Concept Opportunities

Considering the mentioned aspects, Figure 24 shows different conceptions of potential debugging dashboards, comprising features like animated motion planning, robot statistics, typical errors, and planning steps. All proposals utilize the idea of expanding the existing error page to give the user more information. Considering that users might advance and not always need additional help, or that they might be overwhelmed if directly presented with a magnitude of information, this appeared like a sensible choice.

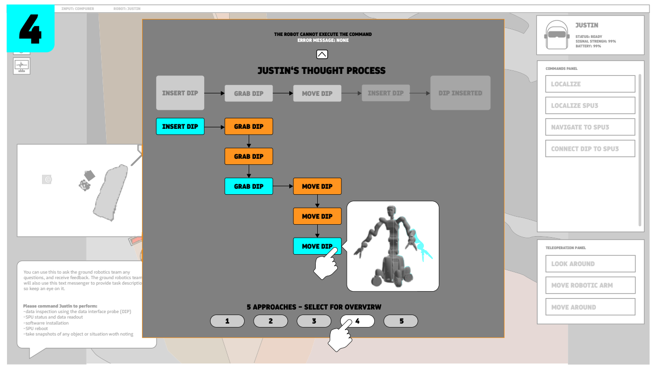
The main features of (1) cover a playable animation of the robots planning process including a slider that allows the user to move though the motion on their own terms. Since the robot plans its actions beforehand and only executes them if it finds a viable solution, the planning can be animated with the robot. This shows the user the motions the robot tried out and where it got stuck. This could for example help with collision problems, showing the user that the robot tried to put down an object randomly but could not find a free spot. Moreover, it could make reachability or orientation problems more apparent. Here the question lays in how to display this motion, from what vantage point and with how many additional elements. In the visualized example, the user can select how many additional objects they want to display. Choosing the vantage point beforehand and playing out a video or letting the user freely move the view by creating a 3D animation are further design choices in this example.



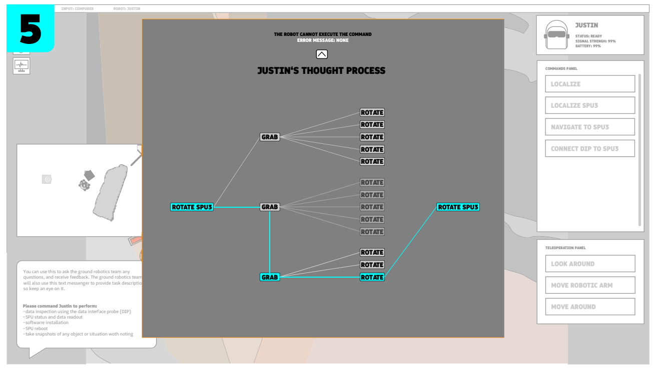
The illustration in (2) takes a different approach and explores how robot statistics could help in understanding errors. The idea of showing additional information, like the robot's motions limits, the number of located objects, and the proximity to the next object can indicate error origin. For example, reachability problems could be deducted by looking at the closeness to the next object. Moreover, the example showcases the robot's initial stage and last planning stage as an image. The question in this example is how easily understandable this information truly is, especially to a novice user and how to layout and visualize the information to get the most out of it as suggested by [38].



Moving on to (3) which tries to offer direct problem assumptions and solution directions. While it is difficult to predict or identify every possible type of failure, there are common errors that repeatably cause problems. For example, a common source of error in Justin is the improper localization of objects. By providing the user with typical problems and solution direction for this issue, this feature aims to create more awareness and offer direct feedback to the user. Again, the layout and design of this will determine the effectiveness, as well as the availability of known errors. Moreover, filtering the typical errors based on the robot's situation and only giving the user context specific suggestions will likely improve the helpfulness of this solution, but is also a factor of limitation.



Lastly, (4) and (5) aim to visualize, in different ways, the in-between steps the robot takes to solve a commanded task. When Justin is commanded a task, he splits it up into little sub-tasks and tries to find a viable solution for each. If he cannot find a viable solution for one of the sub-tasks he moves back to the prior sub-task. This information can be extracted and visualized. However, layout and choosing what information to display and what to remove is vital in this example as certain commands can be split in up to 10 sub-tasks and each sub-task could be iterated multiple times. Too much information could overwhelm and confuse rather than help the user. Nevertheless, showing at which subtask Justin got stuck, especially since they have human readable names, might be easier to understand than animating Justin's motion to that point.



Moreover, the animations could be split up for each step, giving a better overview of where and why Justin tried a certain movement, see (4). The idea of showing the subtask could also help the user understand Justin's working process more.

As with the third person perspective, this concept and the correlated features will be evaluated further under Reducing the Options, see p. 82.

Usability Add-Ons

This concept direction encompasses a range of small-scale changes and additions to the GUI and beyond that aim to enhance usability and provide contextualized support. Examples of these additions include a chatbot for immediate assistance, introduction videos for beginners, a heatmap overlay to improve teleoperation, or practical tips and do's / don'ts on a loading page. By offering assistive features throughout the GUI, this concept direction aims to help users increase their robot awareness, better utilize existing tools, and improve their overall experience with the system.

Why Add-ons?
The usability add-ons are a collection of smaller ideas that could improve the user experience and mitigate errors. On their own they might not all be sufficient as a solution, and they are not big enough to be called a concept in themselves. However, a combination of them or adding them to the already mentioned concepts might be a good solution. This makes it nevertheless hard to give an overarching reason to explain this concept direction, especially as some of them are very different.

In a nutshell, all the ideas from this direction are generally based on mental models.
Mental models are structures of organized knowledge that allow individuals to interact with their surroundings. For a human, they play a critical role in describing, explaining, and forecasting events in the environment [105].

Users of commercially available robots usually do not have any specialized knowledge of the robot's internal workings, and they often rely on observable factors such as appearance and behavior, which together form a "system image". These factors serve as the foundation of their mental models that explain and predict robot behavior [71]. Users can form wrong impressions of the robot's sensing abilities due to preconceived ideas, obscured sensors, or the robot's deceptive actions. Those who have more precise mental models of intricate systems are better equipped to use them effectively [71].

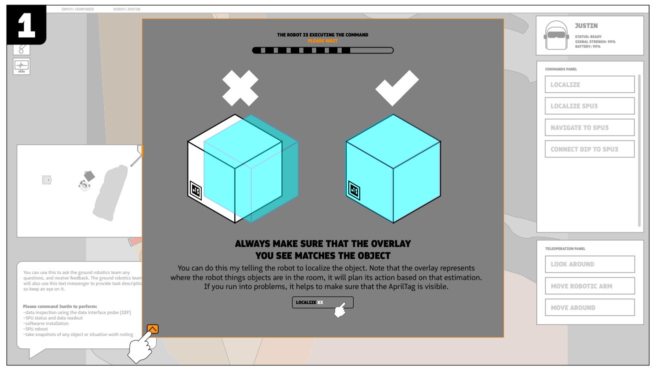
However, as mentioned in the related work section, most of the research regarding mental models in HRI focus on adjusting the robots' behaviors based on user actions or recursive reasoning, where a robot forms a belief about a human's model of the robot [105].

By adding better cues throughout the user interface, providing more information to the user, or setting them up with a better understanding of the system before the start, these mental models could be shifted in the right direction.

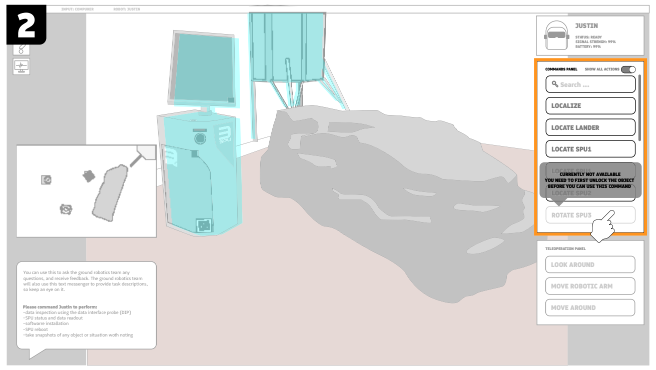
Figure 25: Overview of concepts for the Usability Add-ons: (1) Loading screen, (2) Hover indicators, (3) Tutorial introduction, (4) Chatbot, (5) Heatmap for depth perception, (6) utilizing cameras of other Robots and (7) robot embodiment through manikin.

Concept Opportunities

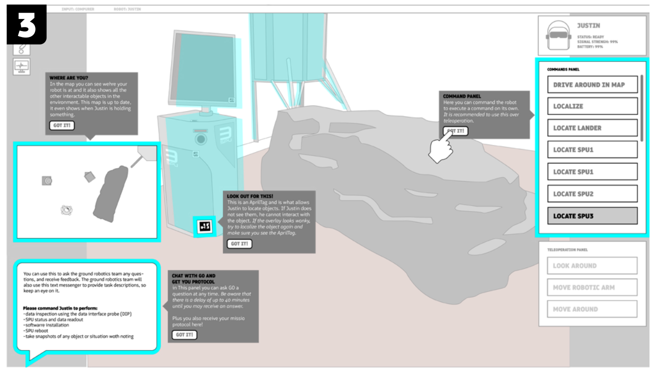
In this section all the features and smaller concepts of this concept direction will be discussed. Figure 25 provides an overview of all of them, including loading pages, showing hidden information, better head starts through tutorials or intro videos, chatbots, disparity maps, utilization of other robots, and a manikin version of Justin, in that order.



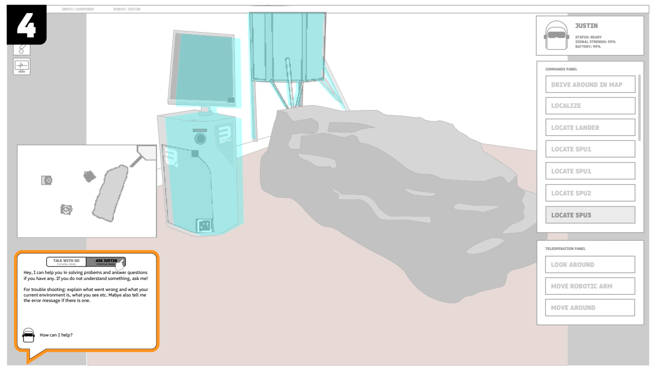
The first, (1) showcases a loading screen and is one of the ideas from this direction that could stand alone. The idea behind this is to utilize the unavoidable waiting time when the robot is planning. First this idea makes the user roughly aware that they need to wait and for how long, ideally reducing confusion in the user when nothing happens. Second, this idea tells the user typical problems and useful usage tips about the robot, ideally creating more awareness, mitigating errors and making use of time that would be otherwise wasted.



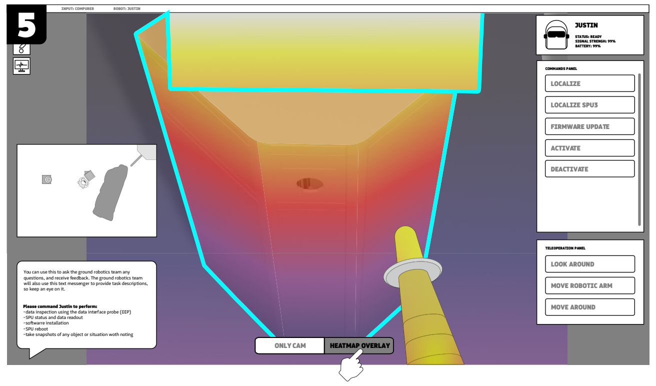
On the other hand, (2) is a more subtle approach and should be combined with other solutions. It tackles one specific error type which is preconditions. The idea behind this is to allow the user to show all possible commands and gray out currently unavailable ones. If a user then wants to do an action, they can still search for it with the added search function of this idea and hover over it if it is unavailable, to receive an explanation.



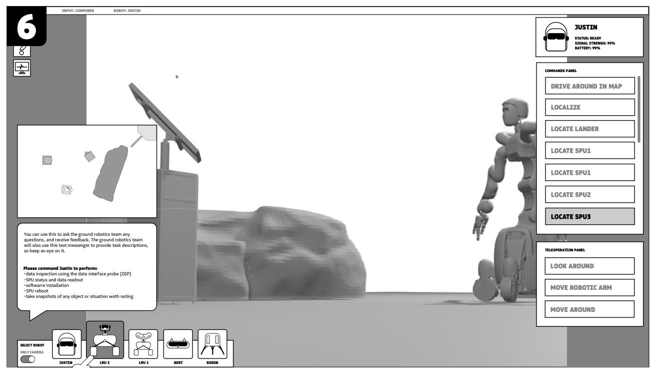
Like this (3) should too be combined with other solutions. It encompassed the idea to give the user a better head start, by either integrating a tutorial like introduction into the GUI or simply providing a video that explains the most important things shortly. According to literature, setting expectations by forewarning participants of the robot's abilities can improve evaluations of the robot and judgments of the quality of the service [54]. Moreover, participants who saw a robot stating its limitations before asking for help reported liking the robot more than those who only saw control statements in the moment of error [38].



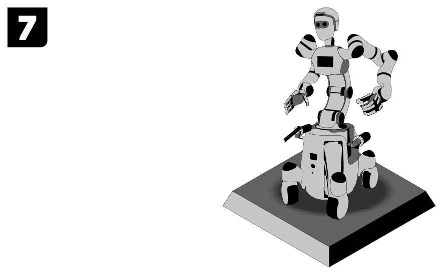
Moving on, (4) is a solution that could stand alone. It makes use of the rapid improvement of natural language processing (NLP). The idea here is to give the user the option to talk to Justin and ask him questions even in the case of errors. If the NLP-model used to do this is trained on the robot's source code, typical problems, and general functionalities, it should be able to give easily understandable solution directions. Moreover, this would reduce the need to ask GO for help and remove the waiting time of receiving an answer.



Another more specific solution approach can be seen in (5) that utilizes disparity maps to give the user a better depth perception. This could be helpful during teleoperation tasks when the user needs to step in, since the current camera makes depth perception difficult. Moreover, it could be used as an analysis tool to understand better why the robot failed. However, this solution focuses only on a small aspect and can only be considered as a supplement to other ideas.



Similarly, (6) tries to give the user more environmental awareness by allowing them to switch to the camera perspective of other robots in the surrounding. Considering that in the future not only Justin but also other robots should be controlled via the GUI, the idea is not far-fetched. However, this again only focuses on a small criterion that needs improvement and can too only be seen as an addition and not a full solution.

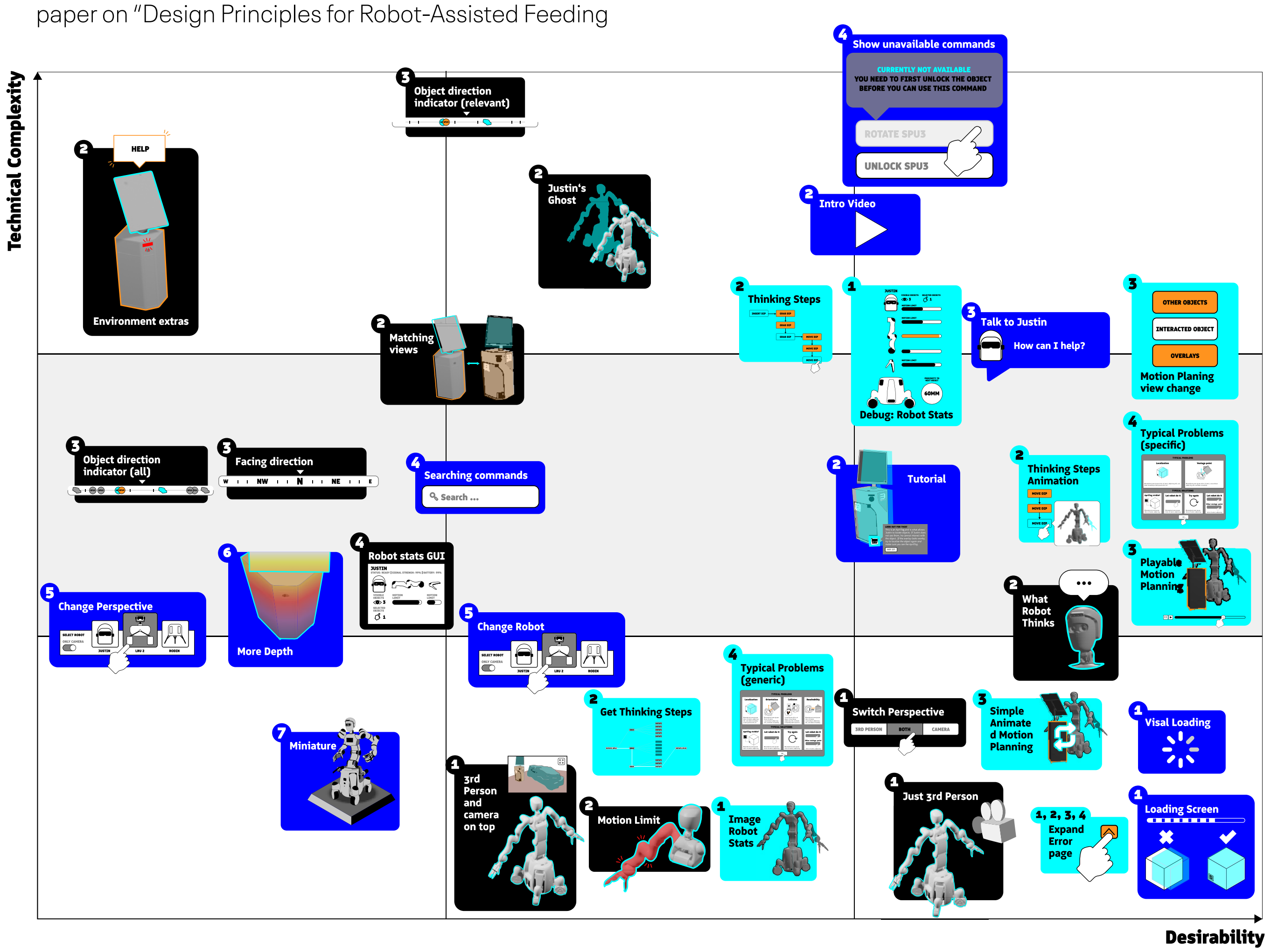


Lastly (7) plays with the concept of embodiment by providing the user with a to-scale manikin of Justin. For this idea, the manikin is moveable and allows the user to explore the motion limitation and joint configurations of Justin. One paper explored how the Mars Exploration Rover Mission team used visualization and embodiment techniques to create a sense of presence and connection between the scientists on Earth and the rovers on Mars. While the paper does not directly mention physical models, it does discuss the importance of embodiment and visualization in creating a shared understanding of a robotic system. The authors suggest that embodying the robotic system through visualization tools can help users better understand and engage with the system [114]. Therefore, by physically interacting with the manikin, users could visualize themselves as the robot, gain a better sense of what it might be like to perform tasks as the robot and learn about its way of moving (including the limitations). This could ultimately lead to a better understanding of the robot's capabilities and limitations, and better utilization of the robot in various tasks.

Again, whether this concept and any of its features are feasible or even helpful will be further evaluated in the next section Reducing the Options.

Reducing The Options

Considering that not every idea from the three concept directions has the same potential impact or feasibility, the need to be weighted. This further helps in deciding which of the ideas to peruse and develop further. Therefore, each concept including the associated features have been evaluated on a matrix of implementation complexity and desirability, see Figure 26.
The concepts are color-coded and the numbers in the top left corner of each colored shape indicate correlation. For instance, there are two approaches of animating the motion plan, either by creating a video file or by making it interactable. Both implementations have different advantages but also contrast in their implementation complexity. The matrix for evaluation is based off a paper on "Design Principles for Robot-Assisted Feeding in Social Contexts" [8] where a team of researchers with implementation experience evaluated ideas from participant sessions based on their "Technical Complexity" and "User-Expressed Priority". This evaluation of the concept allows to narrow down the options and gives a good overview for further decision making. To evaluate the technical complexities of each concept and the connected features, a session with Dr.-Ing. Daniel Leidner [55], leader of the group for Fault-Tolerant Autonomy Architectures at DLR was conducted. During the session all visuals presented in the section before were evaluated with a focus on each feature and the requirements to implement them.



The evaluation yielded interesting results. Overall, it showed that the general idea of having a third person perspective, the option to switch perspectives, an error page with additional information, a loading screen with tips (or simply a loading screen), producing a video of the robot's motion planning as well as augmenting the robot with additional information like a thinking bubble are highly desirable and technically easy to implement. Other more complex but desirable ideas are robot statistics in the debug dashboard, showing the planning steps, suggesting typical fixes tailored to the problem, making a more interactive motion planning animation, making a 3D motion planning that allows view change, having a robot chatbot or creating an introduction tutorial.

Moreover, the session with Dr.-Ing. Daniel Leidner pointed out previously unknown limitations. For example, a ghost preview of Justin is highly complex. Justin takes some time for planning and during the planning phase it is not possible to directly showcase that. So, a ghost preview could only be possible after the planning. While the ghost preview might still be interesting if an error occurs, users must wait either way and it would still be technically complex to implement. Another interesting limitation is that not all missions are predictable, meaning that augmented information on for example a broken SPUs is barely possible to implement. While this might be possible on earth, on mars everything is a bit more limited. Ground Control simply cannot know which object in the environment might need tending, the user has to actively look for it. Additionally, it is currently not possible to clearly state in natural language why certain actions are unavailable, ruling out the option to show and explain unavailable commands. The same goes for clear error messages that explain the user why an error occurred and highlights the need to create better awareness of the robot as well as other means of communicating errors. Another interesting insight came from the chatbot idea. This option could be trained to a degree that it would help the user significantly, but the helpfulness of the robot scales directly to the labor it takes to train the NLP model. The better the chatbot, the more effort is required. A more surprising insight came from the idea of having a video introduction. According to Dr.-Ing. Daniel Leidner, the Surface Avatar system changes constantly. Therefore, they would need to adjust the video material continuously making this idea highly infeasible for them. Moreover, the information required for a disparity map is already in the system as the robot needs it. However, showing that in the interface is limited by bandwidth and would slow down transmission.

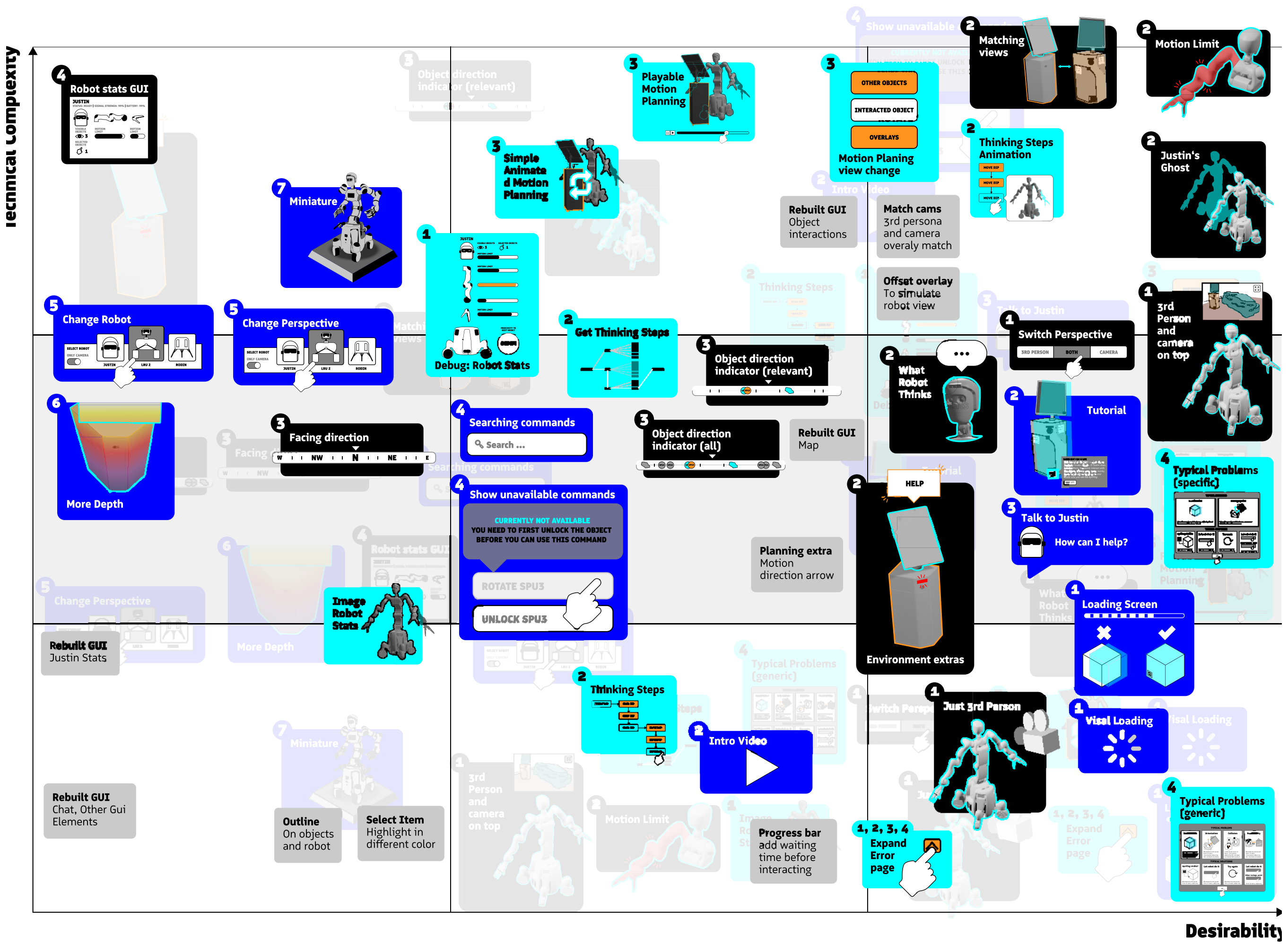
On more general term, it was discussed that the current coding of the GUI system makes it difficult to layout and create interactive solutions. The general retrieval and display of certain features might not be technically complex but certain visualizations might take more effort. For instance, highlighting the robot's arm in red when it gets close to singularity is difficult but generally informing the user that the right arm is close to singularity is easy.

Figure 26: Evaluation of concepts and respective features based on technical complexity and desirability: Black boxes are all features of the 3rd Person Perspective, cyan boxes are all part of the Debugging Page concept and the blue boxes all belong to the Usability Add-ons.

Mapping Prototyping Feasibility and Desirability

Aside from this investigation, the same matrix was implemented based on prototyping complexity. While some ideas might be easily implemented in the real system, testing cannot easily be done with the existing setup, at least not all the time. Therefore, it helps to keep in mind the amount of effort it takes to create a testable prototype for each concept and feature. Additionally, not all concepts seemed equally desirable to the expert from DLR and the author of this work. The matrix for this evaluation can be seen in Figure 27.
This visualization further includes prototyping steps in gray, that are required to create the impression of the prototype being the same or similar to the real system but are not directly connected to any of the concepts. For easier comparison, the mapping on implementation is semi-transparent underlaid.

There are a few overlaps with the implementation matrix. For example, the loading screen, expandable error page and general idea of a third person perspective are easily prototyped, implemented and perceived equally desirable. Also, the tutorial, direction indicators, suggesting typical fixes tailored to the problem, disparity map and searching commands are similarly situated.



Differences show in the motion planning animation, no matter what form they take. It is considerably harder to set up all the requirements to read out and animate the available model of the robot for a prototype than it is to simply implement it in the system. The Surface avatar system has this already set up, but it cannot be easily transferred. While people at DLR could set up the animations and create them as mp4's to use in a wizard of OZ experiment, they themselves are currently struggling to do so. A simpler solution to test this idea is to fake the motion in a standard 3D animation environment instead of using actual error logs. This could be feasible if they resemble the actual motion planning.

Most of the aspects of a third person perspective can be relatively easy implemented with programs like Unity or Unreal engine, both are popular game development platform that allow for the creation of immersive 3D environments. Nevertheless, other aspects like highlighting motion limits or matching the views from first person and third person perspective require a lot of effort. The ghost preview too is not that simple as it is limited for the same reasons as visualizing the motion planning.

Additionally, there are some discrepancies in what the expert from DLR deemed desirable and what the designer (author) expects to be more helpful to users. For instance, animating the motion planning is not expected to be as helpful by the designer as it is by the DLR expert.

The actual helpfulness will be determined by tests with chosen prototypes, see Prototype Development & Elaboration. However, not all concepts will be evaluated further. For the concepts that will not be tested, it is left to the reader to determine which features they find more helpful or desirable based on the two matrices provided.

Figure 27: Evaluation of concepts and respective features based on prototyping complexity and desirability (implementation evaluation underlaid in semi-transparent): Black boxes are all features of the 3rd Person Perspective, cyan boxes are all part of the Debugging Page concept and the blue boxes all belong to the Usability Add-ons.

Prototype Elaboration and Development

In order to explore and evaluate design ideas for enhancing error handling, two prototypes were chosen to be implemented: the third person view and the debug page concept. Prior to their implementation, extensive research was conducted to inform the development of these prototypes.

By conducting thorough research and creating visual overviews of the ideas explored, informed decisions were made regarding the detailed implementation of the chosen prototypes. The subsequent sections will delve into the particulars of each prototype, discussing their further research, purpose, development process, limitations, and testing purpose.

Mitigating Errors: Third Person View

Introducing the Third Person View Prototype: This section explores the further research done for the implantation of a high-fidelity prototype. The research is summarized in the form of a morphological chart [113] of potential communication aids to enhance error detection, planning, execution, and object interactability. Moreover, the prototype implementation and testing are further discussed to assess the impact of situational awareness on user error handling capabilities.

Further Game Element Research:
For the third person view prototype, the research focused on investigating game elements, visual cues, control mechanisms, and embodiment concepts. A matrix was created to explore different versions of the third person view with various additions that could potentially aid in error detection, planning and execution waiting time, object intractability, and localization.

Why Gaming: Video games are not just widely enjoyed and financially prosperous; they are increasingly integrated into mainstream cultural consciousness [11]. Joyce (2019) showed that Jakob Nielsen's 10 heuristics for user-interface design [85] can be applied to video game elements, allowing designers to make informed decisions to utilize game elements in user interface design [47]. Moreover, research by Granic et al. (2014) found that, particular game rules and mechanics support the advancement of cognitive abilities like neural processing and efficiency, problem-solving proficiency and spatial skills [7].

Choosing Game Elements: Research by Dillman et al. (2018) analyzed 49 contemporary videogames and classified game elements by (1) Purpose (are the cues there to help discover, guide etc.), (2) Markedness (How do the elements "stick out"/grab the user's attention), and (3) Trigger (what is the source of the cue, e.g., other actors or sounds) [50]. Rueben et al. (2021) focused on the use of video game cues in communicating agent vision to humans. They identified effective design cues, including the behavior and facial expressions of other actors, communication cues, atmospheric sounds, user perspective changes, and visual effects, as established methods within gaming. [72]. Other Research suggests that for human robot interaction embodied cues, such as verbal-, gesture-, social-, or gaze cues, have a significant impact on learning and development, attention and engagement, motivation, compliance, and persuasion [77]. Lastly, others discuss the choice of camera perspective and use of cinematographic elements to guide the user's attention [13].

Applying it to the Robot: Based on the researched game elements and embodied cues, a morphological chart of potential cues and additional elements for the third person perspective of Rollin Justin were composed. The Chart can be seen in Figure 28 and is split into different states and communication needs of the robot: (1) General interaction cues (planning and execution time, communication of the protocol, finding targets, showing interactivity of objects, showing localization of objects). (2) error communication (general planning and execution errors, specific cues per type of error, communication after fixing an error), and (3) Setting up the user (motion limitations of the robot and games cues to facilitate correct usage). An enlarged version of Figure 28 can be observed on the next page.

It was determined that implementing game elements into the prototype would fall outside the scope of this research and could potentially interfere with comparing it to the first-person view. As a result, the decision was made to visualize selected aids from the matrix and gather participant preferences through a questionnaire during the evaluation sessions. This allows for informed recommendations for future research, see Further Recommendations, p. 134.

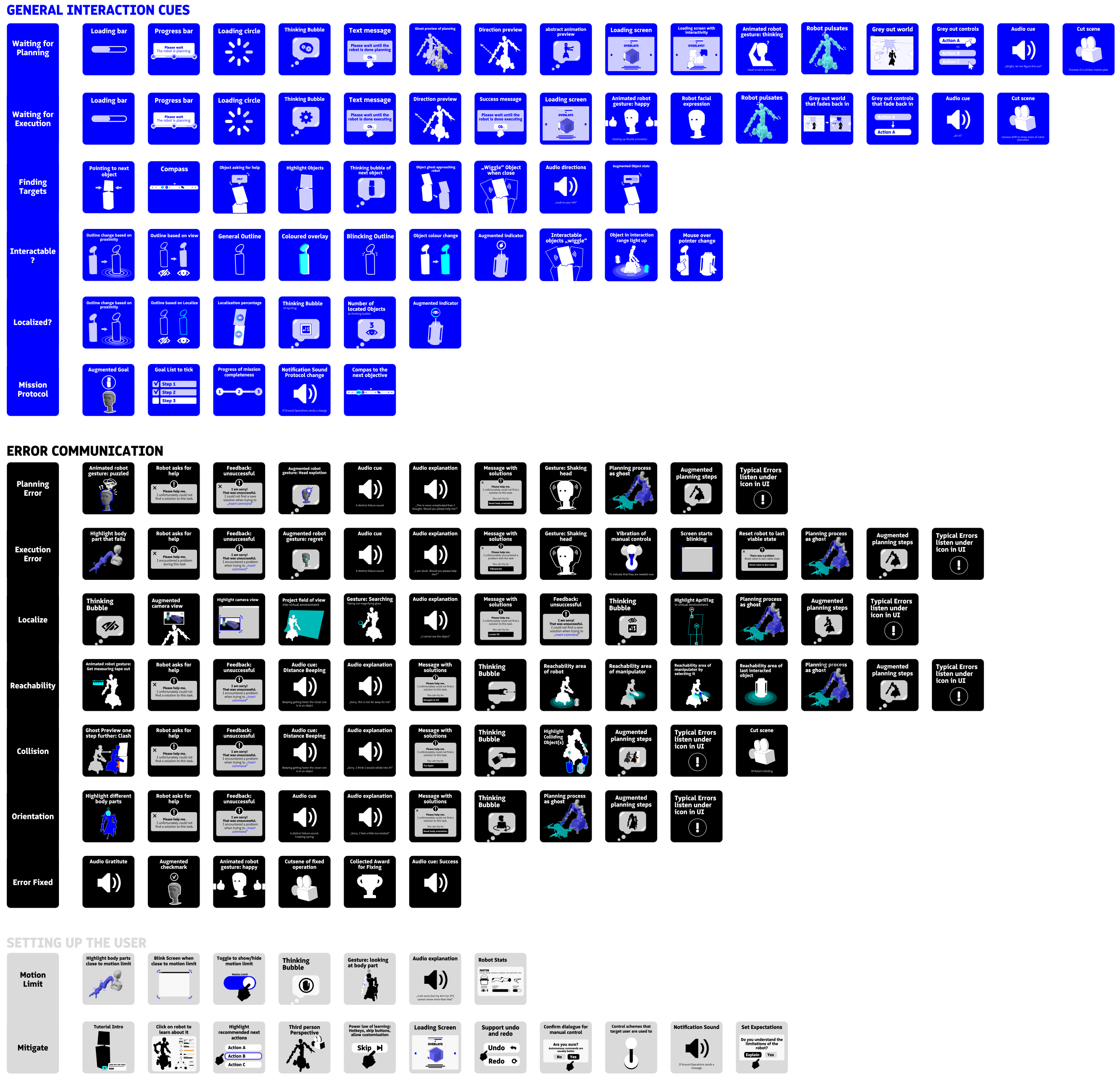


Figure 28: The Chart, illustrating the robot's various states and communication requirements: (1) General interaction cues (planning and execution time, communication of the protocol, finding targets, showing interactivity of objects, showing localization of objects), (2) error communication (general planning and execution errors, specific cues per type of error, communication after fixing an error), and (3) Setting up the user (motion limitations of the robot and game cues to facilitate correct usage).

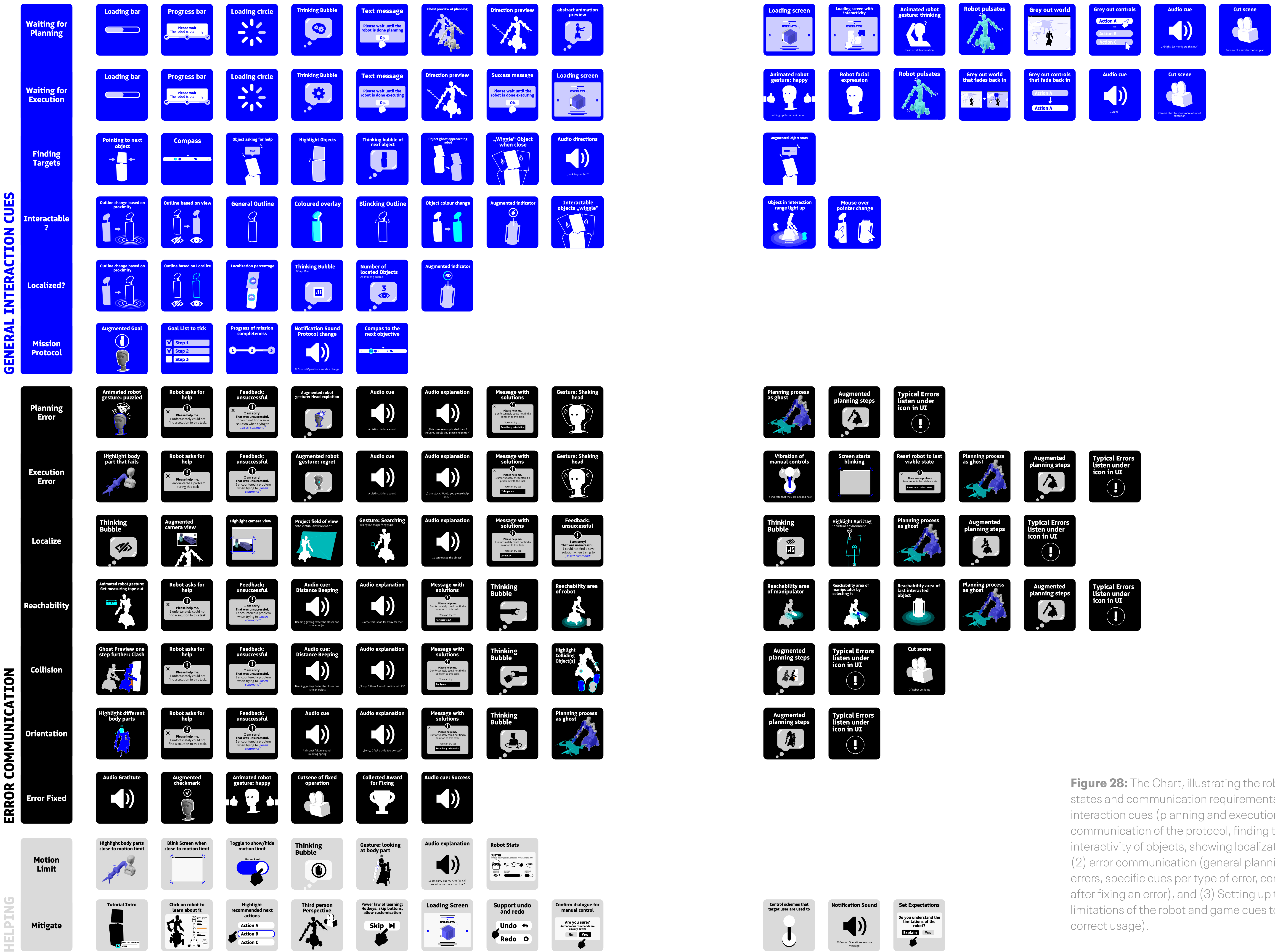
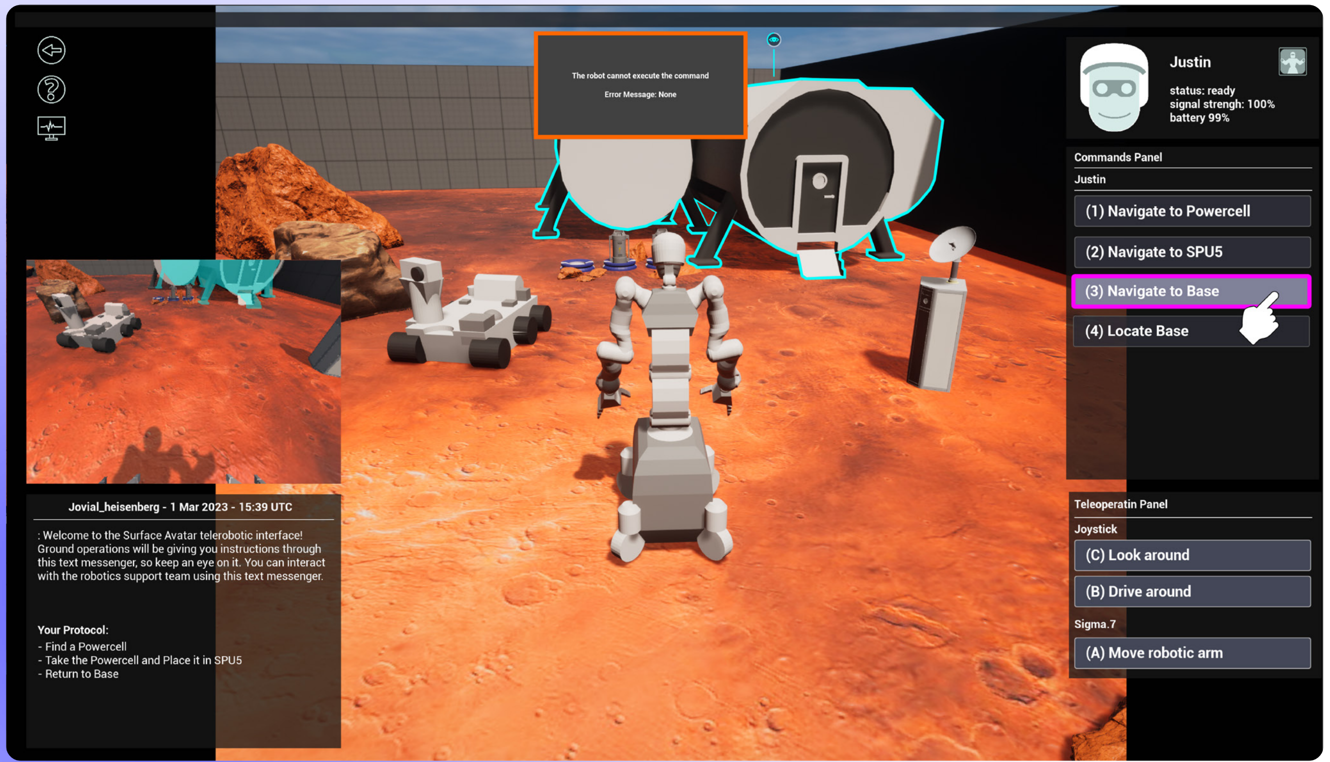


Figure 28: The Chart, illustrating the robot's various states and communication requirements: (1) General interaction cues (planning and execution time, communication of the protocol, finding targets, showing interactivity of objects, showing localization of objects), (2) error communication (general planning and execution errors, specific cues per type of error, communication after fixing an error), and (3) Setting up the user (motion limitations of the robot and game cues to facilitate correct usage).

Development of the Third Person Perspective
The third person view prototype was implemented with the intention of leveraging elements from the gaming industry to enhance user experience and error detection. Its main purpose was to assess the effect of situational awareness, induced by perspective on the error handling capabilities of the user.
To accurately compare, two versions of the prototype were developed using Unreal Engine [110]. The first version replicated the existing system's first-person perspective and a top-view camera map, while the second version encapsulated the proposed solution with a third-person view and the first-person camera view instead of the map.

Implementation: The prototype included autonomous navigation, object interaction, manual head and base control, camera change, contact with Ground Operations, view of the robot's camera in the virtual environment, and a virtual third-person perspective with the freedom to change perspectives, see Figure 29. The entire user interface was rebuilt based on the existing system, except for teleoperation of the arm, which remained in its original place but did not have functional interaction.

Limitations: To streamline testing, certain features like waiting times for planning and teleoperation were removed. Additionally, localization was not fully added due to implementation limitations and manual control was simplified using the WASD keys instead of a joystick.



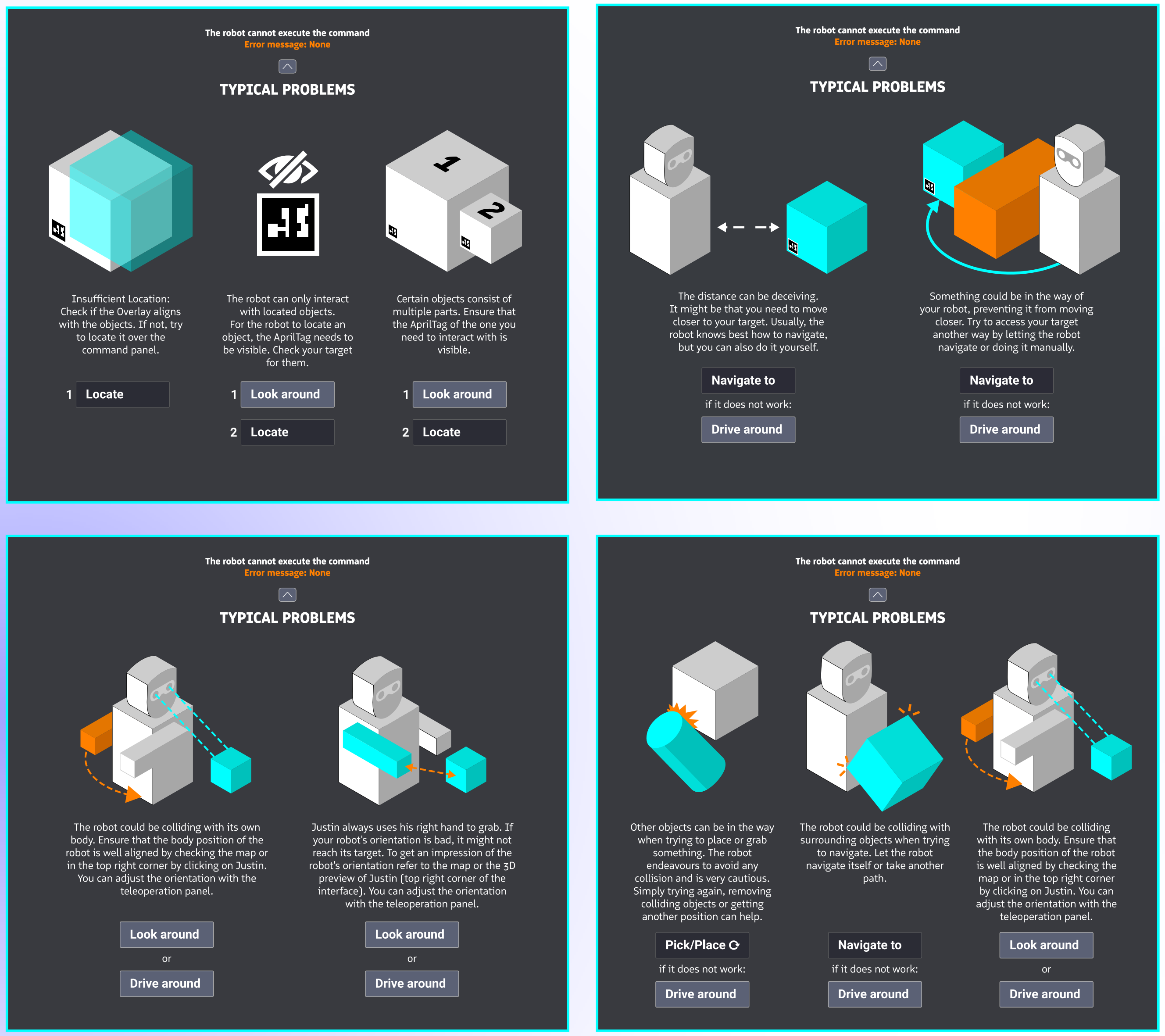
Virtual Environment: The simulated environment incorporated real Mars surface scans, rovers, tools, and other objects relevant to the mission. The prototype was designed to replicate the experience of astronauts encountering unfamiliar objects while using Justin. As the participants were not going to be actual astronauts but roboticists from DLR who were somewhat familiar with Justin's development, additional objects were introduced in the environment to simulate the presence of unknown elements. These included a Mars base, loading stations, power cells, SPUs with different appearances, drones, and an increased number of rock formations compared to the current testing site at GSOC. This approach aimed for participants to face and interact with objects they were not previously familiar with, resembling the situation astronauts might encounter during their missions.

Testing purpose: A mission protocol was created for testing the prototype, involving a realistic Mars mission task. Mission control provided the following protocol: Searching for a power cell, installing it in SPU5 (Solar panel unit that looks different to the current), and returning to the base. Each subtask was designed to force participants into specific planning errors, such as approaching the power cell from an unreachable side or encountering collision issues during navigation. The same protocol was tested in both views (first-person vs. third-person view) with different participants to assess the effects of the perspective.

Figure 29: Third person simulation built in Unreal, with the same User Interface as the current System

When Errors are Unavoidable: Debugging Help

Introducing the Debugging Help Prototypes: This section explores the development of two types of debug pages for Rollin Justin. The first approach includes error-specific typical issues and solutions, providing visual representations and concise explanations. The second approach involves animated motion planning videos that showcase the robot's failed actions and interactions after an error occurs.



Further Research on Error Communication
The development of the debug page also further investigated research on error communication, however, due to the limited research on error communication already mentioned in the Related Work section, some ideas of the game cue morphological chart were further considered. Below a quick summary of the existing research is shown.

Existing Research: As mentioned in the Related Work section, to improves failure identification and solution scores incorporating the history of recent actions and environmental reasoning in explanations has shown to be effective [24]. Using multiple modalities, such as audio and visuals, enhances comprehension and attention. Limiting information overload and applying Gestalt principles by combining visuals with text or spoken words improves attention and understanding [38, 40]. Organized visual warnings with ample white space are more effective in capturing and maintaining attention compared to dense blocks of text [119].

Development of Two Types of Debug Pages
The debug page prototype consists of two approaches: **(1)** an error page that provided to some degree error specific typical issues and solutions and **(2)** a debug page with animated motion planning of the robot, based on the failed plan.

(1) Typical errors and solutions: The error messages were designed to provide simple visual representations of the error, short explanations, and direct solution propositions, see Figure 30.

Implementation: This debug page was designed using Illustrator and Figma. Figma allowed the creation of an interactive prototype of the current user interface, including the proposed error messages.

Limitations: The Figma prototype provided more static feedback through a click dummy compared to the dynamic simulation of the third person view. However, its purpose was to investigate if the additional information aided participants in solving the presented problems, and therefore did not require this level of control.

Testing Purpose: This prototype was design for participant sessions to evaluate how efficiently the provided information aided in different error scenarios. Moreover, animated videos of this prototype were used in a questionnaire environment to compare to the (2) animated motion planning, see section Questionnaire Results, p. 126.

Figure 30: Typical errors and solutions design: localize, reachability, collision and orientation from top left to bottom right.

(2) Animated motion planning: This prototype offers more error-specific features by animating and rendering the robot's planning process as a video after an error occurs. Users can witness the robot's actions, like attempting to reach objects, while having the option to control and explore the animation using a slider. Additionally, objects that the robot last interacted with are included in the animation, see Figure 31.

Implementation: The concept of using motion planning information as an error message involved Blender and After Effects. A model of Justin was created, rigged, and animated in 3D using Blender to visualize the robot's motion planning. The motion planning videos were then combined with After Effects to integrate them into the user interface. The resulted videos showed how a user interacted, had an error and used the motion plan error message.

Limitation: There is no interactivity or control for the user for this prototype, however it was not needed, as this prototype was only tested in the questionnaire environment.

Testing purpose: This prototype was tested in a questionnaire environment to compare its success to its (1) typical errors and solutions counterpart, see section Questionnaire Results, p. 126.

Overarching Testing Purpose:
Both prototypes were designed for four error scenarios: Localization, Reachability, Collision, and Orientation errors. This way, participants could be presented with each scenario and identify the cause of the issue.

Figure 31:
Robot motion planning design: localize, reachability, collision and orientation from top left to bottom right.

Prototype Testing

A testing plan was set up to assess the effectiveness of the two prototypes in mitigating errors and improving error detection rates, cognitive workload, user experience, and engagement. For the Third person perspective prototype, an additional objective is to examine the impact on situational awareness and user behavior. For the Debug pages, the main objective is to determine their effectiveness in relation to different error causes and compare their performance. Additionally, the research aims to evaluate the sufficiency of the provided information and identify areas for further improvement. The following sections outline the research setup, participant demographics, conducted tests, and data analysis approach.

Research Setup

Sessions conducted at the DLR facilities at Munich, Oberpfaffenhofen, and (2) A/B between-subjects testing conducted through an online questionnaire.

The research conducted in this study was approved by the Human Research Ethics Committee (HREC) at TU Delft. All research procedures adhered to the approved data management plan, and participants were requested to provide informed consent by signing a consent form. It should be noted that participants volunteered for the study without receiving any form of remuneration.

The participant sessions spanned five days and involved a total of 16 participants from DLR. Eight sessions were dedicated to detailed prototype testing of the third person simulation and the typical error and solution debug page. These individual sessions lasted approximately 1 hour and included participants using both prototypes, completing questionnaires, and participating in interviews (see next page more details). The remaining eight sessions served as a comparison group for the third person prototype, where participants used the same prototype in first person mode and completed identical questionnaires. These sessions took approximately 15-20 minutes each.

In addition to the participant sessions, A/B between-subjects testing was conducted entirely online and did not involve DLR employees. This testing aimed to compare the existing system with the proposed solutions based on correct error identification rate and choice of action to solve the error. In this questionnaire, the debug concept of animated motion planning, the debug concept of typical errors and solutions and the third person perspective were included. Participants were randomly assigned to either the existing system or one of the proposed solutions and were asked to complete the questionnaire based on their assigned condition.

In the context of the online questionnaire, participants were informed that their anonymized data would be used and that by sending their answer they consent to the participation.

Participant Demographics

For the test sessions, all participants were employees at the DLR, encompassing individuals from various backgrounds and expertise. They included members of the development team, individuals with prior dry-run experience using the existing system, and those who had observed astronaut trainings. Some participants had no previous experience with the robot. Both the prototype testing group and the comparison group were evenly distributed among the participants, see Appendix p. 186.

Regarding the online questionnaire, specific participant criteria were applied. The target demographic consisted of individuals aged between 26-60, reflecting the age range of astronauts. Participants were sought with moderate to high levels of technology experience, given it is expected of astronauts. Preferred occupational categories included engineering, science, health, biology, and related fields, which closely align with occupations commonly associated with astronauts and individuals involved in space-related activities. For an overview of the participant distribution, refer to Appendix p. 191.

It is important to note that none of the participants who took part in the test sessions were asked to complete the questionnaire. This decision was made to prevent potential biases or influences on their responses, ensuring an independent evaluation of the questionnaire results. The assignment of participants to Version A or Version B of the questionnaire was randomized. Participants were not provided with any specific information about the differences between the two versions to minimize potential biases resulting from preconceived notions or expectations.

Prior experience with the robot was not a requirement for participation in the A/B study. This decision was made considering that astronauts typically have limited exposure to the surface avatar system too.

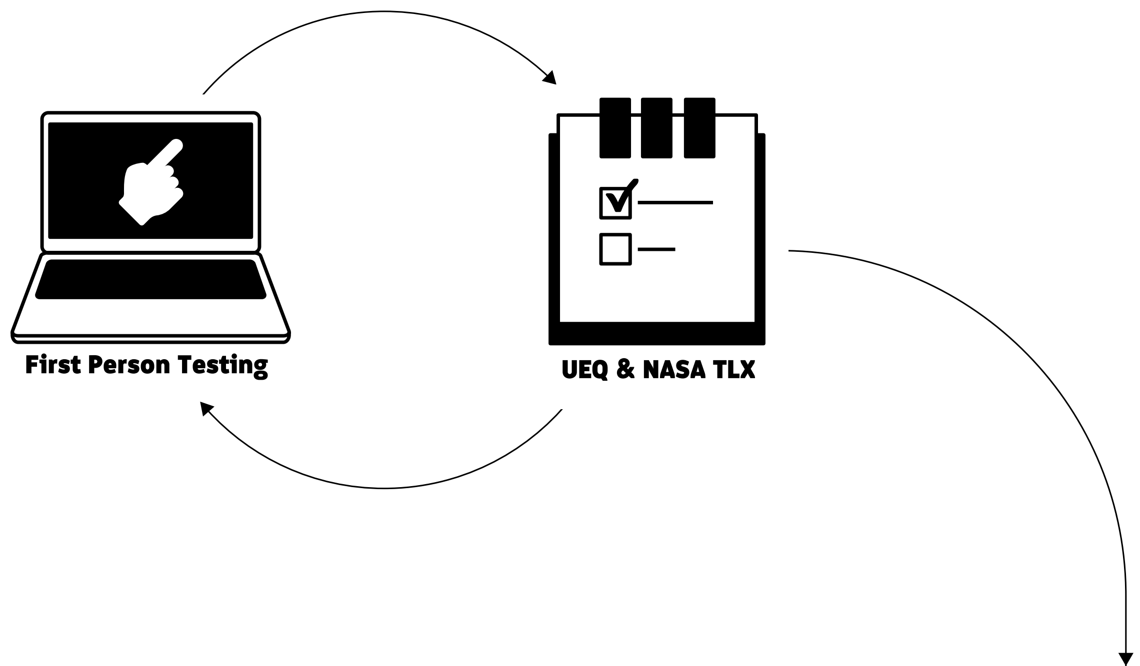
Prototype Testing Sessions

The overall process of the sessions can be seen in Figure 32. It was split into two parts, (1) the testing of the third person prototype and (2) the testing of the debug page with typical errors and solutions.

The sessions started off with the consent form followed by the testing of the third person prototype where each participant followed the same protocol. After finishing the protocol, they received two questionnaires, the NASA Task Load Index (NASA TLX) [35] and User Experience Questionnaire (UEQ) [102], to fill out based on their experience with the prototype. This was followed by a discussion on additional game elements that could be incorporated into the third person prototype and ended in a more detailed interview.

The second part allowed the participants to interact with the debug page prototype in four scenarios, followed by the same questionnaires, NASA TLX [35] and UEQ [102] and ended with a detailed interview on the overall experience.

To compare the results of the third person prototype another set of different participants did the same protocol in first person mode and completed the NASA TLX [35] and UEQ [102] based on their experience.



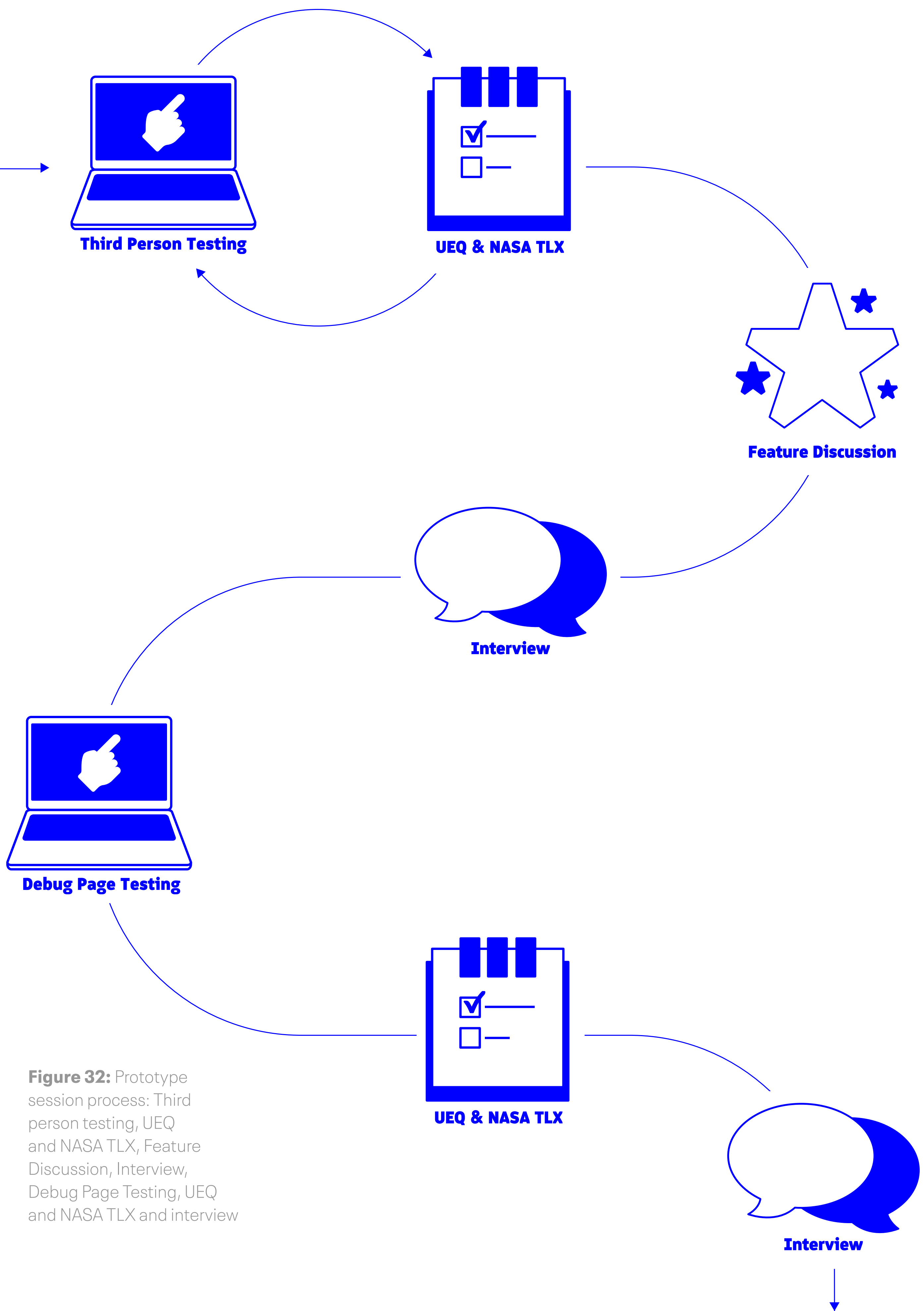


Figure 32: Prototype session process: Third person testing, UEQ and NASA TLX, Feature Discussion, Interview, Debug Page Testing, UEQ and NASA TLX and interview

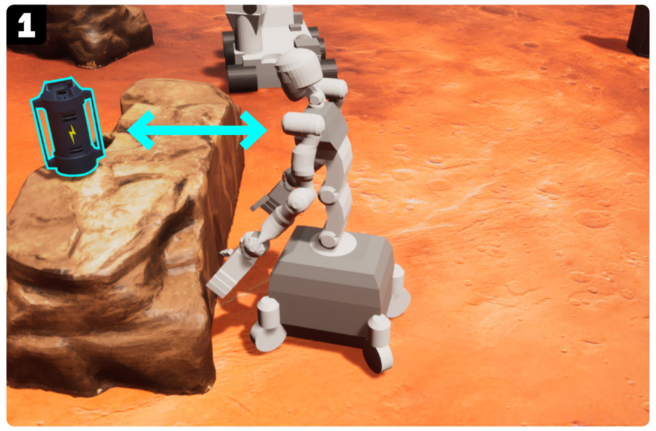
Figure 33: The three visuals highlight the potential error causes that participants encounter in the prototyping test: (1) Reachability, the robot has too much distance, as highlighted by the arrow and needs to approach the power cell from the other side. (2) Orientation, the robot's body parts (head in cyan and body in black) are misaligned, as highlighted. (3) Collision, the two highlighted objects block the path, so the robot stops to avoid the collision.

NASA TLX and UEQ

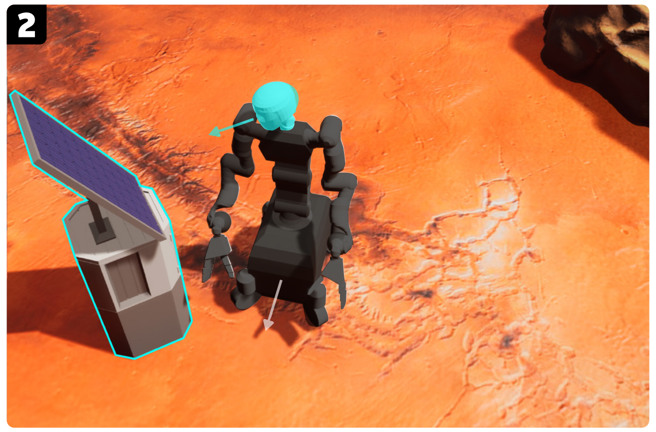
For each prototype and the comparison of the existing system, each participant was asked to fill out a NASA TLX and a UEQ immediately after completing their Tasks. The NASA TLX is a subjective tool used to assess mental workload during task performance [35], refer to Method Glossary p. 4 for a detailed explanation. The UEQ scales offer a holistic evaluation of user experience [102], for more explanation on this tool, refer to Method Glossary, p. 4. With these two Questionnaires the workload and the overall user experience can be measured and compared.

Third Person Prototype Testing

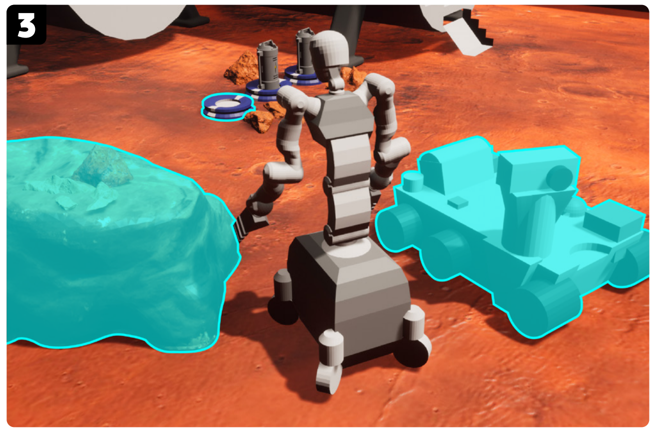
For interacting with the third person prototype, participants were asked to follow a protocol, provided to them through the chat of the interface. This is the same situation astronauts are put in when testing the current system. The protocol asked participants to (1) search a blue power cell, (2) install it in SPU5 and then (3) return to base. For every step of the protocol a potential error was forced onto the participants to see their reaction and evaluate their error handling capabilities.



(1) Finding the Powercell: When automatically navigating to the powercell, the robot would purposely navigate to a point not close enough, causing a reachability issue if not fixed by participants. To solve this issue participants had to go around and grab the powercell from another side, see Figure 33 (1).



(2) Installing it in SPU5: Once they obtained the powercell they continued to SPU5 to install it. SPU5 is a solar panel unit (SPU) that looks different to the once from DLR (like SPU3) but consist of the same components. Upon arrival to the SPU5 the robot was automatically forced into a bad orientation, which can happen since the head and base of the robot are independent. So, when users navigate manually and with autonomous commands interchangeably, this is not unlikely and was checked with the head of the Team [55] prior to testing. When trying to install the powercell, the robot threw again an error, see Figure 33 (2). This could be solved by manually bringing the robots head back into center. Once the users did that, installing of the powercell worked.



(3) Returning to Base: The last task asked the participants to navigate back to base. From their current position, next to SPU5, the shortest path to base included a potential collision cause. Therefore, if they chose to let the robot navigate, or manually took that path themselves, the robot would stop automatically close to the area of risk for collision and throw an error, see Figure 33, (3). Users could recover by either manually driving to base or positioning themselves somewhere else so that the new path was more collision free and from there let the robot navigate autonomously again.

Why this scenario: The scenario was chosen based on its similarity to other protocols, because the objects were logical for a mars mission but unknown to the users (similar to what astronaut's experience) and they covered three error types: reachability, orientation and collision.

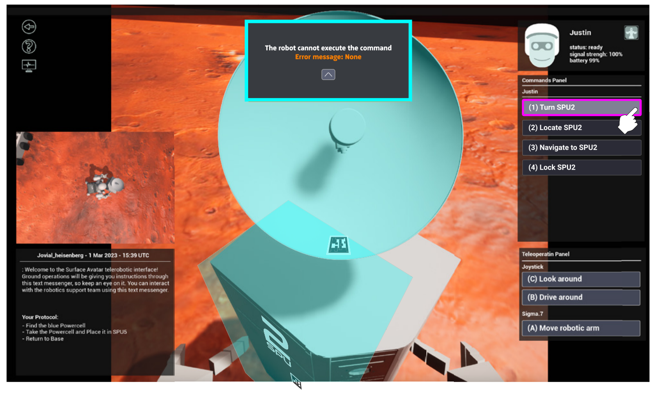
Limitation: The scenario does not cover localization errors since they could not be implemented to their full extend in the scope of this work.

Data capturing: While participants where interacting, the simulation screen was recorded, an additional back up camera was set up to film and audio was recorded on another device. The simulation what setup to have a tiny key log message every time users did certain interactions, so that one can derive every action from the screen recording. The additional camera was only aimed at the screen and keyboard and did not record the participants. It was there to capture any other usage of the participants that was not covered by the key logs. For example, if users mistakenly used wrong keys that had no action mapped to them etc. The audio was recorded to listen back to participants questions during the interaction, as well as when they were trying to explain their assumptions of error causes.

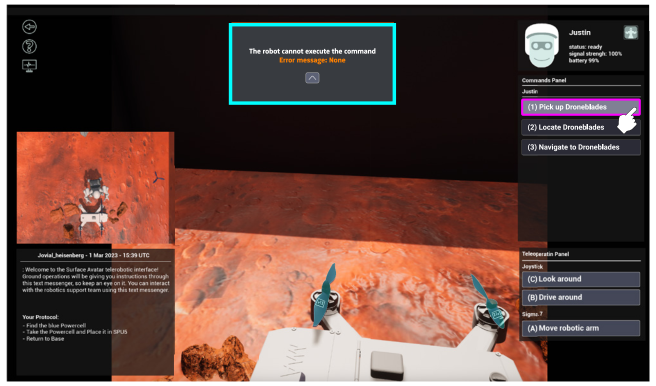
Game Element Evaluation:
Three visuals each were created to represent different game cues/elements related to seven topics: localization, collision, reachability, orientation, planning- and execution waiting time, and indicating object interactability in the virtual third person environment. Participants were shown these visuals for each topic and asked to select their preferred game cue. For localization and the three error types, participants were asked to choose a cue for both general information and error-specific information. They had the opportunity to explain their choice and provide feedback on the other propositions. Participants were asked to fill out a questionnaire to indicate their choice, see Appendix p. 172 - 175, but they could also further elaborate by explaining their preferences.
The game cues used in this study were derived from the matrix mentioned under Mitigating Errors: Third Person View, p. 88. By collecting participants' preferences and perspectives, this information will inform recommendations for improving the third person perspective and guide future research.

Interview Third Person
As the final stage of the third person prototype testing, participants were interviewed to gather additional insights. The interview questions covered several topics, including: (1) Positive or negative aspects that stood out during the interaction, (2) Assessment of the prototype's potential in mitigating errors, including specific examples, (3) Evaluation of cognitive workload and engagement during the interaction, (4) Identification of features that stood out or were lacking and (5) Suggestions for improvements.
These interview questions were designed to complement the results obtained from the NASA TLX and UEQ assessments, provide a deeper understanding of the prototypes' mitigating power, and give participants an opportunity to provide feedback and suggestions.

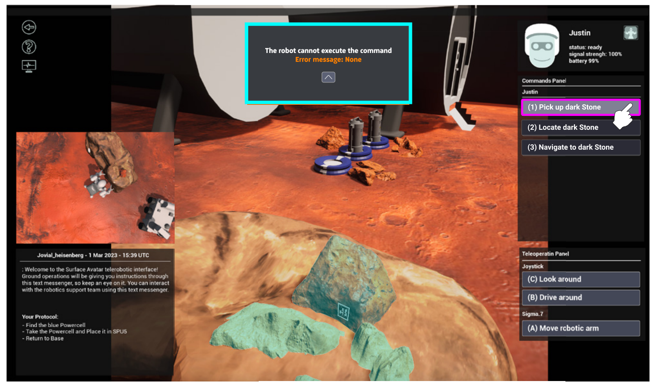
Debug Page Prototype Testing
Participants in the debug page testing were exposed to four different situations, each involving a distinct planning error. The current User Interface (UI) screen was displayed, highlighting the last attempted action, and presenting a planning error message. By clicking on the UI, participants could expand the error message to access information about typical errors and solutions, as well as utilize various UI commands, including teleoperation commands. Upon finding a solution and implementing it through a command, participants would proceed to the next situation. In total, participants encountered four scenarios, each featuring a different error type: localization, reachability, collision, and orientation. Please refer to Figure 34 for an illustration of the four scenarios.



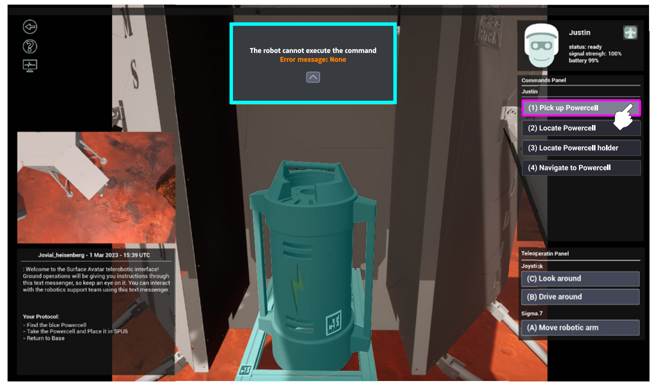
(1) Localize: The Object consists of multiple parts, the one on the bottom is not localized well, because the AprilTag is not fully in view. Moving the head to view the Tag better will help to localize it correctly.



(3) Reachability: The drone blade in the center of the view is too far away to be grabbed. Approaching the blade from the other side will solve the issue.



(2) Collision: The other stones in front of the one that the robot is trying to grab are causing a collision. Trying it again or simply approaching the desired Object from another position will help to grab it.



(4) Orientation: The robot's orientation is bad, which is not visible in the first-person mode. Bringing the robot back into an idle position will make it easier to execute any command (move head or body).

Figure 34: Overview of the four error scenarios from the Debug page prototype: localize, collision, reachability and orientation from top left to bottom right.

The objective of this prototype testing was twofold. Firstly, it aimed to assess participants' ease of recovery and ability to identify the cause of a planning error using the additional debug page information. Secondly, it is there to determine the number of attempts required to solve the problem.

Interviews Debug Prototype

Following the completion of the debug page prototype testing, participants underwent interviews to gather additional insights. The interview questions covered various topics, including:
(1) Notable positive or negative aspects experienced during the interaction, (2) The extent to which participants felt guided and supported by the provided information, (3) Evaluation of the informativeness and helpfulness of the debug pages, (4) Assessment of the prototype's potential in mitigating errors, with specific examples, (5) Evaluation of cognitive workload and engagement during the interaction, (6) Identification of standout features or areas lacking in the prototype and (7) Suggestions for improvements.
Similar to the interview questions for the third person prototype, these interview questions were designed to complement the results obtained from the NASA TLX and UEQ assessments. They aimed to provide a deeper understanding of the prototypes' effectiveness in error situations and offered participants an opportunity to provide feedback and suggestions.

Baseline: First Person Prototype Comparison

The comparative first-person prototype testing followed the same protocol as the third-person prototype testing, with identical errors and scenarios. The only difference was that participants experienced the prototype from a first-person perspective, where the map was displayed on the side instead of the camera view. The User Interface was rebuilt as it is in the current system and no additional features were integrated. This system allowed participants to navigate and interact with the environment as if they were experiencing the current surface avatar system. This setup was tested by different participants to the ones testing the prototypes.

By conducting the comparative testing in both third person and first-person perspectives, it was possible to assess and compare the effectiveness of the prototypes in error identification and problem-solving from different visual perspectives.

Comparative Study Questionnaire

The comparative questionnaire was conducted online and divided into two, labeled as Questionnaire A and Questionnaire B. Before the questionnaires began, participants were provided with an introduction explaining the purpose of the research and assuring them that it was acceptable if they had no prior experience with the system or did not know the answers. Participants were informed that their honest guess regarding error causes was what the research aimed to learn. A brief introduction to the system was given, including explanations of localization and the user interface (as astronauts would receive), while potential error causes were not explained (as astronauts would not receive such information).

Third Person vs. First Person
Both versions of the questionnaire followed a similar structure. Participants were first asked to provide demographic data to filter out participants. They were then presented with 12 error scenarios, shown as images, either in the third-person or first-person perspective. Each scenario included a description of the last action taken, and participants were asked to determine the cause of the error and select the appropriate action to solve it from multiple-choice options. An "Other" option was also available.

The 12 errors consisted of three reachability, three localization, three collision, and three orientation errors, presented in random order for each participant.

Attention Check
Following the 12 error scenarios, an attention check was included to assess participants' alertness and knowledge of the robot. Two questions were asked to confirm their understanding of teleoperation and localization.

Debugging pages vs. Current Error Message
Next, participants were shown the same four situations from the debug page testing conducted in the participant sessions. First, the situations are presented without any additional help. Subsequently, participants are shown either a video demonstrating typical errors and solutions or the motion planning of the robot for each situation. The same question as with the first 12 errors scenarios were asked: what the cause of the error is and what to do to solve the issue.

A/B Questionnaire Difference
Questionnaire A primarily focused on the third-person perspective. It included the first 12 questions in the third-person mode and followed by eight error situations in the first-person mode. Participants viewed the situations without help initially and then with either the motion planning or the typical errors and solutions. In contrast, Questionnaire B featured the same 12 errors in the first-person mode, followed by the same setup as Questionnaire A. However, where Questionnaire A had motion planning help, Questionnaire B provided typical errors and solutions help, and vice versa.

The questionnaire intentionally included multiple instances of the same error types across different situations to ensure that error identification was independent of the provided error scenario. By alternating between motion planning and typical errors and solutions, the study aimed to assess which form of assistance was more helpful. The initial presentation of the situations without help aimed to determine if the provided assistance influenced participants' perception of the error cause and if it facilitated error identification compared to the absence of help.

The comparative questionnaire design enabled an evaluation of participants' ability to identify errors given different perspectives and assistance methods.

Data Analysis

Statistical Analysis: The ratings obtained from the NASA TLX and UEQ assessments were subjected to a simple statistical analysis. This involved accumulating the ratings and comparing them between prototypes and the comparison test. By examining the statistical differences, insights were gained into the subjective mental workload and user experience in relation to the different prototypes.

Error Analysis: The audio and video footage captured during the sessions were utilized to derive participants' guesses regarding the cause of the errors. These guesses were transcribed and examined to determine common themes and patterns. Additionally, the number of attempts it took participants to solve each error type, as well as the time it took them was recorded, allowing for the calculation of an average number of attempts and time spent per error type. This analysis provided insights into the participants' problem-solving strategies and the challenges they encountered during error resolution.

Interview Analysis: The interviews conducted with participants were transcribed, and the responses were analyzed by clustering them into overarching themes using statement cards [14]. This qualitative analysis allowed for a comprehensive exploration of participants' perceptions, feedback, and suggestions. By identifying common themes, patterns, and trends, it provided a deeper understanding of participants' experiences and perspectives. Moreover, selected relevant quotes or excerpts from the interview data support the interpretations and findings. These quotes provide a human perspective on the statistical data.

Comparative Questionnaire Analysis: The comparative questionnaire data analysis focused on comparing the correct error identification rates and the correct actions to fix the issues between Questionnaire A and Questionnaire B. By examining the response data from both versions, one can assess the impact of different perspectives and assistance methods on participants' ability to identify errors and determine appropriate actions.

Prototype Results and Eval-uation

The following section presents the results of the prototype testings. These results provide insights into the effectiveness and usability of the developed prototypes in addressing the identified core issues related to error handling.

UEQ: User Experience Results

The results of the UEQ are displayed in Figure 35. The UEQ scores were analyzed for the first-person perspective comparison (baseline), the third person perspective prototype and typical errors and solutions prototype. The grey beams in the graphs indicate the mean value per category, while the dark lines show the variance. It should be noted that all participants encountered error scenarios and had to manage errors in each prototype and the comparative study.
Both the baseline and the third person perspective had the User Interface as it is in the current system and no additional features were integrated for the prototype testing. Features for the third person perspective were evaluated separately, see Feature Discussion, p. 120.

Overall: A notable finding is the overall difference in scores observed in the first-person perspective (baseline) compared to the prototypes. In the comparison, all five categories of the UEQ fall within the neutral range, whereas both the Debug page prototype and the third person perspective prototype show all scores within the positive range.

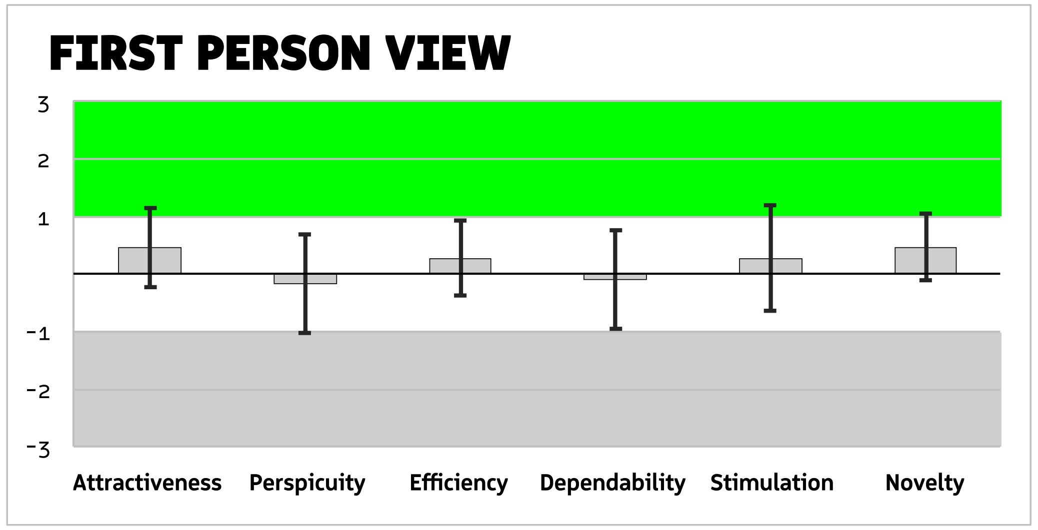
Debug Page: For the Debug page prototype, the UEQ results were consistently high across all categories (Attractiveness 2,104, Perspicuity 2,281, Efficiency 2,125, Dependability 2,125, Stimulation 2,094, Novelty 1,531). The UEQ provides a benchmark comparison based on data from 21,175 individuals in 468 studies involving various commercial products [102]. As Rollin Justin is not a commercial product, direct comparisons must be made with caution due to differing standards and expectations. Nevertheless, the "Excellent" scores in five out of six categories, with a "Good" rating for novelty, suggest a satisfactory user experience. It should be noted that the error situations for the Debug prototype were similar but not exactly the same as those in the comparison study.

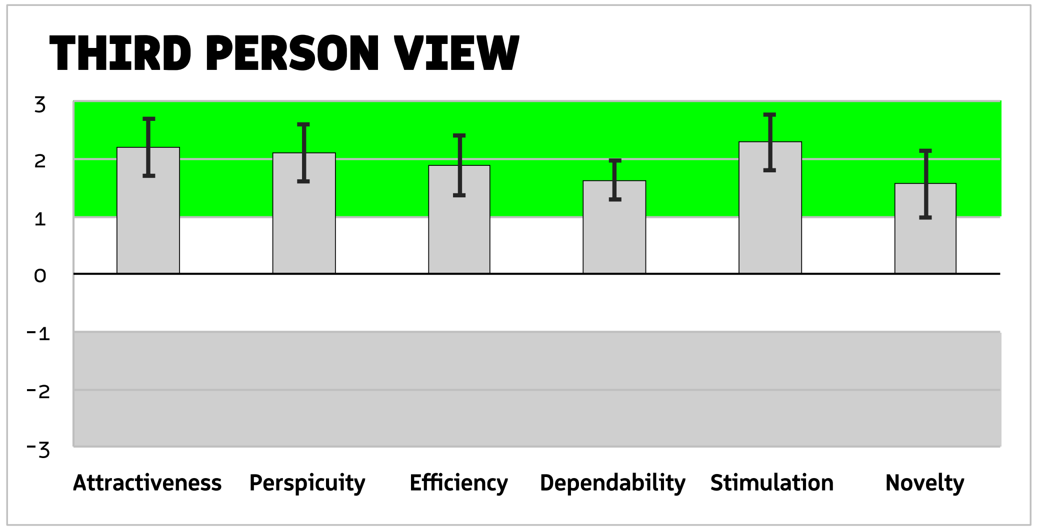
Third person: The third person perspective prototype also demonstrated positive results, although slightly lower in the "Dependability" and "Efficiency" categories. This outcome is not unexpected, as participants received no error-solving help in the third person prototype, unlike the Debug page prototype, where concrete potential causes and solutions were provided.

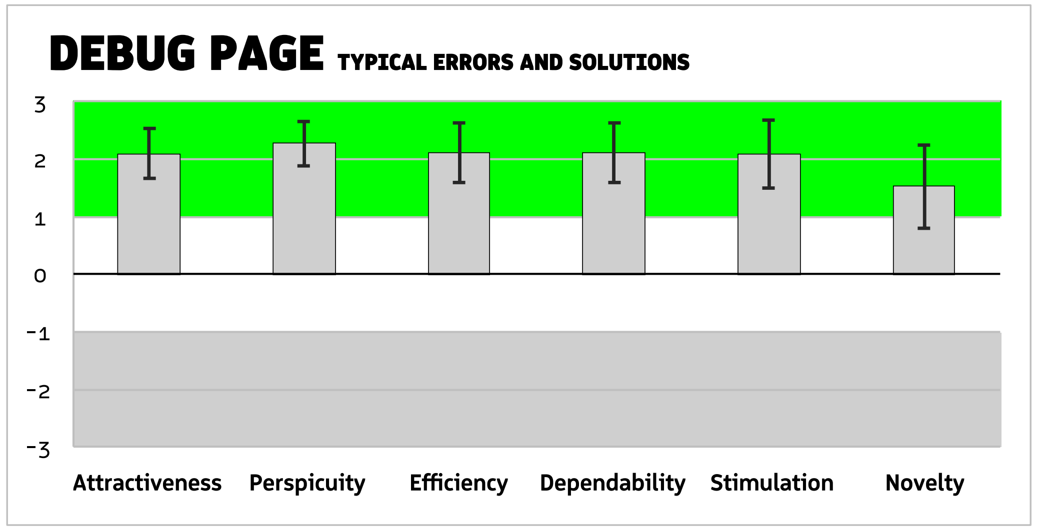
Comparing each scale of the third person UEQ results to the ones of the baseline, all categories show a marked improvement (Attractiveness 2,188|0,458, Perspicuity 2,094|-0,156, Efficiency 1,875|0,281, Dependability 1,625|-0,094, Stimulation 2,281|0,281, Novelty 1,563|0,469). Furthermore, when considering the Benchmark ratings, there is a substantial enhancement (Attractiveness "Excellent"|"Bad", Perspicuity "Excellent"|"Bad", Efficiency "Good"|"Bad", Dependability "Good"|"Bad", Stimulation "Excellent"|"Bad", Novelty "Good"|"Below Average").

Overall, the results indicate a notable improvement in the user experience for both prototypes, as reflected by considerably higher UEQ scores, compared to the first-person perspective (baseline).

Figure 35: UEQ results for each prototype and the comparison, from left to right: the first-person perspective comparison (baseline), the third person perspective prototype and typical errors and solutions prototype







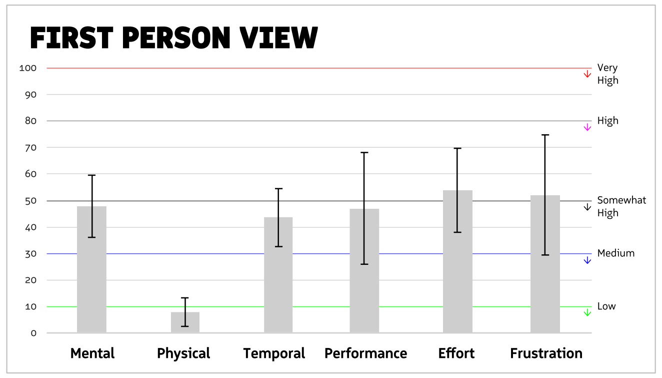
NASA TLX: Workload Results

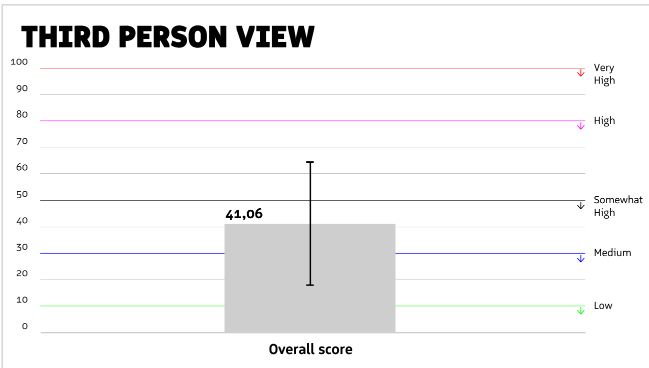
The NASA TLX scores are presented in Figure 36, displaying the results for the Debug page with the first-person perspective comparison (baseline), the third person perspective prototype and typical errors and solutions. The top row of the graph shows the scores split by category, while the bottom row represents the overall score.
The original scale of the NASA TLX, as well as its subsequent research, have not yielded normalized data that allows researchers to determine what constitutes an acceptable or unacceptable workload. However, from the original validation studies [35], a limited set of benchmarks can be derived, which are indicated on the left side of the graphs (e.g., low, medium, somewhat high, high, very high) to provide some reference points.

Overall: Considering the participants encountered multiple error scenarios the workload scores appear relatively good based on the reference points. Both the Debug page and the third person perspective prototypes outperform the first-person comparison. The overall NASA TLX scores for the Debug page and the third person perspective prototypes fall within the "Medium" category, while the first-person comparison scores are categorized as "Somewhat High." This trend is consistent when examining the individual scores, with the first-person view consistently performing worse in every category compared to the other two prototypes. It is important to note that there is a relatively high standard deviation in the overall scores for all the results. However, even accounting for this variability, the prototypes still demonstrate better performance. The physical demand score for all prototypes and the comparison is rated as "Low," likely due to the participants interacting with a computer screen while seated throughout the entire testing session.

Debug page: The Debug page prototype consistently demonstrates better results with lower scores overall. Notably, in the 'Frustration' category, the Debug page prototype shows a marked improvement compared to the first-person comparison (19.28 - 'Medium' | 51.73 - 'High') and performs better than the third-person prototype (19.28 - 'Medium' | 29.38 - 'Medium'). Another noteworthy finding is the score for 'Effort,' which exhibits better values compared to the first-person view (26.88 - 'Medium' | 53.46 - 'High') and similar values to the third-person view (26.88 - 'Medium' | 28.13 - 'Medium').

Third person: The results of the third person perspective prototype are consistently better than the first-person view in every category, with the most notable differences observed in "Frustration" (29.38 - "Medium" vs. 51.73 - "High") and "Effort" (28.13 - "Medium" vs. 53.46 - "High"). The improved cognitive workload scores are further supported by participants' feedback. Seven out of eight participants mentioned that the third person perspective reduced their cognitive workload, emphasizing the benefits of having a better overview of the environment. Some participants (P01, P02, P05, P06) specifically highlighted the improved situational awareness and reduced motion sickness associated with the third person perspective:





"So, I think that because you saw more, it was easier than if I had only seen the camera, because I have a better awareness of where things are around me, what the robot is like, where can I navigate to?" – P06 (translated from German)

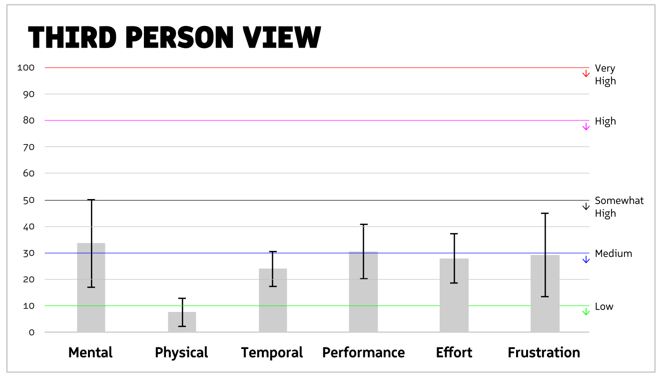
"Yes much, much less. So, you could do a lot more with this world... This situational awareness helps so much, I think it's just good. I'm just wondering a bit what you need the 1st person for. Probably for direct teleoperation." – P01 (translated from German)

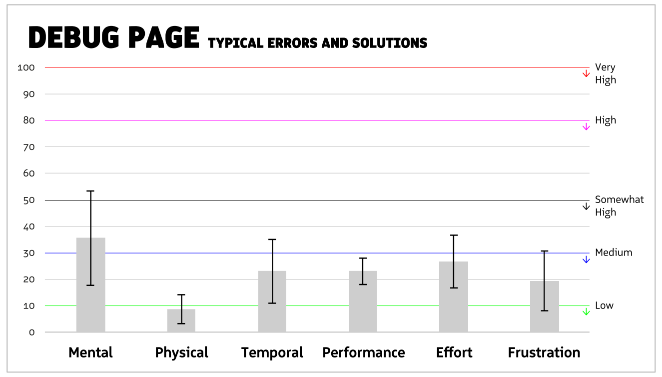
"I would say it will definitely be a bit less workload. Simply because you have a better overview." – P02 (translated from German)

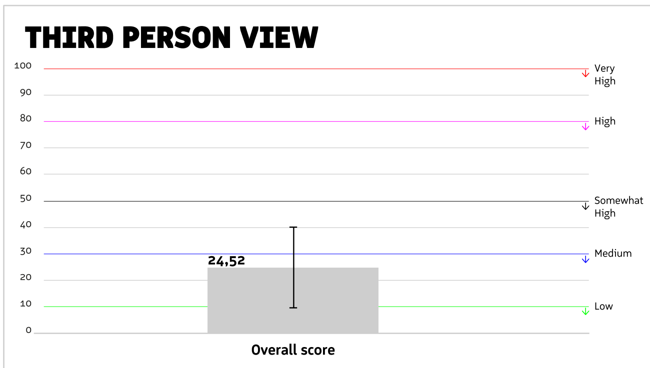
"In principle, I find it increasingly difficult with a first-person perspective. [...] I actually get sick quickly when it's in the first-person perspective." – P05 (translated from German)

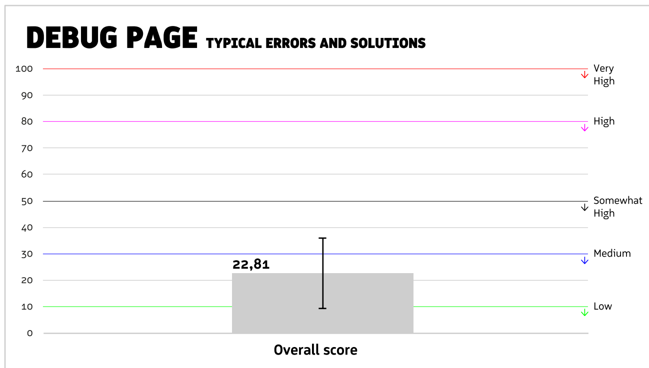
In conclusion, the results indicate an overall improvement in cognitive workload for both the Debug page and third person perspective prototypes. Participants' feedback regarding the situational awareness provided by the third person perspective further strengthens the understanding of the observed improvements.

Figure 36: NASA TLX scores for each prototype and the comparison, from left to right: the first-person perspective comparison (baseline), the third person perspective prototype and typical errors and solutions prototype. The top row shows the results split by category and the bottom row shows the overall cognitive workload.









Attempt Rates and Time Spent

This section presents the average number of attempts and time spent (in seconds) on each error encountered by the participants. Figure 37 displays the results for the first-person comparison prototype (baseline), the third person perspective and the typical errors and solutions pages. The top row of the graphs represents the time spent per error, while the bottom row shows the average number of attempts per error. The time spent was measured from the first attempt to the final completion. If a participant asked for help, the time was measured until their last attempt, since in the prototype sessions all participants solved the problem latest after asking for help. It is important to note that the errors presented in the Debug page prototype were different from those in the third and first-person perspectives.

Overall: The difference between the third person and first-person perspectives may not appear substantial at first glance. However, when considering factors such as the time spent on each error and the average attempt rate, it becomes evident that the Debug page prototype exhibits notably lower values. A detailed comparison of the results for the Debug page, first-person view without help, and motion planning prototypes will be discussed under Questionnaire Results, p. 126.

Debug page: Overall, resolving errors with the debug page appears to be much faster and with a markedly higher success rate compared to the third- or first-person view. This outcome aligns with the intended purpose of the prototype, which focuses on providing immediate error assistance, whereas the third person perspective concentrates on situational awareness and mitigating errors. Participants were able to identify and resolve Reachability, Collision, and Orientation errors with their initial guesses, while Localization errors required an average of 1.25 attempts to solve. No external help was requested by the participants for any of the error scenarios. On average, participants spent approximately one minute per error, with Reachability being the fastest to resolve at 27.125 seconds. Although the time spent may seem slightly high, it could be attributed to the amount of text provided in the Debug page prototype. Three out of the eight participants (P03, P08, P01) expressed a desire for less text to read, while another participant (P02) found the amount of text to be satisfactory, see participant quotes:

"The text might be too long" – P03

"It's a bit on the long side sometimes. Yeah. Um, but [...] I've seen how, how the crew do this, they, they do read." – P08

"It maybe is too much text. I think I would like less text." – P01 (translated from German)

"It does everything it should and is not overloaded." – P02 (translated from German)

Third person: In general, the desired outcome is for participants to efficiently resolve an issue. However, in cases where errors are not easily identifiable, it is beneficial for participants to invest more time and make multiple attempts before seeking assistance. Therefore, for the reachability and collision errors, better results are indicated by a smaller number of attempts and less time taken to resolve the issues, as all participants were able to successfully solve them. Conversely, for the orientation error where not all participants were able to resolve it, a higher number of attempts and more time invested can be seen as positive indicators. A higher number of attempts and more time invested can be seen as positive indicators as it demonstrates greater engagement and a stronger willingness to persist and figure out the issue independently.

When comparing the third person perspective to the first-person view, the average attempts and time spent are not notably different. No external help was needed for the collision or reachability error and the numbers for first person and third person are roughly the same. For the collision error, participants from the third person required less time and attempts until completion than the ones

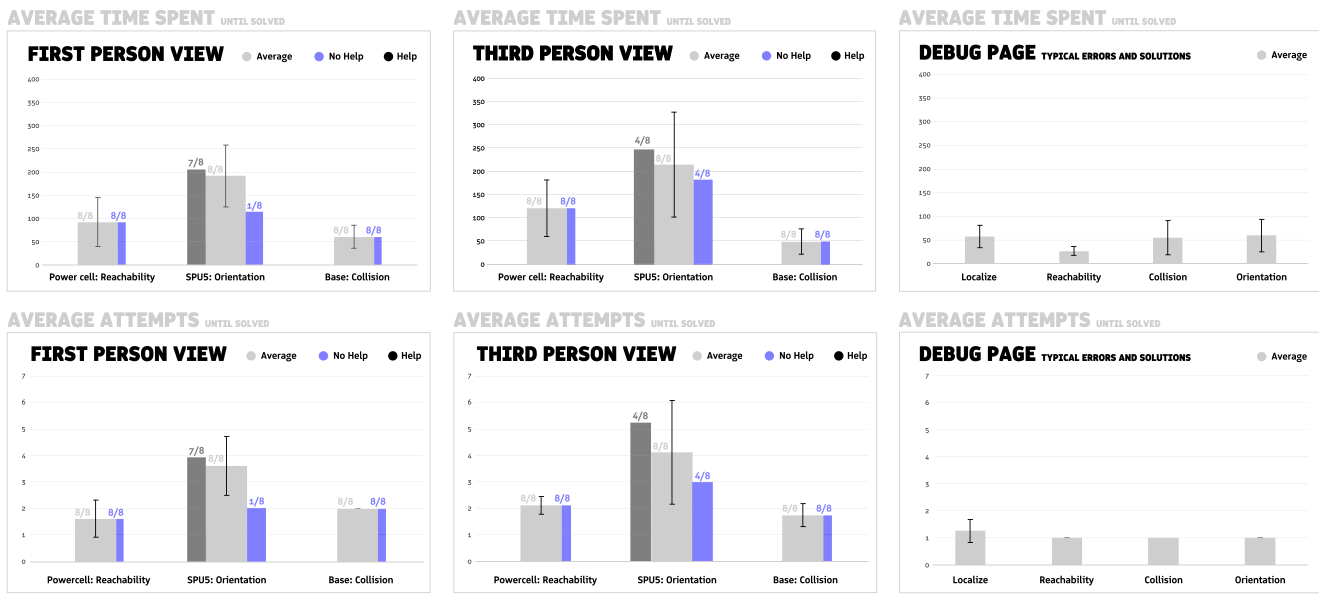


Figure 37: Overview of average attempts and time spent per error for each prototype and the comparison, from left to right: the first-person perspective comparison (baseline), the third person perspective prototype and typical errors and solutions prototype. The top row shows the time spent until completion and the bottom row the overall attempts at solving.

from the first person (Third person| First person: 1,75|2 and 48,25s|60,625s). In comparison, for reachability, the first person showed lower numbers (Third person | First person: 2,125|1,625 and 120,525s|92,125s).

For the Orientation error, four of the participants from the third person prototype needed help, while seven of the first person required it. Moreover, the participants form the third person perspective spent more time and attempts at resolving the error (Third person | First person: 4,125|3,625 and 214,75s|191,75s). Especially the participants that in the end required help spent more attempts in the third person perspective than the ones in the first-person view (Third person | First person: 5,25|3,86). Overall, more time and attempts were spent in the third person view. This can be explained by the fact that the participants from the first-person comparison were quicker at asking for help while the third person participants wanted to keep trying. For instance, participant P04 and P07 were reminded after a couple of failed attempts that they can ask for help, but insisted to continue:

"I would love to try more" – P04

"Okay, just try one more time" – P07

This was not exhibited by the participants from the first-person perspective test.
Moreover, the willingness to continue trying in the third person perspective could further be attributed to the lower level of frustration (as indicated by the NASA TLX score) and the high engagement reported by seven out of eight participants when asked about it. For example, participant mentioned the interaction to be fun (P05, P07), engaging (P03), intuitive (P02) and wanted to continue solving task problems (P01), see below:

"Totally. I would love to continue and solve some more tasks"– P01 (translated from German)

"Switching the view, I think it's already, um, yeah, it's already like engaging" – P03

"It, it was actually fun. It was engaging, yes." – P07

"Yes, it is. It's quite fun to drive around there." – P05 (translated from German)

"It was quite good that way. So especially if you gamed a bit, it's just very intuitive." – P02 (translated from German)

In terms of error resolution, the Debug page prototype demonstrated the most favorable results, with notably better performance compared to the third person and first-person perspectives. While the third person perspective required more time and attempts to resolve errors, it also relied less on external assistance and exhibited a higher level of participant engagement compared to the first-person perspective.

Feature Discussion

As mentioned under Prototype Testing Sessions, p. 102, participants were asked to provide feedback on further proposed features for the third person perspective. The preference results of the suggested features for each discussed contextual scenarios can be seen in the Appendix, p. 175.

The participants' preferences regarding specific features were strong, and a summary of the top-ranked features can be observed in Figure 38. Seven of the eight participants wanted to have at least some of the provided features implemented. Some specifically mentioned features included the progress bar (P07), indicating visible object with augmented information (P6) or showing the reachability of the robot (P01). Other participants simply wanted all proposed features (P03, P08), see below for quotes of the mentioned examples:

"I think the progress part with steps and execution, that is one thing that was great. " – P07

"All of them. All of them." – P03

"I think this is already, uh, adding a lot of good, uh, features that I would like to see there." – P08

"Especially the perspective of what is visible what is not: so augmented information I found pretty pretty cool." – P06 (translated from German)

"(On Reachability map) I think that's good. We can also project that well. We have the skills to do it and I can do it quite well in theory." – P01 (translated from German)

Overall, the combination of the presented features showcased in Figure 38, offer valuable insights on how to effectively enhance the third person perspective.

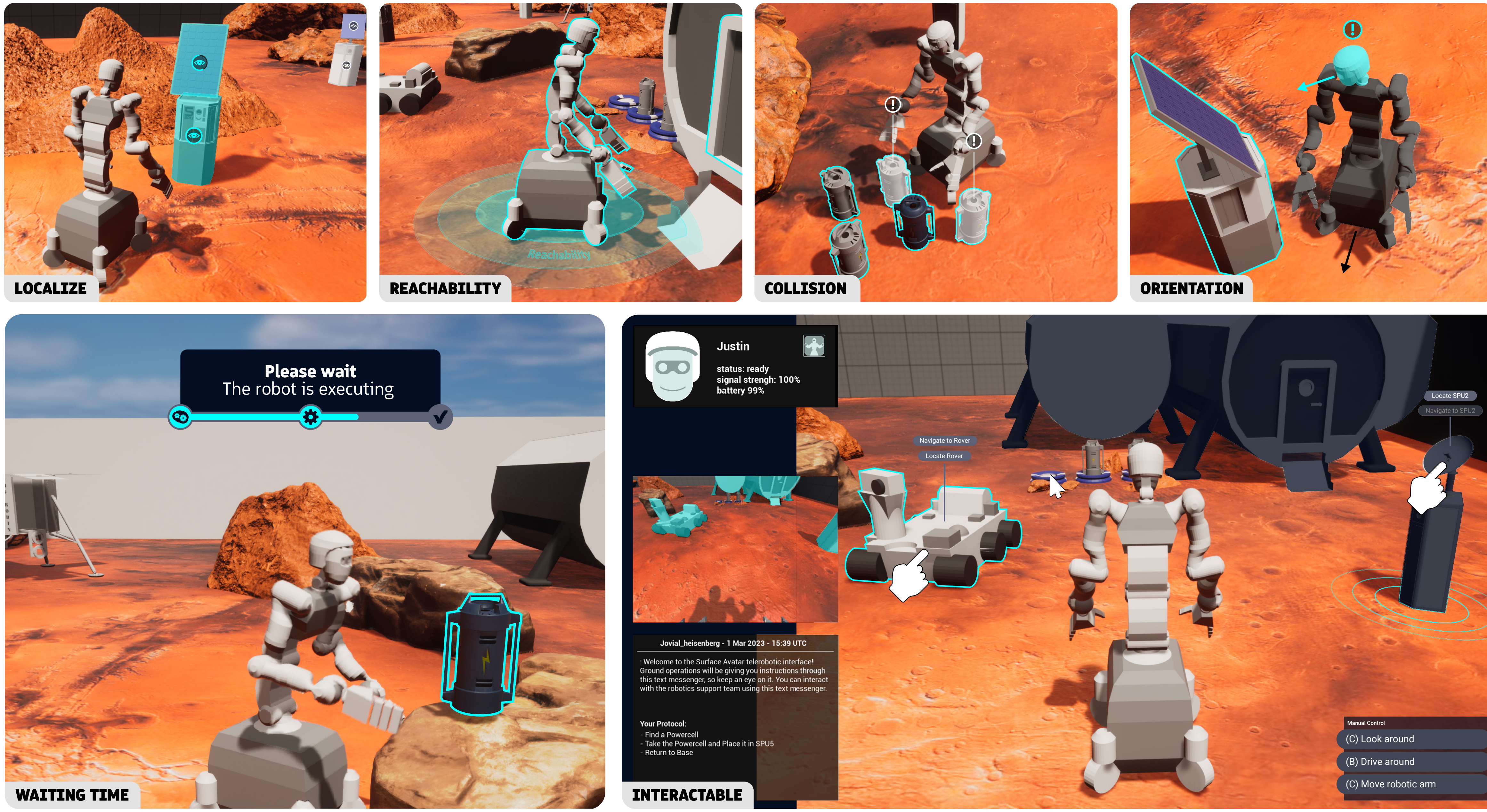


Figure 38: Overview of preferred features: Localize (showing the percentage to indicate how well an object is highlighted), Reachability (projecting the robot's reach), Collision (highlighting objects that cause collision), Orientation (if body is misaligned, highlighting the different body parts to draw attention to it), Waiting Time (showing a progress par that goes from planning time to execution time to being done) and Interactable (moving interaction onto objects instead of having them on the side, if an object is not in vie clicking on it will deactivate actions and tell user to bring object into view)

Interview Results

This section presents the results of the interviews, focusing on various topics discussed during the interviews. These topics include the participants' perceptions of both positive and negative aspects of the prototype interaction, the effectiveness of the prototypes in mitigating challenges, the level of guidance provided, and other relevant insights.

The Debug Page Prototype
Debug page design: Within the discussed aspects of the debug page prototype four main themes regarding the content were identified: providing solutions, providing clear explanations, the text length, and the use of visuals in the debug page. Table 1 provides a comprehensive overview of the core themes identified during the interviews, along with quotes from participants. It is important to note that not all quotes are included in the examples, as some participants expressed agreement without providing additional details. Furthermore, certain themes are covered by a single quote, which would result in repetition if included as separate examples.

Debug Page Design
Among the design themes, the most prominent one, was mentioned by seven out of eight participants, focused on the helpfulness of the provided solution strategies. Participants see Table 1 for examples of participants reasoning.

Another strong aspect mentioned by six of the eight participants focused more on the general theme of having instructions making it easier to identify the error cause, even as a novice user, see Table 1 for examples.

The last identified themes focus more on the design, layout and visuals given in the debug page. Four of the participants mentioned the visuals and layout. The main aspects cover that the abstract visuals make it easy to use the system robot independent and that having the overall design helped in figuring out the issue.

As mentioned under Attempt Rates and Time Spent, p. 118, the amount of text had more diverging opinions. However, one of the participants (P08, with the highest experience with astronauts) mentioned that it was not necessarily bad to have more text considering that astronauts tend to read a lot:

"For me it's a system that, that, that aims at people who like to read manuals, <laugh>, and that's perfect for a space [...] deployment. [...] They will read, they will read it." – P08

Debug Page Helpfulness and Impact on the User:
The strongest cluster among participant covers the aspect of creating a general better understanding about the robot, see Table 1. All participants believe to be better equipped for future scenarios and errors. The last mention characteristic is the level of guidance and support participants felt when using the Debug page prototype. Six of the eight participants felt that it helped and guided them to solve the errors scenarios.

Providing solution strategies (7/8 mentions)	Instruction design (6/8 mentions)	Visuals and layout (4/8 mentions)
"In the end you have just some, [...], one or two very specific instruction. I really like that." – P02 "It was cool that there were explanations right away about what it could be and then also suggestions for solutions. I think that's pretty cool" – P05 (translated from German) "Um, I, I like the, uh, suggestions. That was good that I had suggestions Okay. On how to move forward." – P08 "positively stood out to me. Was the hints you'd get from the pop-up thingy. I think that's very cool." – P04 "I liked that it was nicely presented with how many possible solutions there are and that they were kept abstract." – P01 (translated from German)	"I, I like really like this [...] you can get more instructions [...] you can click and get more information and even more concrete instructions." – P02 "The error messages were on point, like the instructions mm-hmm. <affirmative> they were good." – P07 "The Hinweis section where you have so much, uh, details on, on what could be wrong. [...] Very useful, and I would count on it." – P08 "Yeah, I mean, the information that was there was pretty cool. [...] what I found very good, is, that you show the possibilities also, so everything that can be the cause, and then I could solve that already directly." – P06 (translated from German) "I really like it because it just tries to explain and to a normal user that never used Justin, what could possibly go wrong?" – P04	"So these visualizations here were very expressive. So you actually understand directly what is meant. That's why I thought it was very, very cool." – P06 (translated from German) "It's fun. [...], I mean this visualization, I like it" – P03 "I thought it was good that it was nicely presented, [...] and that they were kept abstract. So it's robot-independent, because if you assume that it's a different system, it also works, so it was done just as well." – P01 (translated from German) "The layout of the extended Error message, that's great. So just with the nice pictures and underneath the text plus clickable links to the commands, that's actually really handy. Yes, I would like to have the system." – P02 (translated from German)

Future helpfulness (8/8 mentions)	Guidance (6/7)
"Yes. I think just sticks in your mind and you think about it more so" – P04 "I think by having seen a few of them, um, I will start getting a better idea of this is how this environment works." – P08 "So I personally would read through it and then practically click one of the suggested things. And when I read through it, I think it also helps me in the future to have a better understanding of what possible sources of error are." – P06 (translated from German) "Yeah. It definitely helps in tackling future problems." – P07	"Yes, very intuitive [...] I think in general the instruction are quite, yeah. Supportive." – P03 "I felt confident."– P04 "They were very helpful." – P01 (translated from German) "Yes very helpful. As I said, because on the one hand they are very understandable, they give me a little better understanding of what possible problems could be. I find that helps then very well." – P06 (translated from German) "Very helpful. It gives you context and you can build on that." – P02 (translated from German)

Table 1: Debug Page Prototype discussion clusters, from left to right: Providing Solution strategies (7/8 mentions), Instruction design (6/7 mentions), Visuals and layout (4/8 mentions), Future helpfulness (8/8 mentions) and Guidance (6/7)

The Third Person Prototype
For the third person view, the core identified clusters were the improved overview, a discussion on navigation controls, that the orientation error maintains a difficulty and that all participants believe the third person perspective to mitigate errors, See Table 2 for the overview of all supportive statements for these themes.

Generally, six out eight participants mentioned that the improved overview of the environment and robot stood out to them, because they were able to see more objects, and made the robot usage more intuitive. This also aligns with the reasons the participants gave, why they believe that the third person perspective helps in mitigating errors. As provided in Table 2, all participants believe the perspective to reduce error scenarios, because they are more immersed, have a better situational awareness and a better overview of their surroundings. However, five out of the eight participants expressed that orientation errors remain a challenge. They highlighted the difficulty in recognizing that robot orientation can be a contributing factor to the error, as it is not immediately obvious unless explicitly pointed out. Another repeated point of discussion revolves around the manual control of the robot. In the prototype, participants were instructed to use the WASD keys on the laptop as a substitute for the joystick, which was unavailable during the study. Three participants identified this as an improvement, considering it to make the control more intuitive. However, two other participants found it confusing, with one participant (P08) expressing particular difficulty with the WASD controls. It is important to note that this aspect was not the primary focus of the study, as extensive research has already been conducted on joystick controls to cater to astronaut preferences. Further testing is required to explore this aspect in more depth.

Overall, the themes identified for the Debug Page Prototype focused on providing solutions, clear explanations, text length, and the use of visuals. Participants found the provided solution strategies helpful, and instructions aided in identifying error causes. Visuals and layout were appreciated, but opinions on text length varied. The Debug Page Prototype was perceived as beneficial in enhancing understanding, equipping participants for future scenarios, and providing guidance for error resolution.
For the Third Person Prototype, participants highlighted the improved overview and intuitive robot usage. The perspective was believed to mitigate errors by enhancing situational awareness. However, participants expressed challenges with orientation errors and mixed opinions on manual control using the WASD keys.

Improved Overview (6/8)	**WASD Navigation (3/8 like, 2/8 Dislike)**
"Like the switching view? I like it." – P03 "[...] because it's very important if you want to get a closer look or a different view of things and even meanwhile, just have your robot and you can see at which orientation position it is. I really like that"– P04 "That the map is rotatable. Is great because it is not in the current version and that is super annoying some times." – P02 (translated from German) "The scroll using the mouse part, like the third person view where you can navigate the entire environment using the mouse. Mm-hmm. <affirmative>, that was actually very intuitive." – P07 "I think it's pretty cool. For one thing, because you have a practical, better impression about, about the object you're dealing with, the robot you have a better understanding of" – P06 (translated from German) "Honestly, if a game offers a first person perspective, but also third person, that I'll switch that right over to 3rd person. Skyrim for example." – P01 (translated from German)	"Uh, yes, I like it (WASD), especially with a keyboard. Okay. It's more intuitive I would say."– P04 "actually, quite practical (WASD). Because it's intuitive, because everyone has gamed something now and then." – P02 (translated from German) "So, navigation was amazingly intuitive even with manual, so just like you're used to." – P01 (translated from German) ——— "One more thing. It's not moving forward from this. It's a bit ma weirdly mapped" – P07 "Negative part was really the controls (WASD). [...] that's really the only negative, otherwise it's okay" – P08 "The missing joy stick brought, um, a lot of frustration and uncertainty in the, uh, in, in, in the command" – P08

Orientation still hard (5/8)	**Mitigating errors due to situational awareness (8/8)**
"You need to align your camera view when moving the platform, that is difficult "– P03 I just didn't like that the prototype didn't tell me I need a camera for the action."– P04 "Without error messages, finding out that the alignment of the head is important is difficult." – P02 (translated from German) "The negative part was, during the SPU interaction" – P07 "[...] And that's why this collision was not obvious to me and that was particularly negative" – P01 (translated from German)	"Yeah, I think it helps cause you can have a better, uh, view of the surrounding environment and it helps entire operation mode a lot, I think." – P03 "Yes, because you get more overview yourself and the whole map is displayed."– P05 (translated from German) "So for my example, I was looking at where should it be so I can face it correctly. So, I think it supposed to be helpful. [...] it guided me to debug the action."– P04 "I think so, because you also have a better overview of the whole. [...] with a direct camera view it's also awkward to look there more often and then you don't have something else in the picture. From that point of view, it's helpful." – P02 (translated from German) "Yes. God views is good. I like that. I wish I had that." – P08 "I would, say yes, it is helpful because I think during the SPU as well, my first guess was, it's not in the direct orientation, so I could actually use a third person perspective to look around the object, which actually helped me orient itself." – P07 "I think for sure, [...] this immersive experience, it's much more present. [...] if you see the robot all the time, then you know exactly how it looks and if something looks unhealthy. Then it's directly present to you. And that's why I think that's super valuable. [...] So I thought, with this inserting into SPU, I thought it was cool that you could look around like this, is there any space? And the perspective definitely helped." – P06 (translated from German) "Yes. Yes, definitely. For example, I saw this mistake with the "local minima", and the next moment I already knew. In addition, I would never have seen, if I were now in first person, that behind me there was the wall, the stones, where I had to navigate around, because I saw how the robot reacts, I then went directly this half step for me and the rest I still let the robot do. And that was already. It was already very good."– P01 (translated from German)

Table 1: Third person prototype discussion clusters, from left to right: Improved Overview (6/8), WASD Navigation (3/8 like, 2/8 Dislike), Orientation still hard (5/8) and Mitigating errors due to situational awareness (8/8)

Questionnaire Results

The questionnaire of the A/B between Subjects study collected answers from a total of 42 participants (Group A: 21, Group B: 21). The collected data includes correct and incorrect error identification results, as well as participants suggested next actions.

Error Identification

The correct error identification results are presented in two Figures. Figure 39 displays the correct error identification results categorized by error type and scenario. Figure 40 shows the difference in correct error identification among the three conditions: (1) No help (current system), (2) Typical errors and solutions debug page, and (3) Motion Planning of robot in Debug page, for four tested error types.

Third person vs first person: Looking at Figure 39, a notable finding is that the third person scenarios consistently show a higher rate of correctly identified errors across all 12 scenarios. While the difference is relatively small for localization, correct error identification for collision and orientation, is notably higher for the third person. For reachability, although still better in every scenario, certain situations appear to be almost equally well identifiable with the first-person view. Orientation stands out as the most challenging error cause to identify, which aligns with the findings from the participant study. On the other hand, localization seems to be the easiest error cause to identify.

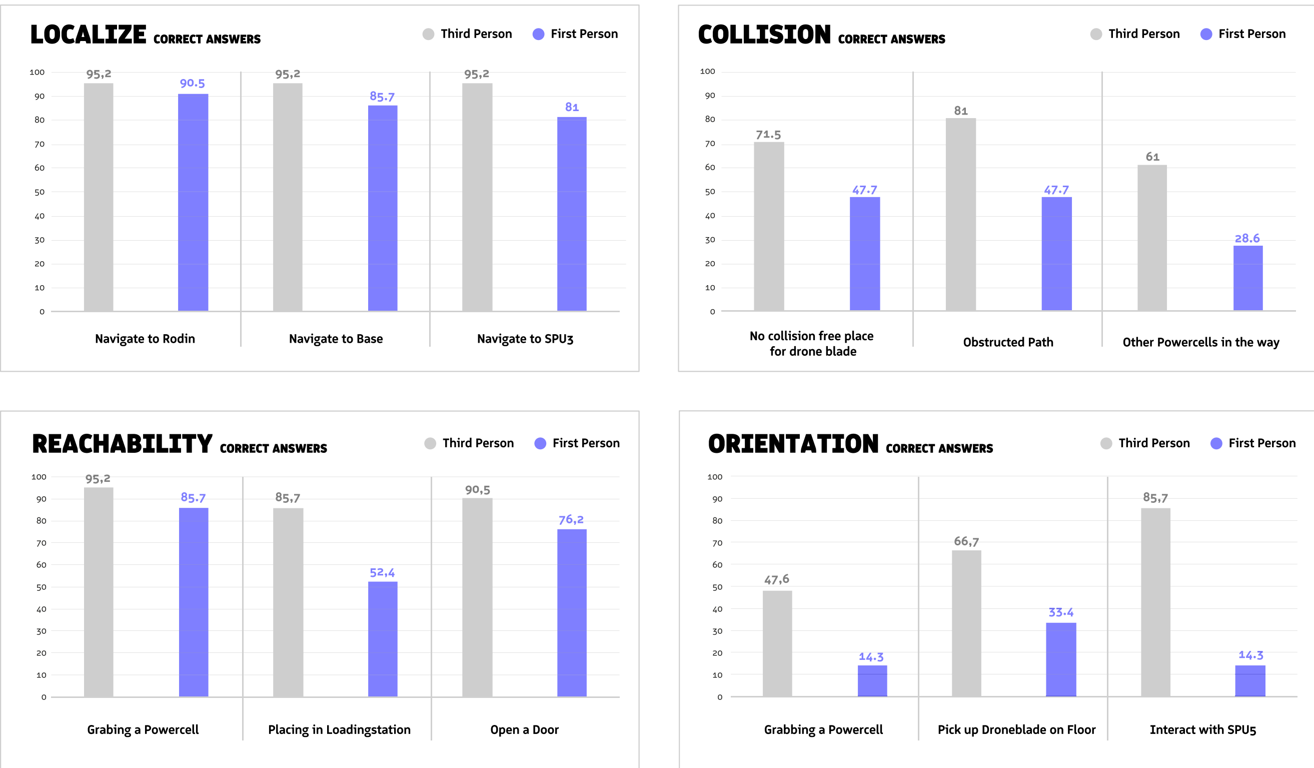


Figure 39: Questionnaire results third person vs first person, from top left to bottom right: 3 Localize scenarios, 3 Collision scenarios, 3 Reachability Scenarios and 3 orientation scenarios.

Debug Page Comparison: Before referring to Figure 40, it is important to note that the "No Help" scenario results represent the mean of all 42 responses as all participants from both questionnaires received these scenarios. The dark lines in the graphs indicate the standard deviation between Questionnaire A and B. Observing the "No Help" statistics for correctly identified errors, it becomes apparent that the chosen scenarios were more challenging compared to the average scenarios from the third person vs. first-person perspectives. Nevertheless, the results strongly support the effectiveness of the typical errors and solutions prototype, particularly when compared to the current system with no help. The typical errors and solution debug page consistently performs the best across all four error types, achieving almost perfect scores. Introducing motion planning of the robot shows mixed results, with a decline in results for localization but slight improvements for collision and orientation. In the case of reachability, motion planning performs at a comparable level to the typical errors and solutions page, which aligns with its inherent visual indication. While the motion planning results for orientation are still relatively low, they considerably enhance the correct error identification by almost threefold.

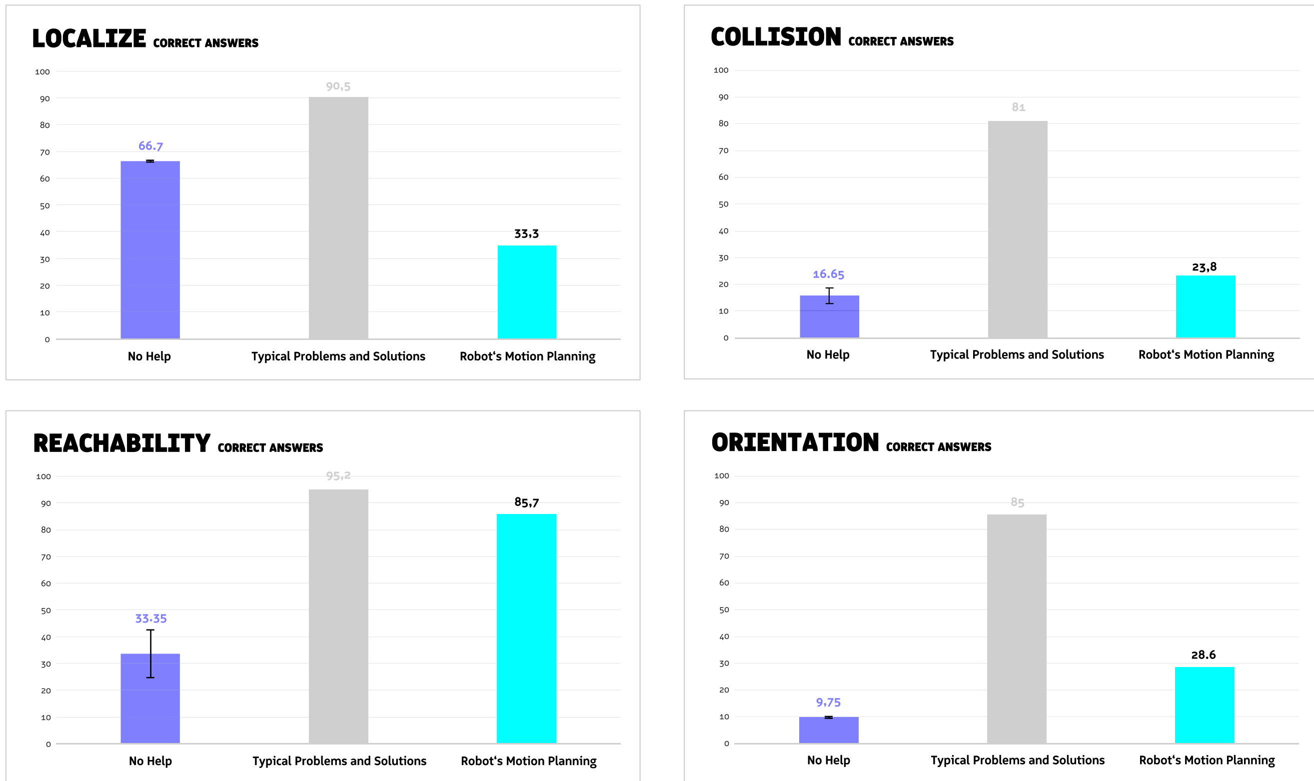


Figure 40: Questionnaire results debug pages. Blue bars indicate correct error identification without any additional debug page or feature (baseline), grey indicate with the help of the typical error and solution debug page and dark grey indicate with the help of the robot's motion planning debug page. From top left to bottom right, the results are shown for the scenarios: Localize, Collision, Reachability and Orientation.

Suggested next Actions
Participants were additionally asked to select actions to address the identified errors, and their responses were categorized as "Ok," "Bad," or "No Effect". Actions classified as "Bad" included asking for help or teleoperating the arm, both of which were time-consuming and unnecessary for the given scenarios. Responses that did not include bad actions but also did not contribute to problem resolution were categorized as "No Effect," while answers with actions that could potentially solve the error were classified as "Ok." It should be noted that "Ok" responses do not necessarily indicate that participants would successfully resolve the problem, as the questionnaire did not capture the specific order or method in which participants intended to perform those actions.

Third person vs first person: Figure 41 to the right presents the questionnaire results for both the third person and first-person (baseline) perspectives, with percentages rounded to one decimal place (or two decimals if necessary).

The results show that the third-person perspective generally exhibits a higher rate of "Ok" actions, indicating a greater tendency for successful problem resolution compared to the first-person perspective, except for one localization error, "Navigate to Rodin." However, the difference between the participants' chosen actions is not substantial for all localization errors, indicating that both views are almost equal in terms of resolving localization errors.
When examining the other error types, the results become more varied. For collision and reachability errors, "Ok" actions are consistently higher for the third-person perspective, and participants in this group chose overall fewer bad actions compared to the first-person perspective.
Notably, the orientation scenarios displayed the most noteworthy disparity, with the highest divergence in "Ok" answers favoring the third-person view.
Additionally, the third-person perspective demonstrated a notable decrease in "Bad" actions, suggesting a reduced dependence on seeking help or manual arm teleoperation for orientation errors. These findings further align with the outcomes of the user studies discussed before.

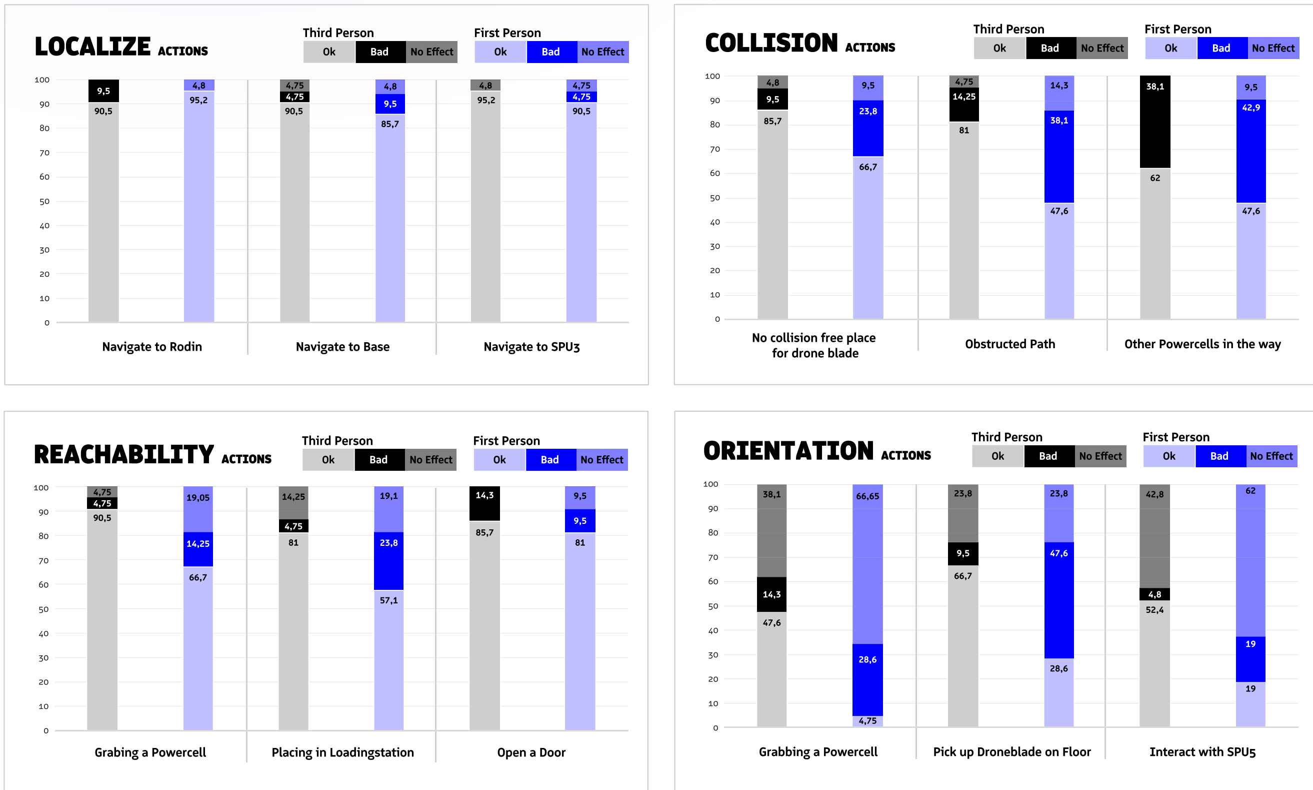


Figure 41: Action results for the 12 error scenarios categorized into localization, collision, reachability, and orientation errors. The results are depicted using grey beams for the third-person perspective and blue beams for the first-person perspective. The beams are further divided into the categories "Ok," "Bad," and "No effect" for each scenario, indicating the percentage of participant responses in each category.

Debug Page Comparison: In addition to comparing the third person and first-person perspectives, participants were also asked to identify appropriate actions using "No Help," the "Typical Error and Solution Debug Page", and the "Robot's Motion Planning". Figure 42 below presents the results for the four tested error scenarios, with the data categorized into the three respective categories: "Ok," "Bad," and "No Effect."

Overall the results for the three situations varied, with the robots motion planning resulting in the most "Ok" actions for Reachability, while " Typical Error and Solution Debug Page" shows the most "Ok" actions compared for Orientation and Localization errors. Surprisingly, in the Collision scenario, "Ok" actions were highest with "No Help" despite the availability of solutions in the typical error and solution page. It should be noted that the typical error and solution page offers multiple solutions, emphasizing the importance of trying autonomous commands before resorting to manual navigation. In this specific collision case, autonomous navigation did not resolve the issue, necessitating manual navigation. Examining the results in more detail (see Appendix p. 198), it becomes evident that participants shifted towards autonomous navigation in the " Typical Error and Solution Debug Page" condition, explaining the reduction in "Ok" actions in this scenario.

Turning back to the overall findings, the "Typical Error and Solution Debug Page" exhibited the lowest occurrence of "Bad" actions. Conversely, the "Robot's Motion Planning" approach notably increased the number of "Bad" actions, particularly in the Collision and Orientation scenarios.

In summary, the A/B between subjects' study collected data from 42 participants (Group A: 21, Group B: 21) regarding error identification and suggested next actions. The third-person perspective consistently outperformed the first-person perspective in correctly identifying errors, particularly in collision and orientation scenarios. The typical errors and solutions debug page proved to be the most effective in error identification across all error types. Motion planning of the robot in the debug page showed mixed results, with a decline in localization error identification but improvements in collision and orientation. Regarding suggested next actions, the third-person perspective exhibited a higher rate of "Ok" actions and a notable decrease in "Bad" actions, suggesting a greater tendency for successful problem resolution compared to the first-person perspective. The typical error and solution debug page showed the highest number of "Ok" actions, while the robot's motion planning increased the occurrence of "Bad" actions, particularly in collision and orientation scenarios. These findings highlight the potential benefits of the third-person perspective and the effectiveness of the typical errors and solutions debug page in addressing errors.

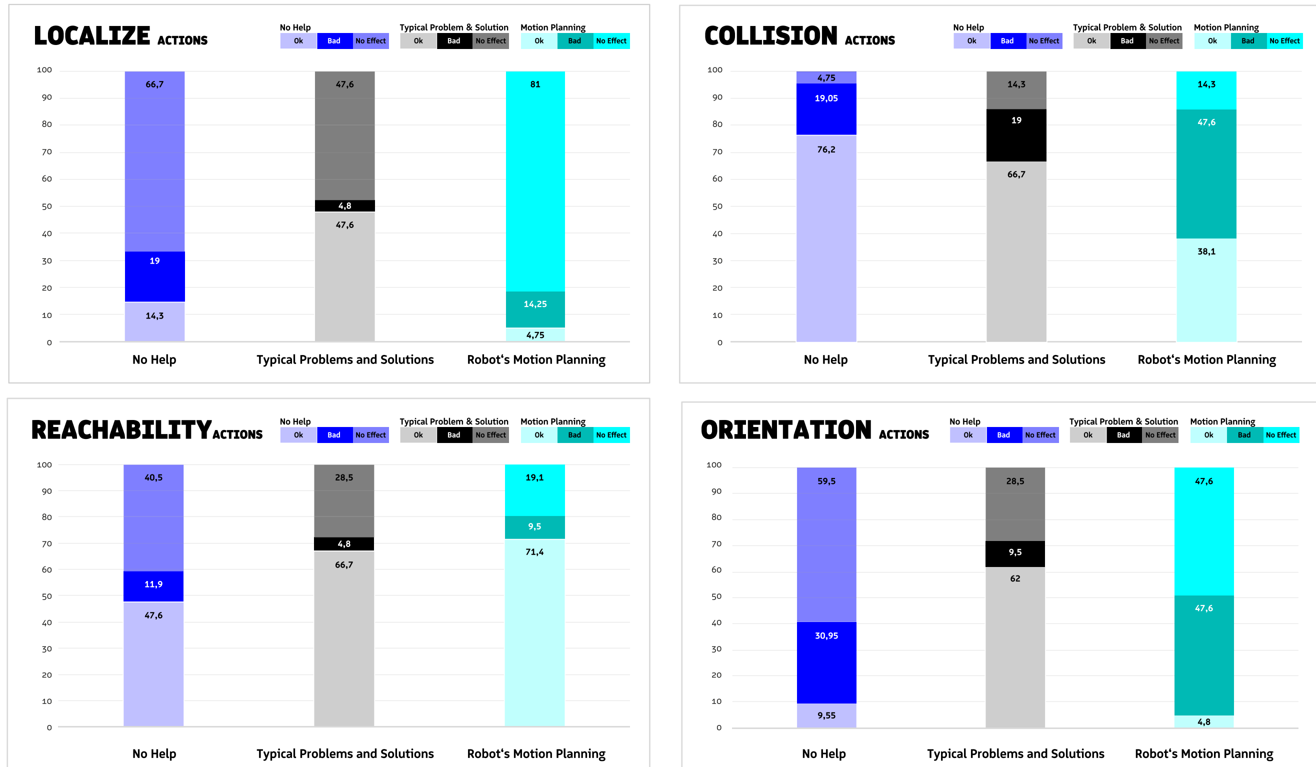


Figure 42: Action results for the 4 error scenarios categorized into localization, collision, reachability, and orientation errors. The results are depicted using blue beams for baseline, grey beams for "Typical Error and Solutions" and cyan beams for the Robots motion planning. The beams are further divided into the categories "Ok," "Bad," and "No effect" for each scenario, indicating the percentage of participant responses in each category.

Prototype Testing: Result Overview

Overall, the user study results for both prototypes indicate notable improvements in user experience and a reduction in cognitive workload compared to the to the first-person perspective baseline.

User Study: Typical Errors and Solutions Debug Page
Concerning error resolution, the "Typical Errors and Solutions" Debug Page was effective across all error types. Participants successfully identified and resolved Reachability, Collision, and Orientation errors based on their initial guesses and Localization errors typically required an average of 1.25 attempts. Overall, approaching errors with this prototype took participants roughly one minute to solve. The effectiveness of this prototypes is further underlined by participant comments such as the helpfulness of the provided solution strategies (mentioned by 7/8), the provided guidance through the instruction design and layout (mentioned by 6/8) and the potential of the prototype to aid users in understanding future error scenarios (mentioned by 8/8).

User Study: Typical Errors and Solutions Debug Page
When comparing the third person perspective to the first-person view, the average attempts and time spent are similar. Participants from both groups (third person and baseline comparison) were able to resolve Reachability and Collision error without external help. The most notable difference in usage and error resolution was exhibited for the Orientation error. Regarding the orientation error, participants from the third person perspective required less external help (4/8 needing help) compared to the first-person participants (7/8 needing help). Participants form the third person prototype also spent more time and attempts to resolve the error. This result indicates a higher level of engagement and a stronger willingness to persist independently, which is further underlined by the demonstrated lower level of frustration (as shown by the NASA TLX scores) and participants specific comments of engagement (mentioned by 7/8). Overall, all participants believe the third person perspective to aid in mitigating errors as it helps their situational awareness (mentioned by 8/8) and general improved overview of the environment and robot (mentioned by 6/8). Nevertheless, orientation error remains difficult without additional guidance. For the remaining difficulties, the discussed game elements provide directions how one could further aid users in the third person perspective, see Further Recommendations (next page) of how that could look like.

Questionnaire Results:
Error Identification and Resolution Actions
Overall, the A/B between subjects study involving 42 participants (Group A: 21, Group B: 21) yielded valuable insights into error identification and suggested next actions.

The results consistently favored the third-person perspective over the first-person perspective in accurately identifying errors, especially in collision and orientation scenarios.
When considering suggested next actions, the third-person perspective exhibited a higher rate of "Ok" actions and a notable decrease in "Bad" actions, indicating a greater tendency for successful problem resolution compared to the first-person perspective.

The typical errors and solutions debug page emerged as the most effective in error identification across all error types. While the motion planning of the robot in the debug page showed mixed results, with a decline in localization error identification but improvements in collision and orientation.
Concerning suggested next actions, the typical error and solution debug page demonstrated the highest number of "Ok" actions for Orientation and Localization errors, whereas the robot's motion planning was best for Reachability errors. However, the robot's motion planning also increased the occurrence of "Bad" actions, particularly in collision and orientation scenarios.

Further
Recommendations

This section discusses how the proposed prototype could be improved and provides suggestions on how to integrate them into the current system. The prototypes not only include additional features but also significant user interface changes, particularly the integration of the third-person perspective.

Virtual vs. Live Stream

Taking the Best of Both
The third-person perspective offers a virtual environment where users can freely explore and change perspectives without directly controlling the robot, while the current system displays a live camera feed from the robot's first-person view. Previous research [78, 87] has shown a simple solution for integrating a third-person view by mounting a camera on a robotic arm or attaching it to a long stick, which preserves the live feedback aspect. However, a virtual environment presents advanced capabilities that have been tested before [87, 108], as discussed in the Related Work section.
The virtual environment provides several advantages. Users can explore the environment without any time delay since it doesn't require the robot to physically move. They can investigate unfamiliar objects by changing their viewpoint, enabling them to make more informed decisions before commanding the robot. This feature is particularly useful for navigation tasks, allowing users to assess whether manual teleoperation or autonomous navigation is preferable based on the presence of obstacles. Moreover, the orientation of the robot relative to other elements and the world becomes more apparent and easily investigable. These freedoms are not as easily achievable with mounted cameras due to bandwidth limitations, increased movement required form the robot, and associated operational delays. Additionally, camera footage is not always clear, and the weather and light conditions on Mars can restrict visibility in certain areas.

However, there are contexts where a camera perspective is essential and cannot be replaced. For fine-grained manual control tasks or investigations of broken/damaged objects that cannot be accurately represented or captured in a virtual environment, camera views remain necessary. Additionally, as the environment becomes more dynamic, such as with multiple robots working or significant changes in weather conditions, it becomes increasingly challenging to maintain an accurate virtual representation, making camera views more favorable in such situations.

To maximize the benefits of both virtual and live stream environments, this work suggests integrating a combination of the two, allowing users to switch between the third person virtual environment and the robot's camera view as needed. The third person prototype already facilitates this by presenting both views, with one smaller on the side and the active view in the center. The switching mechanism can be implemented through a button click or within the user interface. Implementing the virtual environment is relatively straightforward, considering the static nature of the current testing environment without weather influences and consistent lighting conditions. For future mars missions current research is already working on how to implement such virtual environment, even considering photorealistic representations by accurately matching 2D images to 3D shapes [87].

Figure 43: Recommended features for integration, including a Localization percentage overlay (1) for easier object verification, a progress bar (2) indicating the robot's current state and activity, the ability to click on objects directly for commands (3), and the visual indication of unavailable commands on objects (4) to guide the user's actions.

Additional Features for the Third Person Perspective
This virtual environment could be further improved by integrating discussed features from the Feature Discussion, p. 120. Here participants expressed strong preferences for certain features, making it easier for them to understand certain limitations of the robot in context and receive more intuitive feedback throughout the usage. Therefore, the recommended features to integrate include a (1) Localization percentage overlay, so that user can more easily check whether an object was located correctly. Additionally, displaying a (2) progress bar that goes from planning to execution to being done, so that the user receives more feedback of what state the robot is currently in and that something is happening even if the robot does not move. Furthermore, allowing users to (3) click on objects directly for commands instead of searching for commands on the side panel (as the current system does it). By having actions directly on object, users can also see if actions are (4) currently not available, as commands are grayed out and indicator text will tell them to bring the object into view, guiding the user in what they must do. An illustration of how this could look implemented can be seen in Figure 43.

Features specifically for aiding in error situation include displaying (B) a reachability map of the robot in case of reachability error, (C) highlighting potential colliding objects to indicate what might be in the way, (D) highlighting the different body parts of the robot in case of an orientation issue so that users' attention gets drawn more to the misalignment and (E) Augmented highlighters on AprilTags in case AprilTags are not fully visible. In terms of error communication, (A) generally, rephrasing errors to be more apologetic and asking for help instead of simply stating that an error occurred is recommended, see Figure 44, next page.

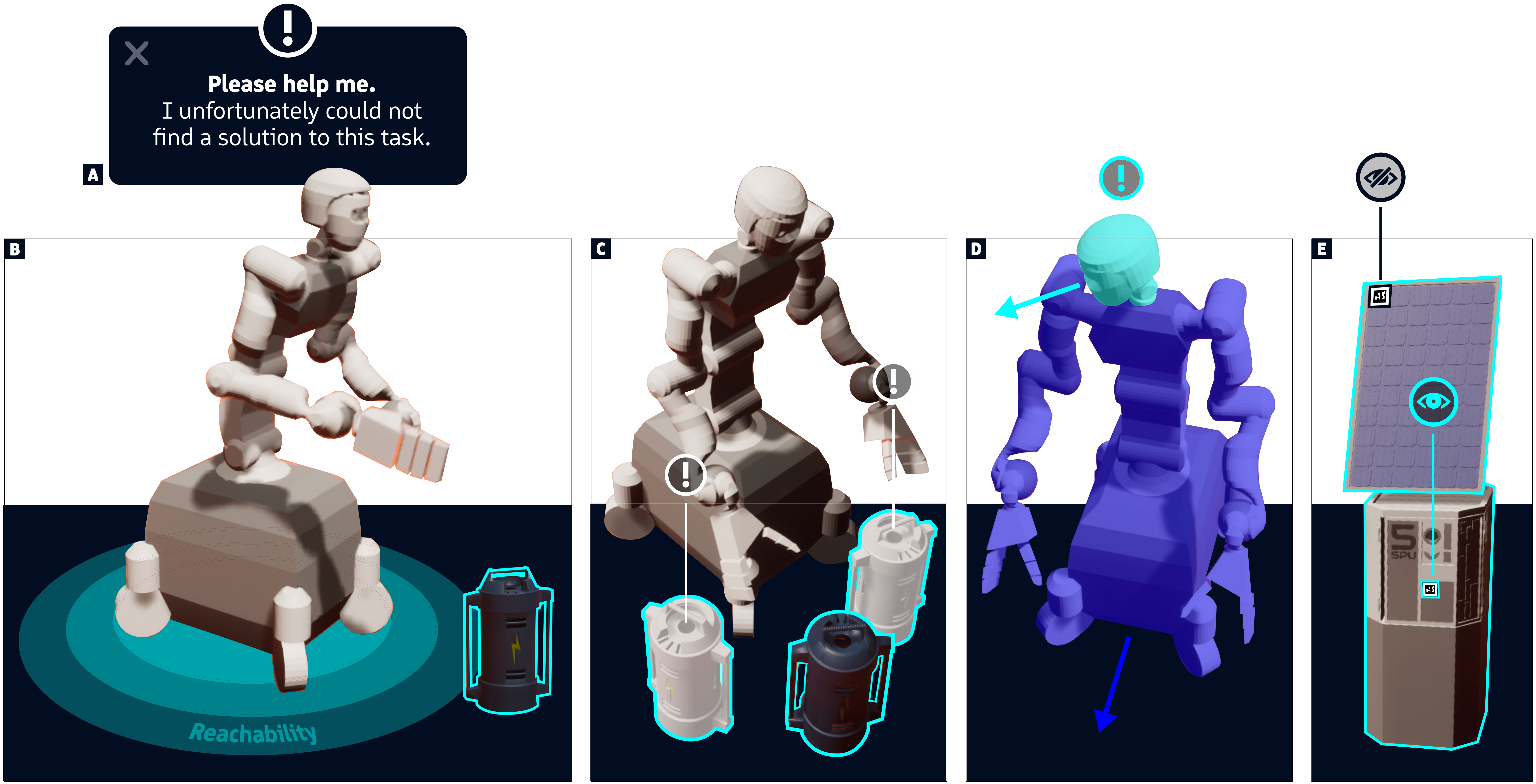


Figure 44: proposed features aimed at aiding in error situations. These include: To rephrase errors to be (A) more apologetic, a reachability map (B) to address reachability errors, highlighting potential colliding objects (C) to indicate obstructions, highlighting different body parts of the robot (D) for orientation issues, and augmented highlighters on AprilTags (E) when they are not fully visible.

Debug Page: Combining it with the Third Person Prototype

The other core aspect of this work focused on additional features that directly help with error identification and resolution. Building on the effectiveness of the third person perspective in error mitigation, the remaining errors that simply require more guidance, such as orientation errors, could be handled by integrating the "Typical Error and Solution" debug page. The advantage of the "Typical Error and Solution" debug page is that it provides the user with information in the moment of need, but only if they need it, as it is expandable, therefore considering user expertise. The information of this debug page would be too much to teach to astronauts, as the training are limited in time and dealing with errors is not the core of what they are supposed to learn. After all, the robot barely throws an error. By integrating the debug page, users receive more assistance in error situations (if not mitigated by the third person perspective). The results from the testing also suggest that users are helped very effectively, with this tool.

Improving the Debug Page

To enhance the debug page's usability, reducing the amount of text can minimize reading requirements, while providing more detailed instructions as foldout text can offer additional clarity, see Figure 45. Moreover, to increase efficiency, the proposed solution strategies currently shown in the debug page could be made interactive, allowing users to directly command actions from the debug page. Another approach could be to highlight the proposed solutions in the User interface if clicked on them in the debug page. Both options would aid the user more in executing the recommended commands.

Figure 45 showcases a summary of how the recommendation and suggestions could look accumulated, with a third person perspective that allows users to switch, additional features and the integration of the adapted Debug Page.

Generally: Helping Astronauts Understand

Lastly, there are crucial aspects that, if known by astronauts, would enhance their overall usage of the system and streamline current experiment testing. These core aspects include:

1. Understanding the robot-specific quirks: Providing a brief explanation of the robot's idiosyncrasies, such as its preference for the right arm, its limitation in localizing one object at a time, or its cautiousness to avoid collisions, can prevent confusion and equip astronauts with a better understanding of certain issues and required commands.

2. Emphasizing the use of autonomous commands over manual teleoperation: By explaining to users the advantages of autonomous commands, such as the robot's expertise in positioning itself accurately, astronauts can be encouraged to rely on autonomous functions more frequently. For instance, the camera view currently presents challenges in perceiving the proximity to objects. Assuring users that the robot will position itself accurately through autonomous commands can alleviate stress and simplify the astronaut's experience. This reassurance eliminates the need for constant manual readjustments, making the interaction less burdensome for astronauts and more efficient in general.

3.Introducing potential error causes: Familiarizing astronauts with potential error causes empowers them to use the robot more effectively and prepares them for any unforeseen errors that may occur. This knowledge provides a better comprehension of how the robot operates and allows for more informed decision-making during usage.

It is important to note that introducing these topics is not intended to make astronauts experts, as the proposed features and additional information on the debug page are available to assist them when needed. Recognizing that it is unrealistic to expect astronauts to remember every detail, a one-time mention of these core aspects helps in establishing a foundation of understanding and equips astronauts with a better grasp of the system they are working with.

Figure 45: Un example of an adapted "Typical Errors and Solutions" debug page showcasing enhancements such as expandable text for detailed explanations and direct command interaction. These improvements are integrated into the third-person perspective.

Discus-sion and Conclu-sion

In this final chapter, the work of this master's thesis is concluded with a comprehensive discussion on the overall process, key findings, and limitations. The chapter reflects on the chosen approach and highlights the significance of design in the context of space exploration.

Discussion

The initial research conducted in this study involved user studies and analysis of astronaut videos, revealing limitations with the current surface avatar system. These limitations primarily centered around situation awareness problems, such as a limited view, time delays in system usage, and a general lack of knowledge among users about how the robot operates. Additionally, it became evident that the astronauts involved in the current experiments are novice users who received only initial training, resulting in them frequently seeking assistance. The existing tools in the user interface that were intended to help users resolve errors were either not used or used sparingly, leaving users feeling helpless when confronted with uninformative error messages that simply stated "Error message: None."

In comparison to previous research, this work builds upon the importance of situational awareness, which has been widely recognized in teleoperation contexts. By adopting the third-person perspective, this study adds a more user-centered and user experience-focused approach to the existing research. On the other hand, limited research has been conducted on appropriate error communication and user guidance. This work addresses this gap by proposing methods for effectively communicating errors and providing user guidance, specifically through the prototypes of the typical errors and solutions page and motion planning. The research presented here contributes to the field of teleoperation by introducing novel concepts inspired by gaming, gestalt principles, and existing research, providing future researchers with approaches for designing error communication and mitigation systems.

The findings of this work highlight the impact that design choices have on the user experience in teleoperation, particularly in the context of space robot teleoperation. The results from the prototypes showcased improvements in user engagement, persistence, and guidance when compared to the current setup. These outcomes emphasize how well-designed systems can not only guide users toward making correct decisions but also influence their behavior and interactions with the system as a whole. The implications of this work can extend beyond the space exploration industry and have potential applications in various sectors, including general remote teleoperation or remote service robot operation.

Limitations

It is important to acknowledge the limitations of this work. Firstly, the studies conducted were not carried out directly with astronauts, the core users of the system. While efforts were made to address their needs and preferences, working directly with astronauts would have provided a more comprehensive understanding of their requirements and an ideal setup. Secondly, to determine the actual statistical significance of the gathered results, statistical analyses should be performed, considering the limitations of the participant numbers. Lastly, as the robot and the surface avatar system were not always available and the testing for errors was challenging to simulate, alternative methods such as cognitive walkthroughs and virtual simulations were utilized. Although these setups provided advantages such as faster testing and the ability to simulate various situations, they may not accurately represent the actual surface avatar system, suggesting the need for further research using the real setup.

Looking towards future directions, the morphological chart showcasing potential game elements and cues presents a multitude of ideas that could aid users in teleoperation systems. Further research can focus on testing these ideas, starting with the proposed features from the Further Recommendations section. Additionally, a convergence of multiple concepts, as suggested in this work, could be tested as a whole. Implementing and testing this setup, ideally with astronauts in a virtual environment to streamline the process, would address some of the limitations encountered in this study and serve as a cornerstone for future research.

In conclusion, this work highlights the importance of design choices in enhancing the user experience in teleoperation systems. By addressing the identified limitations, exploring future directions, and conducting further research, significant advancements can be made in error communication and user guidance.

Conclusion

This work has examined the challenges associated with the current surface avatar system in teleoperation contexts, specifically focusing on the limitations of situational awareness and error communication. Through user studies, comparative questionnaires, and the development of prototypes, insights and improvements have been achieved.

The integration of a third-person perspective, as demonstrated in the proposed prototypes, addresses the limitations of limited views and enhances users' understanding of the environment. The results have shown improved user experience, reduced cognitive workload, increased error identification rates, and higher levels of engagement and persistence.
Furthermore, this work contributes to the field by addressing the gap in research on effective error communication and user guidance in teleoperation. The prototype of the typical error and solution page provide users with clear instructions and support in resolving errors. By leveraging concepts from gaming, gestalt principles, and existing research, this work has laid the groundwork for designing effective error communication and mitigation strategies in teleoperation systems.

The impact of well-designed interfaces on user behavior and interaction patterns emphasizes the significance of user-centered design in teleoperation systems.

In summary, this work has contributed to the understanding of improving user experience, situational awareness, and error communication in teleoperation systems. By integrating the proposed prototypes, refining design concepts, and addressing the identified limitations, we can pave the way for more efficient and user-friendly teleoperation systems not only in space exploration but also in other remote teleoperation domains.

Reflection

My personal interest in this project is deeply rooted in my lifelong passion for both science and design. From an early age, I have been captivated by the wonders of scientific exploration and the potential for design to simplify and enhance our everyday lives. I see design as a tool to navigate the complexities of our world and create seamless communication experiences, acting as a bridge between intricate contexts and end-users.

Unfortunately, design is not always recognized as a beneficial force and is often relegated to the role of mere aesthetics. While industrial design, for example, is acknowledged for its usefulness due to the tangible aspects of ergonomics, the significance of design in the context of interface design can be more elusive, as the impact of design on interaction and controls may not be immediately obvious, since it operates on a more abstract level.

However, the true value of good user interface design lies in its convenience, as it minimizes the user's effort and enables even novices to effectively interact with complex systems, such as smartphones. A prime illustration of this is the evolution of computers. In the past, executing tasks like changing folders necessitated the use of terminal commands. Today, thanks to intelligent design decisions, we can effortlessly drag and drop items, create folders with a simple mouse click, and experience numerous other improvements that are often overlooked as the result of design contributions.

In contrast to personal computers or smartphones, where there is a strong commercial drive to enhance usability and user-friendliness, the research domain of teleoperation of robots in space lacks this commercial incentive. Consequently, design opportunities to make a substantial impact are seldom explored. This makes my work particularly significant to me, as the German Aerospace Center afforded me the freedom to leverage the power of design in enhancing the astronaut experience. Not just for space exploration, but in general, there are compelling reasons to embrace design principles in research: by considering human perception, we can optimize system efficiency, reduce user workload, guide user behavior, and ultimately make experiments and explorations—such as teleoperation in the context of space—more efficient and seamless. It is crucial to recognize that design encompasses more than aesthetics; it holds the potential to unlock user safety, speed, and efficiency improvements in various processes.

With my work, I have strived to challenge this misconception and shed light on the transformative power of design. By focusing on the intricate details of user interface design and user experience in the realm of teleoperation of robots in space, I aimed to demonstrate how design can bring about tangible improvements. It is my sincere aspiration that through my contributions, a broader understanding of design's vital role in enhancing functionality and advancing progress will be fostered. Moreover, to start conversation that encourages future researchers to recognize the crucial role of design and actively collaborate with designers to unlock the full potential of their endeavors.

Refer-ences

References

[1] A. Gawel, Y. Lin, T. Koutros, R. Siegwart and C. Cadena, "Aerial-Ground collaborative sensing: Third-Person view for teleoperation," 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Philadelphia, PA, USA, 2018, pp. 1-7, doi: 10.1109/SSRR.2018.8468657

[2] A. Schiele, "METERON - Validating Orbit-to-Ground Telerobotics Operations Technologies," in 11th Symposium on Advanced Space Technologies for Robotics and Automation (ASTRA), 2011.

[3] A. Schiele, M. Aiple, T. Krueger, F. van der Hulst, S. Kimmer, J. Smisek, and E. den Exter, "Haptics-1: Preliminary Results from the First Stiffness JND Identification Experiment in Space," in Proc. of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, pp. 13–22, Springer, 2016.

[4] A. Schiele, T. Kr¨uger, S. Kimmer, M. Aiple, J. Rebelo, J. Smisek, E. den Exter, E. Mattheson, A. Hernandez, and F. van der Hulst, "Haptics-2 - A System for Bilateral Control Experiments from Space to Ground via Geosynchronous Satellites," in IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2016.

[5] Adamides G, Katsanos C, Parmet Y, Christou G, Xenos M, Hadzilacos T, Edan Y. HRI usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer. Appl Ergon. 2017 Jul;62:237-246. doi: 10.1016/j.apergo.2017.03.008. Epub 2017 Mar 22. PMID: 28411734

[6] AKM Bahalul Haque, A.K.M. Najmul Islam, Patrick Mikalef, Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research, Technological Forecasting and Social Change, Volume 186, Part A, 2023, 122120, ISSN 0040-1625, https://doi.org/10.1016/j.techfore.2022.122120

[7] Alexiou, A., Schippers, M.C. Digital game elements, user experience and learning: A conceptual framework. Educ Inf Technol 23, 2545–2567 (2018). https://doi.org/10.1007/s10639-018-9730-6

[8] Amal Nanavati, Patricia Alves-Oliveira, Tyler Schrenk, Ethan K. Gordon, Maya Cakmak, and Siddhartha S. Srinivasa. 2023. Design Principles for Robot-Assisted Feeding in Social Contexts. In Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23). Association for Computing Machinery, New York, NY, USA, 24–33. https://doi.org/10.1145/3568162.3576988

[9] Astronaut selection 2021-22 FAQs. (n.d.). esa.int. Retrieved April 6, 2023, from https://www.esa.int/About_Us/Careers_at_ESA/ESA_Astronaut_Selection/Astronaut_selection_2021-22_FAQs

[10] Baker, A. L., Phillips, E. K., Ullman, D., & Keebler, J. R. (2018). Toward an understanding of trust repair in human-robot interaction: current research and future directions. ACM Transactions on Interactive Intelligent Systems (TiiS), 8(4), 1–30.

[11] Barr, P., Noble, J., & Biddle, R. (2007). Video game values: Human–computer interaction and games. Interacting with Computers, 19(2), 180–195. doi:10.1016/j.intcom.2006.08.008

[12] Bartneck, C., Croft, E., Kulic, D. & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. International Journal of Social Robotics, 1(1) 71-81. | DOI: 10.1007/s12369-008-0001-3

[13] Burelli, P. (2016). Game Cinematography: From Camera Control to Player Emotions. In: Karpouzis, K., Yannakakis, G. (eds) Emotion in Games. Socio-Affective Computing, vol 4. Springer, Cham. https://doi.org/10.1007/978-3-319-41316-7_11

[14] Byrne, M. L. (2001b). Interviewing as a data collection method. Association of perioperative Registered Nurses (AORN) Journal, 74(2), p. 233–235. https://doi.org/10.1016/s0001-2092(06)61533-0

[15] C. Gutwin, G. Stark and S. Greenberg, "Support for workspace awareness in educational groupware," Proc. Comp. Supported Collaborative Learning (CSCL), 1995.
[16] C. Gutwin, S. Greenberg and M. Roseman, Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation, University of Calgary, 1996.

[17] Caroline W. Dos Santos, Nelson L.D. Filho, Danúbia B. Espíndola, Sílvia S.C. Botelho, Situational Awareness Oriented Interfaces on Human-Robot Interaction for Industrial Welding Processes, IFAC-PapersOnLine, Volume 53, Issue 2, 2020, Pages 10168-10173, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2020.12.2744.

[18] Carroll, J. M. (2000). Five reasons for scenario-based design. Interacting With Computers, 13(1), p. 43–60. https://doi.org/10.1016/s0953-5438(00)00023-0

[19] Coates, E. (2023). Waypoints and Markers | Game UI Database. Game UI Database. Retrieved April 9, 2023, from https://www.gameuidatabase.com/index.php?scrn=163

[20] Connor Esterwood, Lionel P. Robert Jr, Three Strikes and you are out!: The impacts of multiple human–robot trust violations and repairs on robot trustworthiness, Computers in Human Behavior, Volume 142, 2023, 107658, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2023.107658.

[21] Cuevas, H. M., and Aguiar, M. (2017). Assessing situation awareness in unmanned aircraft systems operations. Int. J. Aviat. Aeronaut. Aerosp. 4:3. doi: 10.15394/ijaaa.2017.1176

[22] D. Leidner, C. Borst and G. Hirzinger, "Things are made for what they are: Solving manipulation tasks by using functional object classes," 2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012), Osaka, Japan, 2012, pp. 429-435, doi: 10.1109/HUMANOIDS.2012.6651555.

[23] D. Lester, K. Hodges, C. Ower, and K. Klaus, "Exploration telepresence from Earth-Moon Lagrange points," in Proceedings of the IAF/AIAA Global Space Exploration Conference. International Astronautical Federation, 2012.

[24] Das, D., Banerjee, S., & Chernova, S. (2021, March). Explainable ai for robot failures: Generating explanations that improve user assistance in fault recovery. In Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (pp. 351-360).

[25] Desmet, P.M.A. (2003). Measuring emotion; development and application of an instrument to measure emotional responses to products. In: M.A. Blythe, A.F. Monk, K. Overbeeke, & P.C. Wright (Eds.), Funology: from Usability to Enjoyment (pp. 111-123). Dordrecht: Kluwer Academic Publishers.
[26] Display and Graphics Commonality Standard International Space Station Program. (2001). In https://archive.org/details/ssp50313 (SSP-50313). NASA.
[27] Du, N., Haspiel, J., Zhang, Q., Tilbury, D., Pradhan, A. K., Yang, X. J., et al. (2019). Look who's talking now: Implications of AV's explanations on driver's trust, AV preference, anxiety, and mental workload. Transportation Research Part C (Emerging Technologies), 104, 428–442.

[28] Dumas, J. & Redish, J. (1999). A practical guide to usability testing (Revised edition). Exeter, UK: Intellect

[29] E. Dupuis, P. Langlois, J. L. Bedwani, D. Gingras, A. Salerno, P. Allard, S. Gemme, R. L'Archevêque, and T. Lamarche, "The Avatar-EXPLORE experiments: Results and lessons learned," in Proceedings of the International Symposium on Artificial Intelligence, Robotics, and Automation in Space, 2010.

[30] ESA - European Space Agency. (n.d.). Why go to Mars? Retrieved June 14, 2023, from https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Exploration/Why_go_to_Mars

[31] Fink, J., Mubin, O., Kaplan, F., and Dillenbourg, P. (2012). "Anthropomorphic language in online forums about Roomba, AIBO and the iPad," in Advanced Robotics and its Social Impacts (ARSO), IEEE Workshop on 2012 (Munich), 54–59.

[32] Fong, T., Thorpe, C. Vehicle Teleoperation Interfaces. Autonomous Robots 11, 9–18 (2001). https://doi.org/10.1023/A:1011295826834

[33] Gatsoulis, Y., Virk, G. S., & Dehghani-Sanij, A. A. (2010). On the Measurement of Situation Awareness for Effective Human-Robot Interaction in Teleoperated Systems. Journal of Cognitive Engineering and Decision Making, 4(1), 69–98. https://doi.org/10.1518/155534310X495591

[34] Gómez-Albarrán, M., Sarasa-Cabezuelo, A., Sierra-Rodríguez, JL. et al. Authoring and playing interactive fiction with conventional web technologies. Multimed Tools Appl 81, 14705–14747 (2022). https://doi.org/10.1007/s11042-021-11316-x

[35] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Advances in psychology (pp. 139–183). Elsevier BV. https://doi.org/10.1016/s0166-4115(08)62386-9

[36] Heijne, K., & Van Der Meer, H. (2019). Road map for creative problem solving techniques: organizing and facilitating group sessions. Amsterdam: Boom

[37] Hijaz, A., Louie, W.-Y. G., and Mansour, I. (2019). "Towards a driver monitoring system for estimating driver situational awareness," in 28th IEEE International Conference on Robot and Human Interactive Communication (New Delhi), 1–6. doi: 10.1109/RO-MAN46459.2019.8956378

[38] Honig, S., & Oron-Gilad, T. (2018). Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.00861.

[39] Hooman Hedayati, Michael Walker, and Daniel Szafir. 2018. Improving Collocated Robot Teleoperation with Augmented Reality. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18). Association for Computing Machinery, New York, NY, USA, 78–86. https://doi.org/10.1145/3171221.3171251

[40] Houts, P. S., Doak, C. C., Doak, L. G., and Loscalzo, M. J. (2006). The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ. Couns. 61, 173–190. doi: 10.1016/j.pec.2005.05.004

[41] How to become an astronaut. (n.d.). esa.int. Retrieved April 6, 2023, from https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Astronauts/How_to_become_an_astronaut

[42] International Space Exploration Coordination Group, "The Global Exploration Roadmap," www.globalspaceexploration.org, January 2018.

[43] International Organization for Standardization. (2019). Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems (ISO 9241-210:2019(en)). https://www.iso.org/standard/52075.html

[44] J. Artigas, R. Balachandran, C. Riecke, M. Stelzer, B. Weber, J.-H. Ryu, and A. Albu-Schaeffer, "KONTUR-2: Force-Feedback Teleoperation from the International Space Station," in IEEE International Conference on Robotics and Automation (ICRA), 2016.

[45] J. Burns, D. Kring, J. Hopkins, S. Norris, T. Lazio, and J. Kasper, "A lunar L2-Farside exploration and science mission concept with the Orion Multi-Purpose Crew Vehicle and a teleoperated lander/rover," Advances in Space Research, vol. 52, pp. 306–320, 2013.

[46] J. L. Drury, J. Scholtz and H. A. Yanco, "Awareness in human-robot interactions," SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483), Washington, DC, USA, 2003, pp. 912-918 vol.1, doi: 10.1109/ICSMC.2003.1243931.

[47] Joyce, A. (2019, May 19). 10 Usability Heuristics Applied to Video Games. Nielsen Norman Group. Retrieved April 19, 2023, from https://www.nngroup.com/articles/usability-heuristics-applied-video-games/

[48] Kaber, D. B.,Wright,M. C., and Sheik-Nainar,M. A. (2006). Investigation of multimodal interface features for adaptive automation of a human-robot system. Int. J. Hum. Comput. Stud. 64, 527–540. doi: 10.1016/j.ijhcs.2005.11.003

[49] Keenan, C. (2021, August 13). A Beginner's Guide to Video Games: Perspectives — The SMU Journal. The SMU Journal. https://www.thesmujournal.ca/video-games/a-beginners-guide-to-video-games-perspectives

[50] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex Mitchell. 2018. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 140, 1–12. https://doi.org/10.1145/3173574.3173714

[51] Krueger, Thomas und van der Hulst, Frank und Ferreira, Edmundo und Wormnes, Kjetil und den Exter, Emiel und Gherghescu, Andrei und Gerdes, Levin und Hann, Lukas und Pereira, Aaron und Lii, Neal Y. und Conti, Francois (2020) A Newcomer's Guide to the Challenges of a Complex Space-to-Ground Experiment, With Lessons from Analog-1. In: International Symposium on Artificial Intelligence, Robotics and Automation in Space (I-SAIRAS) 2020. International Symposium on Artificial Intelligence, Robotics and Automation in Space (I-SAIRAS) 2020, 19-23 Oct 2020, Pasadena, CA, USA.

[52] Law, E., Cai, Vicky., Liu, Q. F., Sasy, S., Goh, J., Blidaru, A., et al. (2017). "A Wizard-of-Oz study of curiosity in human-robot interaction," in 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2017) (Lisbon), 607–614.

[53] Lee, J., Hirano, T., Hano, T., and Itoh, M. (2019). "Conversation during partially automated driving: how attention arousal is effective on maintaining situation awareness," in 2019 IEEE International Conference on Systems, Man and Cybernetics (Bari), 3718–3724. doi: 10.1109/SMC.2019.8914632

[54] Lee, M. K., Kiesler, S., Forlizzi, J., Srinivasa, S., and Rybski, P. (2010). "Gracefully mitigating breakdowns in robotic services," in 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (Osaka), 203–210.

[55] Leidner, D. (2022, September 13). Institute of Robotics and Mechatronics: Dr.-Ing. Daniel Leidner. rmc.dlr.de. Retrieved April 9, 2023, from https://rmc.dlr.de/rm/en/staff/daniel.leidner/

[56] Leidner, D. S. (2018d). Cognitive Reasoning for Compliant Robot Manipulation: Vol. 1st ed. 2019 Edition. Springer. http://nbn-resolving.de/urn:nbn:de:gbv:46-00106156-15

[57] Lemaignan, S., Fink, J., Mondada, F., and Dillenbourg, P. (2015). "You're doing it wrong! Studying unexpected behaviors in child-robot interaction," in Social Robotics. ICSR 2015. Lecture Notes in Computer Science 9388 LNCS (Cham; Paris: Springer), 390–400. doi: 10.1007/978-3-319-25554-5_39

[58] Lii, N.Y., Riecke, C., Leidner, D., Schätzle, S., Schmaus, P., Weber, B.M., Krueger, T., Stelzer, M., Wedler, A., & Grunwald, G. (2018). The Robot as an Avatar or Co-worker? An Investigation of the Different Teleoperation Modalities through the KONTUR-2 and METERON SUPVIS Justin Space Telerobotic Missions.

[59] Linkov, V., & Vanžura, M. (2021). Situation Awareness Measurement in Remotely Controlled Cars. Frontiers in psychology, 12, 592930. https://doi.org/10.3389/fpsyg.2021.592930

[60] Luo, J., He, W. and Yang, C. (2020), Combined perception, control, and learning for teleoperation: key technologies, applications, and challenges. Cogn. Comput. Syst., 2: 33-43. https://doi.org/10.1049/ccs.2020.0005

[61] M. Bualat, M. Deans, T. W. Fong, C. Provencher, D. Schreckenghost, and E. Smith, "ISS Crew Control of Surface Telerobots," in Proceedings of the IAF/AIAA Global Space Exploration Conference. International Astronautical Federation, 2012.

[62] M. Bualat, T. Fong, M. Allan, X. Bouyssounouse, T. Cohen, L. Flückiger, R. Gogna, L. Kobayashi, Y. Lee, S. Lee, C. Provencher, E. Smith, V. To, H. Utz, D. Wheeler et al., "Surface Telerobotics: Development and Testing of a Crew Controlled Planetary Rover System," in Proceedings of the AIAA SPACE Conference, 09 2013.

1 2 4 5 6 7 8 9 10

[63] M. Cardone, C. Laroque, M. Sarkarati, K. Nergaard, P. Steele, and S. Martin, "MOE: A System Infrastructure for Robotic Experiments," in Space Operations: Contributions from the Global Community, pp. 27–52, Springer, 2017.

[64] M. Panzirsch, A. Pereira, H. Singh, B. Weber, E. Ferreira, A. Gherghescu, L. Hann, E. den Exter, F. van der Hulst, L. Gerdes, L. Cencetti, K. Wormnes, J. Grenouilleau, W. Carey, R. Balachandran, T. Hulin, C. Ott, D. Leidner, A. Albu-Schäffer, N. Y. Lii, and T. Krüger, "Exploring planet geology through force-feedback telemanipulation from orbit," Science Robotics, vol. 7, no. 65, p. eabl6307, 2022.

[65] M. R. Endsley, "Design and evaluation for situation awareness enhancement," Proc Human Factors Society 32nd Annual Meeting, Santa Monica, CA, 1988.

[66] M. Sarkarati, M. Merri, K. Nergaard, and P. Steele, "How to plug-in your rover into a space mission to moon or mars," in Automation, Robotics and Applications (ICARA), 2015 6th International Conference on, pp. 318–324, IEEE, 2015.

[67] M. Stelzer, B. M. Steinmetz, P. Birkenkampf, J. Vogel, B. Brunner, and S. Kühne, "Software Architecture and Design of the Kontur-2 Mission," in IEEE Aerospace Conference, pp. 1–17, 2017.

[68] Mahoney, E. (2022). 2018 Global Exploration Roadmap. NASA. Retrieved April 6, 2023, from https://www.nasa.gov/exploration/about/isecg

[69] Maia Stiber, Russell H. Taylor, and Chien-Ming Huang. 2023. On Using Social Signals to Enable Flexible Error-Aware HRI. In Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23). Association for Computing Machinery, New York, NY, USA, 222–230. https://doi.org/10.1145/3568162.3576990

[70] Mars.Nasa.Gov. (n.d.). Mars Facts. NASA Mars Exploration. Retrieved June 14, 2023, from https://mars.nasa.gov/all-about-mars/facts/

[71] Matthew Rueben, Maja J. Matarić, Eitan Rothberg, and Matthew Tang. 2020. Estimating and Influencing User Mental Models of a Robot's Perceptual Capabilities: Initial Development and Pilot Study. In Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20). Association for Computing Machinery, New York, NY, USA, 418–420. https://doi.org/10.1145/3371382.3378392

[72] Matthew Rueben, Matthew Rodney Horrocks, Jennifer Eleanor Martinez, Nicolas LaLone, Marlena R Fraune, and Z Toups. 2021. [Hidden] / [Caution] / [Danger]: How Video Games Can Inform the Design of Sight Cues for Agents. In Proceedings of the 9th International Conference on Human-Agent Interaction (HAI '21). Association for Computing Machinery, New York, NY, USA, 396–401. https://doi.org/10.1145/3472307.3484682

[73] McDermott, P. L., Luck, J. A., Allender, L., and Fisher, A. (2005). Effective human to human communication of information provided by an unmanned vehicle. Proc. Hum. Factors Ergon. Soc. Ann. Meet. 49, 402–406. doi: 10.1177/154193120504900340

[74] MD Moniruzzaman, Alexander Rassau, Douglas Chai, Syed Mohammed Shamsul Islam, Teleoperation methods and enhancement techniques for mobile robots: A comprehensive survey, Robotics and Autonomous Systems, Volume 150, 2022, 103973, ISSN 0921-8890, https://doi.org/10.1016/j.robot.2021.103973

[75] Mubin, O., and Bartneck, C. (2015). "Do as I say: Exploring human response to a predictable and unpredictable robot," in Proceedings of the 2015 British HCI Conference on - British HCI '15 (Lincoln; New York, NY: ACM) 110–116.

[76] Murphy, R. R., and Tadokoro, S. (2019). User Interfaces for Human-Robot Interaction in Field Robotics. Springer International Publishing

[77] Mutlu, B. (2011). Designing Embodied Cues for Dialog with Robots. AI Magazine, 32(4), 17-30. https://doi.org/10.1609/aimag.v32i4.2376

[78] N. Shiroma, N. Sato, Y.-h. Chiu, and F. Matsuno, "Study on effective camera images for mobile robot teleoperation," in Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on. IEEE, 2004, pp. 107–112.

[79] N. Y. Lii, A. Schiele, D. Leidner, P. Birkenkampf, R. Bayer, B. Pleintinger, A. Meissner, and B. Andreas, "Simulating an Extraterrestrial Environment for Robotic Space Exploration: the METERON SUPVIS-Justin Telerobotic Experiment and The Solex Proving Ground," in 13th Symposium on Advanced Space Technologies for Robotics and Automation, (Noordwijk, The Netherlands), 2015.

[80] N. Y. Lii, D. Leidner, P. Birkenkampf, B. Pleintinger, B. Bayer, and T. Krueger, "Toward Scalable Intuitive Telecommand of Robots for Space Deployment with METERON SUPVIS Justin," in Symposium on Advanced Space Technologies

[81] N. Y. Lii, P. Schmaus, D. Leidner, T. Krueger, J. Grenouilleau, A. Pereira, ... A. Albu-Schäffer (2022). Introduction to Surface Avatar: The First Heterogeneous Robotic Team to be Commanded with Scalable Autonomy from the ISS. Proceedings of the International Astronautical Congress, IAC, IAC-22. International Astronautical Federation, IAF. 73rd International Astronautical Congress (IAC), 18-22 Sep 2022, Paris, France. ISSN 0074-1795.

[82] NASA Space Flight Human-System Standard Volume 1, Revision A: Crew Health | Standards. (2022c, January 5). https://standards.nasa.gov/standard/NASA/NASA-STD-3001-VOL-1

[83] National Aeronautics and Space Administration (NASA). (2016). NASA spinoff 2016. NASA. Retrieved April 6, 2023, from https://spinoff.nasa.gov/Spinoff2016/pdf/Spinoff2016.pdf

[84] Nesnas, I.A., Fesq, L.M. & Volpe, R.A. Autonomy for Space Robots: Past, Present, and Future. Curr Robot Rep 2, 251–263 (2021). https://doi.org/10.1007/s43154-021-00057-2

[85] Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158.

[86] Norton, A., Ober, W., Baraniecki, L., McCann, E., Scholtz, J., Shane, D., et al. (2017). Analysis of Human–Robot Interaction at the DARPA Robotics Challenge Finals. Int. J. Robot. Res. 36 (5-7), 483–513.

[87] Okura, F., Ueda, Y., Sato, T., & Yokoya, N. (2014). [Paper] Free-viewpoint Mobile Robot Teleoperation Interface Using View-dependent Geometry and Texture. ITE Transactions on Media Technology and Applications, 2(1), 82–93. https://doi.org/10.3169/mta.2.82

[88] P. Birkenkampf, D. Leidner and C. Borst, "A knowledge-driven shared autonomy human-robot interface for tablet computers," 2014 IEEE-RAS International Conference on Humanoid Robots, Madrid, Spain, 2014, pp. 152-159, doi: 10.1109/HUMANOIDS.2014.7041352.

[89] P. Schmaus et al., "Knowledge Driven Orbit-to-Ground Teleoperation of a Robot Coworker," in IEEE Robotics and Automation Letters, vol. 5, no. 1, pp. 143-150, Jan. 2020, doi: 10.1109/LRA.2019.2948128.

[90] P. Schmaus, D. Leidner, R. Bayer, B. Pleintinger, T. Krüger and N. Y. Lii, "Continued Advances in Supervised Autonomy User Interface Design for METERON SUPVIS Justin," 2019 IEEE Aerospace Conference, Big Sky, MT, USA, 2019, pp. 1-11, doi: 10.1109/AERO.2019.8741885.

[91] Papadimitriou, E., Schneider, C.h., Tello, J. A., Damen, W., Vrouenraets, M. L., and ten Broeke, A. (2020). Transport safety and human factors in the era of automation: what can transport modes learn from each other? Accid. Anal. Prev. 144:105656. doi: 10.1016/j.aap.2020.105656

[92] Peñín, L.F., Matsumoto, K. (2002). Teleoperation with time delay: A survey and its use in space robotics

[93] Pernice, K. (2017, November 12). F-Shaped Pattern of Reading on the Web: Misunderstood, But Still Relevant (Even on Mobile). Nielsen Norman Group. Retrieved April 6, 2023, from https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

[94] Piotr Fratczak, Yee Mey Goh, Peter Kinnell, Laura Justham, Andrea Soltoggio, Robot apology as a post-accident trust-recovery control strategy in industrial human-robot interaction, International Journal of Industrial Ergonomics, Volume 82, 2021, 103078, ISSN 0169-8141, https://doi.org/10.1016/j.ergon.2020.103078.

[95] Ratwani, R. M., McCurry, J. M., and Trafton, J. G. (2010). "Single operator, multiple robots: an eye movement based theoretic model of operator situation awareness," in Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction (Osaka), 235–242. doi: 10.1145/1734454.1734549

[96] Rea, D., & Seo, S. H. (2022). Still Not Solved: A Call for Renewed Focus on User-Centered Teleoperation Interfaces. Frontiers in Robotics and AI, 9. https://doi.org/10.3389/frobt.2022.704225

[97] Riley, J. M., Murphy, R. R., and Endsley, M. R. (2006). "Situation awareness in the control of unmanned ground vehicles," in Human Factors of Remotely Operated Vehicles, eds N. J. Cooke, H. L. Pringle, H. K. Pedersen, and O. Connor (Bingley: Emerald Group Publishing Limited), 359–371. doi: 10.1016/S1479-3601(05)07025-6

[98] S. Burigat, L. Chittaro and R. Sioni, "Mobile Three-Dimensional Maps for Wayfinding in Large and Complex Buildings: Empirical Comparison of First-Person Versus Third-Person Perspective," in IEEE Transactions on Human-Machine Systems, vol. 47, no. 6, pp. 1029-1039, Dec. 2017, doi: 10.1109/THMS.2017.2693684.

[99] S. Oleson, G. Landis, M. McGuire, and G. R. Schmidt, "HERRO Mission to Mars using Telerobotic Surface Exploration from Orbit," vol. 64, pp. 304–313, 09 2011.

[100] Sanders, E. B., & Stappers, P. J. (2012). Convivial Toolbox: Generative Research for the Front End of Design. Bis Pub.

[101] Schmaus, Peter und Leidner, Daniel und Krueger, Thomas und Grenouilleau, Jessica und Pereira, Aaron und Bauer, Adrian Simon und Bechtel, Nicolai und Bustamante Gomez, Samuel und Köpken, Anne und Lay, Florian Samuel und Sewtz, Marco und Batti, Nesrine und Ferreira, Edmundo und den Exter, Emiel und Bayer, Ralph und Pleintinger, Benedikt und Holderried, Roman und Pavelski, Pedro Henrique und Lii, Neal Yi-Sheng (2022) On Realizing Multi-Robot Command through Extending the Knowledge Driven Teleoperation Approach. In: Proceedings of the International Astronautical Congress, IAC, IAC-22. International Astronautical Federation, IAF. 73rd International Astronautical Congress (IAC), 18-22 Sep 2022, Paris, France. ISSN 0074-1795.

[102] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In: Marcus, A. (Ed.): Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science, Volume 8517, pp. 383-392, Springer International Publishing.

[103] Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. Organizational Behavior and Human Decision Processes, 101(1), 1–19.

[104] Stephen R Ellis, Katerina Mania, Bernard D Adelstein, Michael I Hill, Generalizeability of latency detection in a variety of virtual environments, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 48, no. 23, SAGE Publications, 2004, pp. 2632–2636.

[105] Tabrez, A., Luebbers, M.B. & Hayes, B. A Survey of Mental Modeling Techniques in Human–Robot Teaming. Curr Robot Rep 1, 259–267 (2020). https://doi.org/10.1007/s43154-020-00019-0

[106] Think Aloud Testing | Usability Body of Knowledge. (n.d.-c). Retrieved April 11, 2023, from http://www.usabilitybok.org/think-aloud-testing

[107] Thomas B. Sheridan, William R. Ferrell, Remote manipulative control with transmission delay, IEEE Trans. Hum. Factors Electron. (1) (1963) 25–29.

[108] Thomason, J. M., Ratsamee, P., Kiyokawa, K., Kriangkomol, P., Orlosky, J., Mashita, T., Uranishi, Y., & Takemura, H. (2017). Adaptive View Management for Drone Teleoperation in Complex 3D Structures. Intelligent User Interfaces. https://doi.org/10.1145/3025171.3025179

[109] Turchetti, G., Palla, I., Pierotti, F., and Cuschieri, A. (2012). Economic evaluation of da Vinci-assisted robotic surgery: A systematic review. Surg. Endosc. 26 (3), 598–606. doi:10.1007/s00464-011-1936-2

[110] Unreal Engine | Das stärkste Werkzeug für 3D-Echtzeit-Entwicklung. (n.d.). Unreal Engine. Retrieved June 08, 2023, from https://www.unrealengine.com/de

[111] Usability Testing | Usability Body of Knowledge. (n.d.). Retrieved April 11, 2023 from http://usabilitybok.org/usability-testing

[112] User and Task Analysis for Interface Design by JoAnn T. Hackos and Janice C. Redish

[113] van Boeijen, A. G. C., Daalhuizen, J., & Zijlstra, J. (2020). Delft Design Guide: Perspectives, models, approaches, methods. (2nd ed.) BIS Publishers. https://www.bispublishers.com/delft-design-guide-revised.html

[114] Vertesi, J. (2012). Seeing like a Rover: Visualization, embodiment, and interaction on the Mars Exploration Rover Mission. Social Studies of Science, 42(3), 393–414. http://www.jstor.org/stable/41721324.

[115] W. Carey, P. Schoonejans, B. Hufenbach, K. Nergaard, F. Bosquillon de Frescheville, J. Grenouilleau, and A. Schiele, "METERON: A mission concept proposal for preparation of human-robotic exploration," in Proceedings of the IAF/AIAA Global Space Exploration Conference. International Astronautical Federation, 2012.

[116] Wakabayashi, S., Magruder, F., & Bluethmann, W. (2003). Test of Operator Endurance in the Teleoperation of an Anthropomorphic Hand. Proceeding of the 7th International Symposium on Artificial Intelligence, Robotics and Automation in Space: I-SAIRAS 2003. http://robotics.estec.esa.int/i-SAIRAS/isairas2003/data/pdf/AS13paper.pdf

[117] Waldron, V. R. (2009). Encyclopedia of human relationships. In H. T. Reis, & S. Sprecher (Eds.), Apologies, 1, vol. 3 (1st ed.). (pp. 98–100). Thousand Oaks, CA: Sage Publishing Inc.

[118] Wiegand, G., Schmidmaier, M., Weber, T., Liu, Y., and Hussmann, H. (2019). "I drive - you trust: explaining driving behavior of autonomous cars," in Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, paper no. LBW0163 (Glasgow). doi: 10.1145/3290607.3312817

[119] Wogalter, M. S., and Vigilante, W. J. (2006). "Attention switch and maintenance," in Handbook of Warnings, ed M. SWogalter (Mahwah, NJ: Lawrence Erlbaum Associates).

[120] Xinyi Zhang, Sun Kyong Lee, Whani Kim, Sowon Hahn, "Sorry, it was my fault": Repairing trust in human robot interactions, International Journal of Human-Computer Studies, Volume 175, 2023, 103031, ISSN 1071-5819, https://doi.org/10.1016/j.ijhcs.2023.103031.

Appendix

Contents

Research Plan Initial Research

OVERARCHING
RESEARCH OBJECTIVES

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

MAP CURRENT EXPERIENCE
Map current experience with the system and the emotions of the user

Reason: To be able to design a better experience, compare it to the existing one in evaluation and get a better understanding of where the issues are.

FIND PAIN-POINT AREAS
Find areas that elicit the negative experience towards the planning errors. Find origin of negative emotions. Find reasons for Planning error (human fault or system etc.)

Reason:
- To find opportunities for improvement
- Understand where one could intervene in the system to fix the issue

UNDERSTAND ROBOT'S RELATION/PERCEPTION
Learn about the robot human relationship

Reason: A system that the user trust will be used more often, with More confidence and have a positive impact on the overall experience

UNDERSTAND TARGET GROUP
Understand who the users are, their goals, motivations, problems

Reason:
- To make it more human tailored and fit the design choices to their needs (Understand what the users wish for in the interface, understand how they work and how a typical day looks like)
- Find opportunities where one could intervene in the system to fix the issue
- To design with preconditions in mind, potentially be able to make the transition smoother, keep knowledge alive and reduce need to learn system over

MAP THE ENVIRONMENT
Understand conditions (mentally, physically, context, workplace)

Reason:
- To design an intervention that works under the given extreme conditions and be aware of limitations.
- To ensure feasibility of the final design
- Get a better understanding of the target user (certain behaviours are context dependent and need to be understood to design a fitting intervention)

TRANSFER KNOWLEDGE
Understand what is out there already?
Find opportunities/context with similar problems
Find Inspiration (Exhibition, online, other contexts (like gaming), speculative design)

Reason:
- To increase the design space
- To improve the ideation phase
- To improve understanding of problem from another perspective

1

OVERARCHING
RESEARCH PLAN

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTRODUCTION
Hello, my name is Lilly! During this Session, I will ask you to tell me some stories that you experienced about your experience with interacting with the robot and problems that you came across. We may even find workable improvements on the go.

This interview is open-ended. If you feel that some parts need more research or there is something you would like to mention, please feel free to do so. I am not here to prove something, I am only here to learn from you and your experience.

SIGN AGREEMENTS
Before we start:
Is it ok, if I record this Interview/Test: ☐ Audio ☐ Video ☐ Both ☐ None

☐ Give User confidentiality agreements

SESSION SETUP
1. Introduction

2. Confidentiality agreement

3. First interview Questions

4. If Training familiar: UX survey if time permits

5. Scenario 1 - Navigation
   5.1 look at Scenario visual or do scenario in interface
   5.2 Evaluate emotions in session with PrEmo and Desirability set if time permits
   5.3 UX Survey Navigation if time permits

6. Scenario 2 - Success
   6.1 look at Scenario visual or do scenario in interface
   6.2 Evaluate emotions in session with PrEmo and Desirability set if time permits
   6.3 UX Survey Success if time permits

7. Scenario 3 - Error visuals
   7.1 look at Scenario visuals
   7.2 Evaluate emotions in session with PrEmo and Desirability set if time permits
   7.3 UX Survey Error if time permits

8. Continue Interview until robot section
   8.1 Godspeed survey robot

9. Continue Interview

10. If more time available ask requirement questions

2

ASTRONAUTS
DISCUSSION GUIDE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTERVIEW QUESTIONS
Could you please tell us a little bit about yourself?

- Your position at DLR
- Your relation to the Robot
- Prior use of robots? How many usually?
- You tasks

Have you interacted with the Robot and the current system already?

- **YES**: How was it? Please describe the experience you had. How would you describe the interaction?
- **NO**: How do you image the interaction with the robot?

Could you describe the usual process when interacting with the robot?

- Are you familiar with the prior training and if yes, how would you describe it?
- If not, how do you imagine the training, what do you think should take place during those 15 minutes?

MOVE TO:
If Training familiar +time then 1. Training UX survey
Else:
1. Scenario 1-2 + Premo and Desirability + Error
2. UX Surveys after each Scenario
3. Back to Interview

What Do you think is the main issue with planning errors and user?

- Is it the lack of feedback,
- The interface design,
- Something lacking in the introduction presentation?
- The camera angle,
- The lack of general relativeness to the robot
- Previous interactions etc..

When/where is a better understanding of the robots planning needed in your opinion?

Why do you need to know why the planning has an error?

What do you need to know to be able to deal with planning errors?

Why do you need to use the robot + system in general? Can you give a use example when you would use it?

How would the ideal interaction look like, without constraints? If you could have the control be however you want it.

How Do you perceive the robot when interacting with it?

- Partner, tool, hindrance, help?
- How is your trust level towards the robot?
- Would you describe it as a cooperation and if yes does that help?

MOVE TO:
Godspeed Form for Robot

How would you describe yourself?

- What are your goals (why decide to be an astronaut)? motivations, problems, wishes?
- How do you act in this context? Strict, ordered, passionate, empathetic?
- What do you like/hate, hobbies, friendships...

How does your typical day look like, what happens?

- Many switches between tasks?
- Lots of free time?
- Demanding?

How do you think are you different from everyday people?

When you interact with the robot what usually works well, what goes wrong?

- What Do you do before/after the testing?

MORE TIME?

What are the limitations of this project (due to ISS, Robot, or Planet) in your opinion

What obstacles may arise in your experience?

4

STAFF MEMBERS
DISCUSSION GUIDE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTERVIEW QUESTIONS
Could you please tell us a little bit about yourself?

- Your position at DLR
- Your relation to the Robot
- Other experience with robots?
- You tasks

Have you interacted with the Robot and the current system already?

- **YES**: How was it? Please describe the experience you had. How would you describe the interaction?
- **NO**: How do you image the interaction with the robot?

Could you describe the usual process when interacting with the robot?

- Are you familiar with the prior training and if yes, how would you describe it?
- If not, how do you imagine the training, what do you think should take place during those 15 minutes?

MOVE TO:
If Training familiar +time then 1. Training UX survey
Else:
1. Scenario 1-2 + Premo and Desirability + Error
2. UX Surveys after each Scenario
3. Back to Interview

What Do you think is the main issue with planning errors and user?

- Is it the lack of feedback,
- The interface design,
- Something lacking in the introduction presentation?
- The camera angle,
- The lack of general relativeness to the robot
- Previous interactions etc..

When/where is a better understanding of the robots planning needed in your opinion?

Why does the astronaut need to know why the planning has an error?

What does the astronaut (in your opinion) need to know to be able to deal with planning errors?

Why does the astronaut use the robot + system in general? Can you give a use example?

How would the ideal interaction look like, without constraints? If you could have the control be however you want it.

How Do you perceive the robot when interacting with it?

- Partner, tool, hindrance, help?
- How is your trust level towards the robot?
- Would you describe it as a cooperation and if yes does that help?

MOVE TO:
Godspeed Form for Robot

How would you describe an astronauts personality?

- What are their goals? Motivations, problems, wishes?
- How do they act? Strict, ordered, passionate, empathetic?

How do you think/know their typical day looks like, what happens?

- Many switches between tasks?
- Lots of free time?
- Demanding?

How do you think are they different from everyday people?

When they interact with the robot what usually works well, what goes wrong?

- What Do they do before/after the testing?

MORE TIME?

What are the limitations of this project (due to ISS, Robot, or Planet)

What obstacles may arise

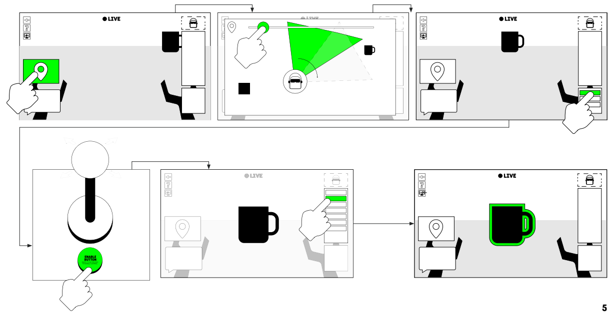
3

EXPERIENCE AND EMOTIONS
SCENARIO 1 – NAVIGATE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Testing
Below you are presented with a scenario to navigate the robot to an object of interest. Please talk aloud while going through the scenario.

Cognitive walk-through
Please explain what you would usually do in this context, look at the images an map your experience with the provided tools

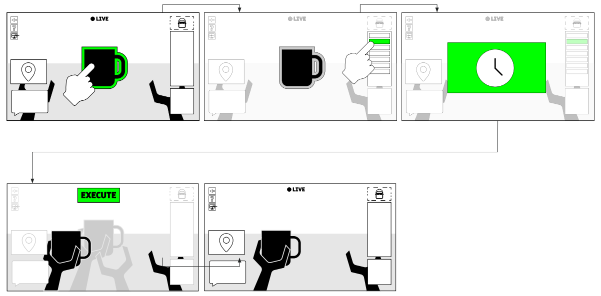


EXPERIENCE AND EMOTIONS
SCENARIO 2 – SUCCESS

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

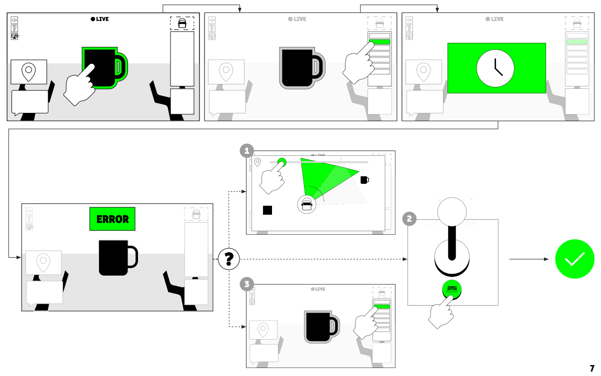
Testing
Below you are presented with a scenario where you have to handle an object in the environment. Please talk aloud while going through the scenario.

Cognitive walk-through
Please explain what you would usually do in this context, look at the images and map your experience with the provided tools



EXPERIENCE AND EMOTIONS
SCENARIO 3 – ERROR

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

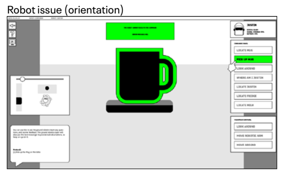


OVERARCHING
ERROR ORIGINS

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Below you are provided 4 scenarios in which a planning error occurred (= The robot was not able to find a solution to the commanded task). Please look at each scenario and answer the following two questions:
What do you think caused the planning error given the scenario you are provided on the left?
Imagine you are using the robot and end up in the scenario, what would you do?

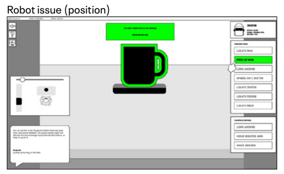
Robot issue (orientation)



What caused the problem?

What would you do?

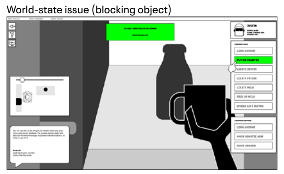
Robot issue (position)



What caused the problem?

What would you do?

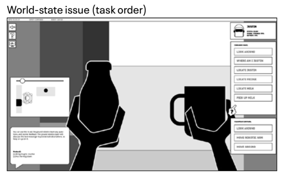
World-state issue (blocking object)



What caused the problem?

What would you do?

World-state issue (task order)

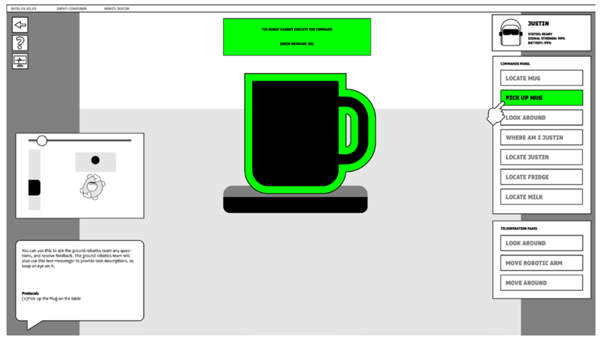


What caused the problem?

What would you do?

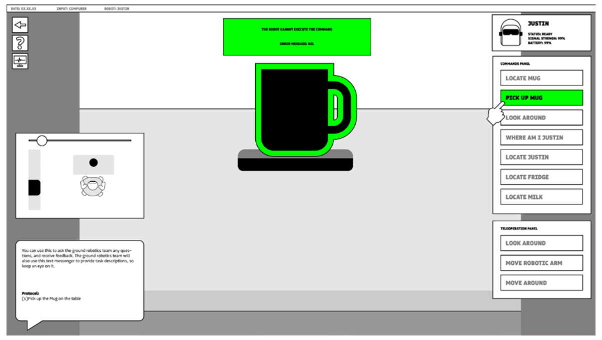
UNDERSTANDING
SCENARIO 3 – ERROR 1

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



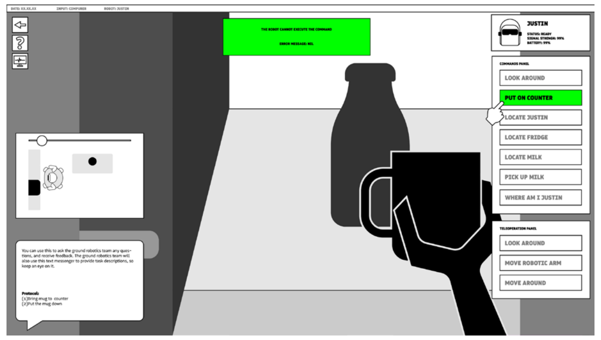
UNDERSTANDING
SCENARIO 3 – ERROR 2

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



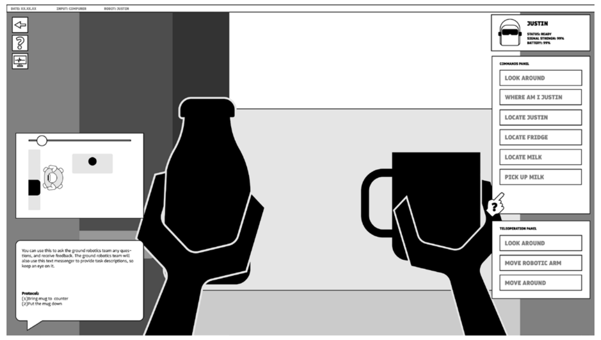
UNDERSTANDING
SCENARIO 3 – ERROR 3

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



UNDERSTANDING
SCENARIO 3 – ERROR 4

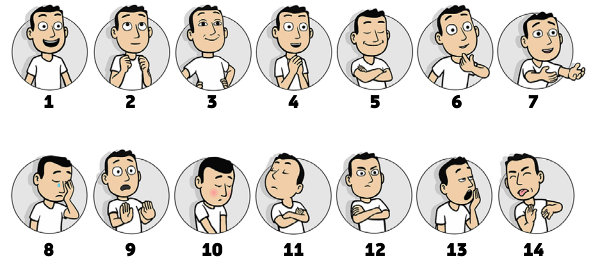
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



OVERARCHING
5X PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

PrEmo measures distinct (pleasant and unpleasant) emotions in a non-verbal manner that is validated cross-culturally. Please look at the images below and use the numbers to write into the scenarios and indicate how you feel about certain aspect in them. Feel free to use as many as you need, there is no limit



OVERARCHING
5X PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

OVERARCHING
DESIRABILITY

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

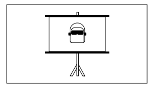
ANNOYING	BORING	TOO TECHNICAL	COLLABORATIVE	EFFICIENT
CONFUSING	DIFFICULT	UNCONTROLLABLE	EASY TO USE	FLEXIBLE
FRUSTRATING	DISCONNECTED	PREDICTABLE	CONTROLLABLE	HELPFUL
IMPERSONAL	GETS IN THE WAY	STABLE	INTUITIVE	ORGANIZED
STRESSFUL	TIME CONSUMING	RELIABLE	PERSONAL	

OPEN

EXPERIENCE EVALUATION
TRAINING

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**



	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

EXPERIENCE EVALUATION
NAVIGATION

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**

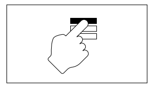
	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

EXPERIENCE EVALUATION

SUCCESS

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**



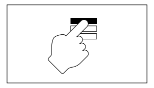
	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

EXPERIENCE EVALUATION

ERROR

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**



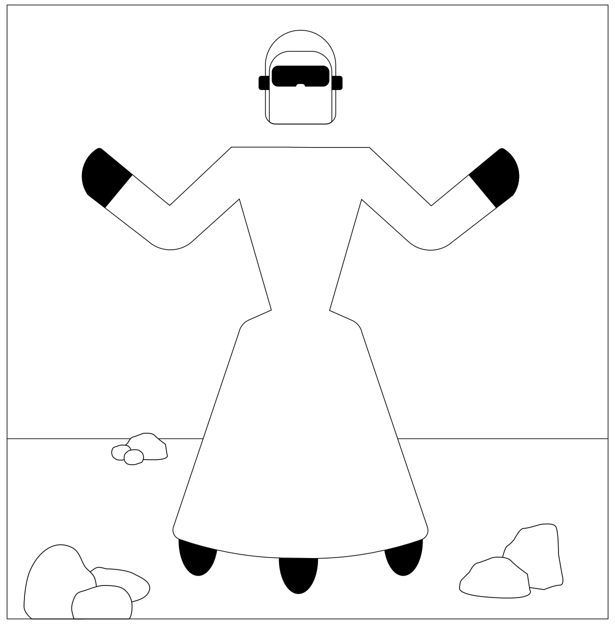
	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

PERCEPTION

THE ROBOT

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Please think about your interaction with the robot. What do you associate with it? Especially when you are controlling it. How does it make you feel? Use the Form on the left to rate your experience with the robot by making a tick in each row closer to the word that fits better. If you feel neither of the words tick 4.



Please rate your impression of Justin on these scales (circle the best fit):

Fake	1	2	3	4	5	Natural
Machinelike	1	2	3	4	5	Human-like
Unconscious	1	2	3	4	5	Conscious
Artificial	1	2	3	4	5	Lifelike
Moving rigidly	1	2	3	4	5	Moving elegantly
Dead	1	2	3	4	5	Alive
Stagnant	1	2	3	4	5	Lively
Mechanical	1	2	3	4	5	Organic
Artificial	1	2	3	4	5	Lifelike
Inert	1	2	3	4	5	Interactive
Apathetic	1	2	3	4	5	Responsive
Dislike	1	2	3	4	5	Like
Unfriendly	1	2	3	4	5	Friendly
Unkind	1	2	3	4	5	Kind
Unpleasant	1	2	3	4	5	Pleasant
Awful	1	2	3	4	5	Nice
Incompetent	1	2	3	4	5	Competent
Ignorant	1	2	3	4	5	Knowledgeable
Irresponsible	1	2	3	4	5	Responsible
Unintelligent	1	2	3	4	5	Intelligent
Foolish	1	2	3	4	5	Sensible
Anxious	1	2	3	4	5	Relaxed
Agitated	1	2	3	4	5	Calm
Quiescent	1	2	3	4	5	Surprised

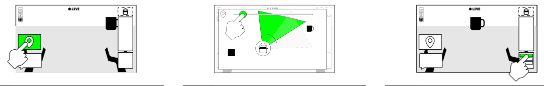
Please indicate what Justin is for you when interacting with it(Cross the word in the table that fits best). If nothing fits, add your description under "Other":

Partner	Co-Worker	Friend	Teacher	Tool	Extension of oneself	Other

SCENARIO 1 – NAVIGATION

PREMO

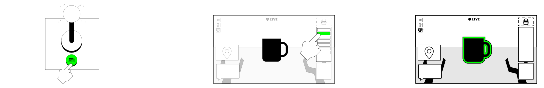
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



PREMO	WHERE

PREMO	WHERE

PREMO	WHERE



PREMO	WHERE

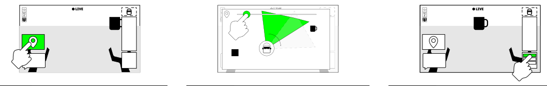
PREMO	WHERE

PREMO	WHERE

SCENARIO 1 – NAVIGATION

DESIRABILITY

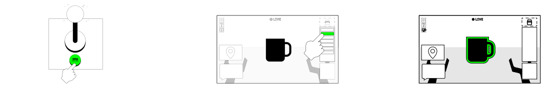
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



DESIRABILITY	WHERE

DESIRABILITY	WHERE

DESIRABILITY	WHERE



DESIRABILITY	WHERE

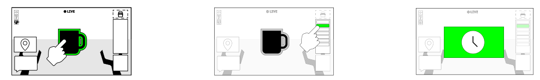
DESIRABILITY	WHERE

DESIRABILITY	WHERE

SCENARIO 2 – SUCCESS

PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



PREMO	WHERE

PREMO	WHERE

PREMO	WHERE

PREMO	WHERE

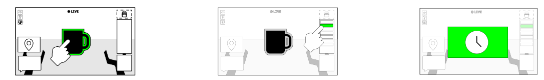
PREMO	WHERE

PREMO	WHERE

SCENARIO 1 – SUCCESS

DESIRABILITY

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



DESIRABILITY	WHERE

DESIRABILITY	WHERE

DESIRABILITY	WHERE

DESIRABILITY	WHERE

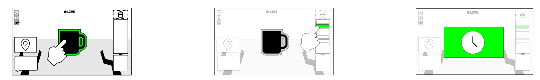
DESIRABILITY	WHERE

DESIRABILITY	WHERE

SCENARIO 2 – ERROR

PREMO

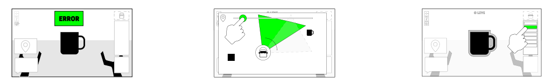
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



PREMO	WHERE

PREMO	WHERE

PREMO	WHERE



PREMO	WHERE

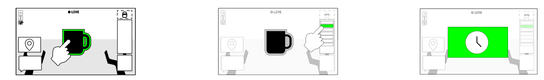
PREMO	WHERE

PREMO	WHERE

SCENARIO 1 – ERROR

DESIRABILITY

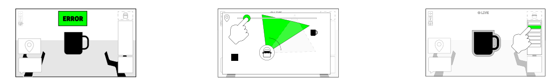
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



DESIRABILITY	WHERE

DESIRABILITY	WHERE

DESIRABILITY	WHERE



DESIRABILITY	WHERE

DESIRABILITY	WHERE

DESIRABILITY	WHERE

Raw Data Inital User Study

GODSPEED

Antromorphism					Animacy						Likeability					Perceived Intelligence					Perceived Savety			Category	Person	Antromorphism	Animacy	Likeability	Perceived Intelligence	Perceived Savety
4	4	2	4	3	4	2	2	3	2	2	5	5	5	5	5	2	4	2	2	3	4	4	4	Other	P09	3,4	2,5	5	2,6	4
5	2	1	1	4	3	4	1	3	4	5	4	5	4	4	5	4	4	5	4	5	5	5	1	Tool	P02	2,6	3,3333333	4,4	4,4	3,6666667
2	2	1	2	5	3	4	1	2	4	2	5	4	4	5	4	4	4	4	4	3	4	4	3	Partner	P04	2,4	2,6666667	4,4	3,8	3,6666667
3	4	4	2	2	3	3	2	2	3	4	4	4	3	4	4	4	3	3	4	3	3	4	3	Tool	P03	3	2,8333333	3,8	3,4	3,3333333
5	5	4	4	4	5	5	4	4	4	5	5	5	5	5	5	4	4	4	4	4	5	5	2	Friend	P08	4,4	4,5	5	4	4
3	4	2	1	2	3	3	1	1	3	4	2	4	3	4	4	4	3	4	3	4	3	4	3	Tool	P01	2,4	2,5	3,4	3,6	3,3333333
4	4	2	2	5	3	4	4	4	4	4	5	5	4	4	4	3	3	3	3	3	4	4	3	Co-worker	P05	3,4	3,8333333	4,4	3	3,6666667
3	2	1	2	2	3	1	2	2	2	2	5	5	4	4	5	4	4	3	2	3	5	4	2	Tool	P07	2	2	4,6	3,2	3,6666667
2	4	2	1	2	4	1	1	2	5	4	5	5	4	4	5	4	4	5	3	4	5	5	3	Tool	P06	2,2	2,8333333	4,6	4	4,3333333
3	3	2	2	3	3	3	2	3	3	4	4	5	4	4	5	4	4	4	3	4	4,2	4,3	2,7							
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,8	0,5	0,9							
M	3						3						4					4				3,7				2,8666667	3	4,4	3,555555556	3,7407407
SD	1						1						1					1				1,1				0,7615773	0,7682954	0,5196152	0,563717818	0,3239418

PREMO

Scenario Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Navigation	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
click on Map	3		1			3						2		
Move Head	1	1	2		1	3								1
Select Navigation	2		4		1			1					1	
Navigate to Object (multiple)	2		5	2	2					1			1	
locate Object	4	1		2		1							2	
Object located	4		2					1						
	16	2	14	4	4	7	0	2	0	1	0	2	4	1

Scenario Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Success	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Select Object	6	2					1							
Command "Pick up"	4	2	1	1										
Wait for Planning	2	1				2	1					2	1	1
Wait/Watch Execution	5		2	1	1				2			1	1	
Have mug in robot hand	6		1		1									
	23	5	4	2	2	2	2	0	2	0	0	3	2	1

Scenario Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Error	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Select Object	6	1				1								
Command "Pick up"	4	1		1		2								
Wait for Planning	2	1				1						2	2	
Encounter Planning Error						2			2			4	1	1
Deal with Problem	1	4			1	4	1	1				2	1	1
Fixed error	4		3											
	7	5	3	0	1	7	1	1	2	0	0	8	4	2

UEQ – ERROR

Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
3	6	5	4	6	5	5	4	7	4	3	2	3	4	6	5	5	6	2	1	6	2	5	4	4	5	P09
2	3	1	3	2	6	7	2	5	2	5	3	3	5	6	4	2	2	3	4	5	4	3	3	3	6	P02
1	1	6	7	7	1	1	1	7	7	1	7	1	1	1	1	5	7	7	1	7	1	5	7	7	1	P04
3	4	4	6	2	4	5	6	6	6	6	3	3	3	3	4	3	4	4	2	5	2	5	5	5	3	P03
4	3	2	4	4	4	6	3	4	2	4	4	3	4	6	4	3	4	2	4	5	4	4	3	2	5	P08
4	4	3	2	3	5	5	5	5	4	4	3	4	5	4	5	3	4	4	4	3	5	3	3	3	4	P01
2	2	4	6	5	4	4	4	6	4	3	5	3	3	4	2	4	6	6	2	6	2	4	5	5	4	P05
2	2	4	6	5	2	4	3	4	5	3	5	4	2	3	2	4	6	5	3	6	3	4	4	4	3	P07
1	1	7	6	6	1	5	1	7	6	1	7	2	3	3	2	2	7	6	1	6	3	4	5	4	4	P06

UEQ – NAVIGATION

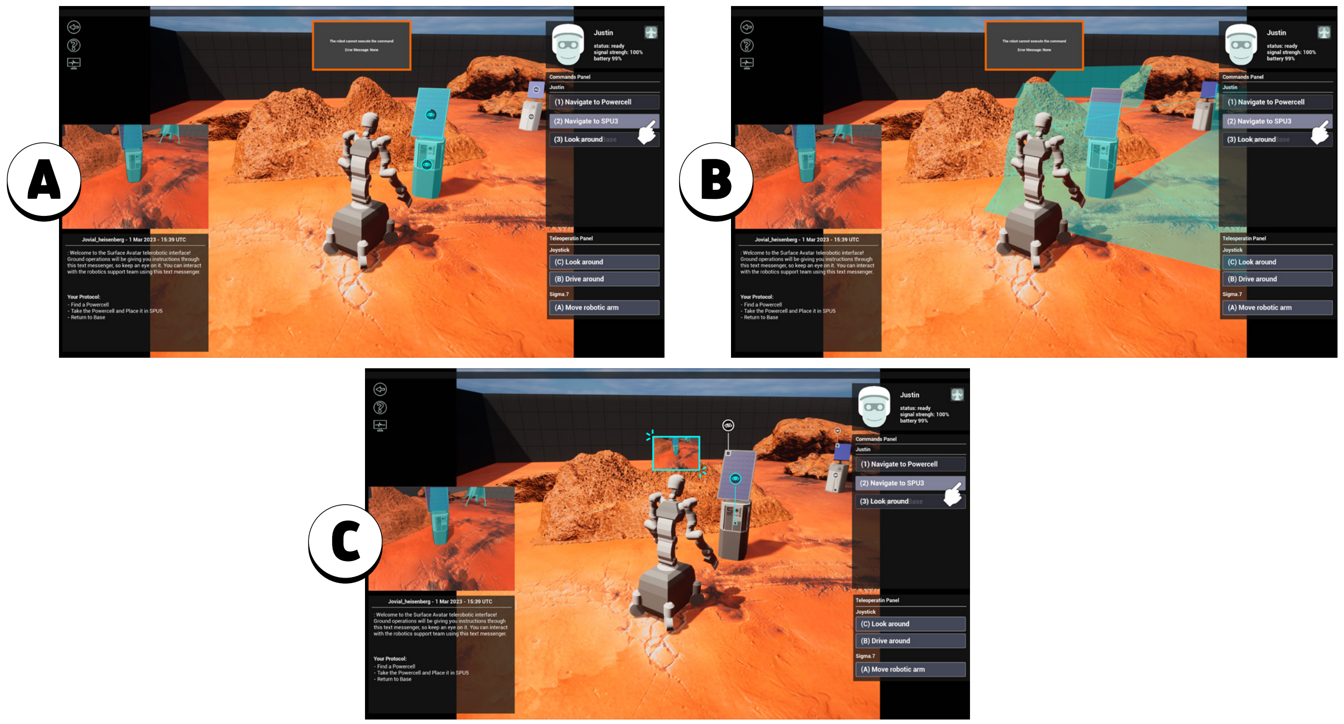
Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
5	5	7	2	7	2	2	7	2	6	5	2	6	5	2	4	2	4	1	4	2	6	2	4	4	2	P09
6	5	3	2	3	5	7	6	4	2	5	1	6	5	5	6	3	1	2	5	3	6	3	4	3	6	P02
6	6	6	6	4	4	2	5	6	6	4	4	7	4	2	4	2	4	5	5	1	6	4	4	3	4	P04
6	6	4	1	2	6	6	6	6	6	6	2	6	6	4	6	2	2	2	5	2	6	2	2	2	3	P03
7	7	3	2	2	4	6	5	4	3	6	2	6	5	6	6	4	2	2	6	3	6	2	2	1	6	P08
6	7	4	2	1	5	5	7	3	4	6	2	6	6	2	6	2	2	2	6	2	6	2	2	2	3	P01
6	6	4	2	2	6	4	6	5	6	4	2	6	6	4	6	1	2	1	6	2	6	2	2	2	4	P05
5	6	5	1	4	5	5	6	5	3	6	2	6	5	5	6	2	4	3	5	3	6	2	3	4	6	P07
6	6	3	1	1	5	3	5	4	3	6	2	6	6	3	5	1	3	2	5	2	5	3	3	2	5	P06

UEQ – OBJECT INTERACTION

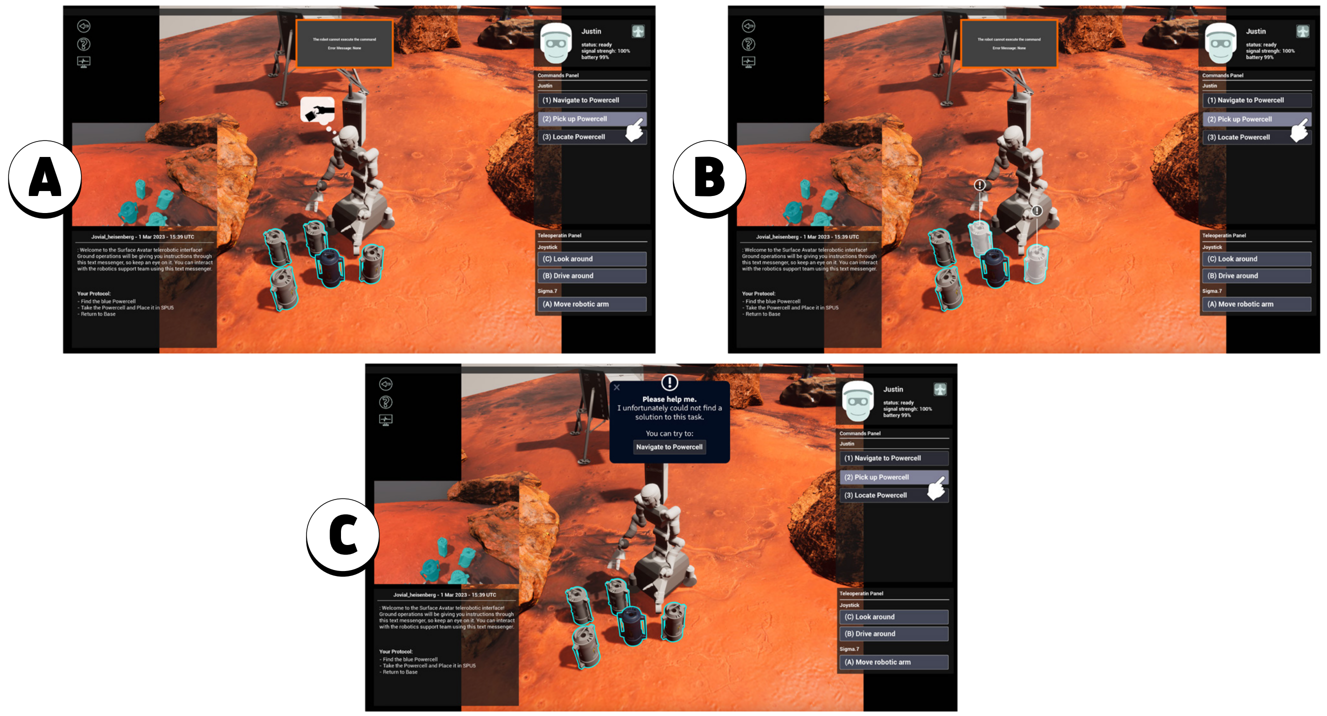
Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
6	6	6	2	6	2	2	7	2	6	6	1	7	4	1	6	1	4	1	6	5	5	3	4	3	3	P09
2	2	3	2	2	5	7	6	5	3	5	2	6	5	6	5	2	2	3	4	3	6	3	4	2	6	P02
7	7	3	3	4	5	6	2	4	5	4	3	6	5	3	5	5	3	4	3	2	5	3	4	4	4	P04
6	6	2	2	2	6	6	6	2	5	6	2	6	6	5	6	2	2	2	6	2	6	2	2	2	5	P03
7	7	2	2	2	6	7	5	4	2	6	2	6	6	6	6	4	2	2	6	3	6	2	2	2	6	P08
5	6	5	1	1	4	5	3	4	3	6	2	6	6	5	5	4	3	2	6	2	6	2	2	2	5	P01
6	6	2	2	2	6	6	6	6	2	6	2	6	6	5	6	2	2	2	6	1	6	2	2	2	6	P05
5	5	4	2	3	5	5	4	5	3	6	3	5	5	4	5	2	3	2	5	2	6	3	3	4	5	P07
7	6	4	2	2	5	5	6	5	3	6	2	6	6	6	6	1	1	1	7	1	7	1	2	1	6	P06

Game Element Overview

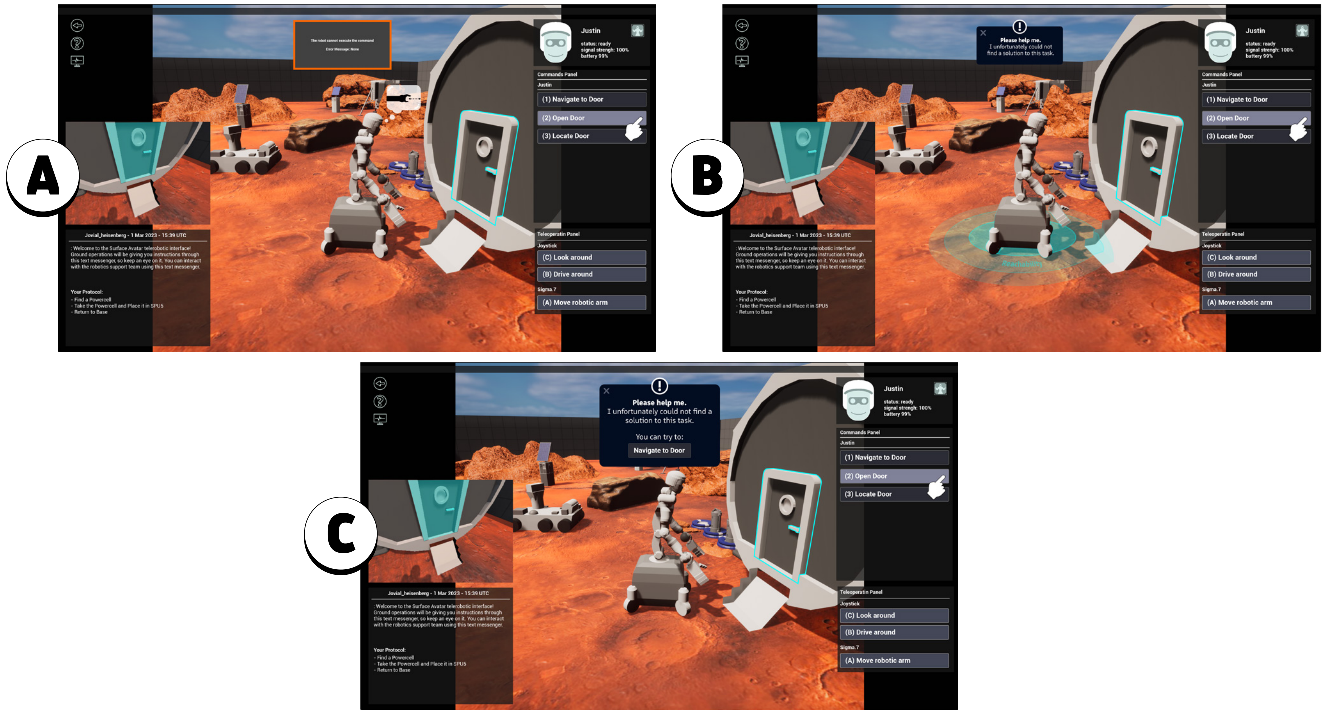
LOCALIZE – ERROR



COLLISION – ERROR

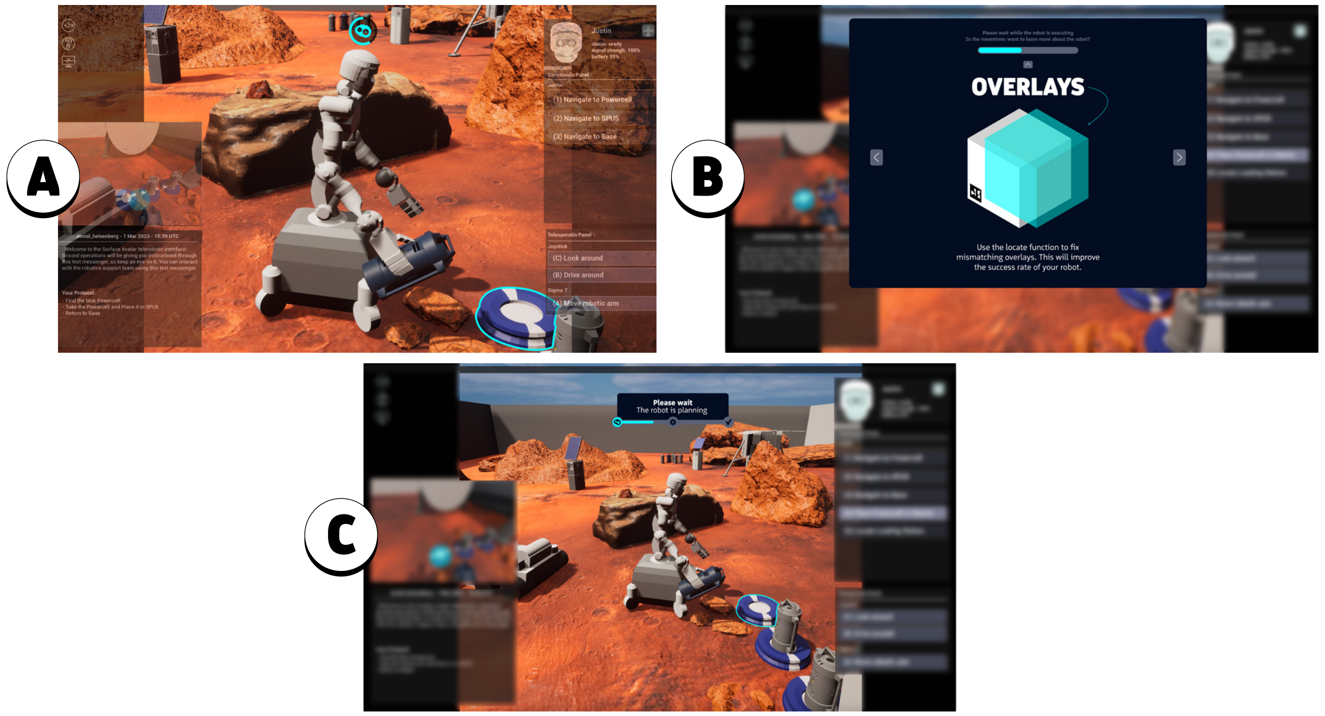


REACHABILITY – ERROR

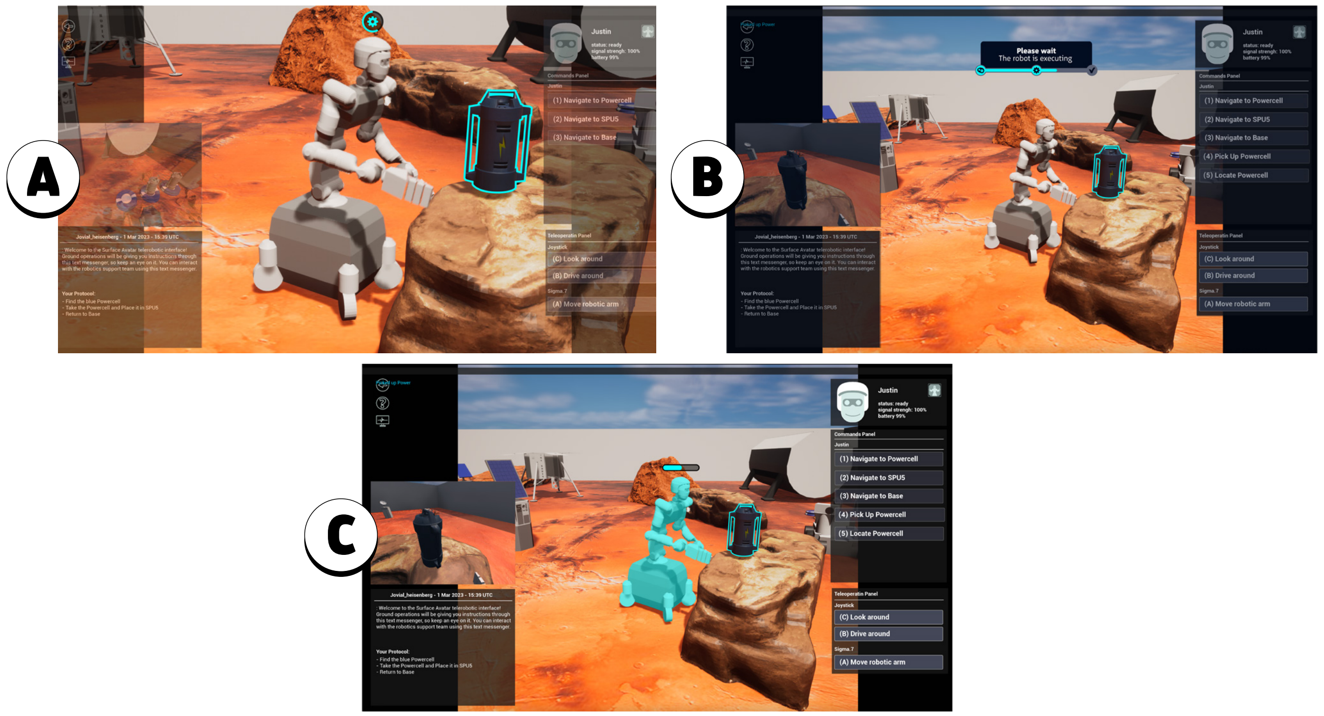


ORIENTATION – ERROR

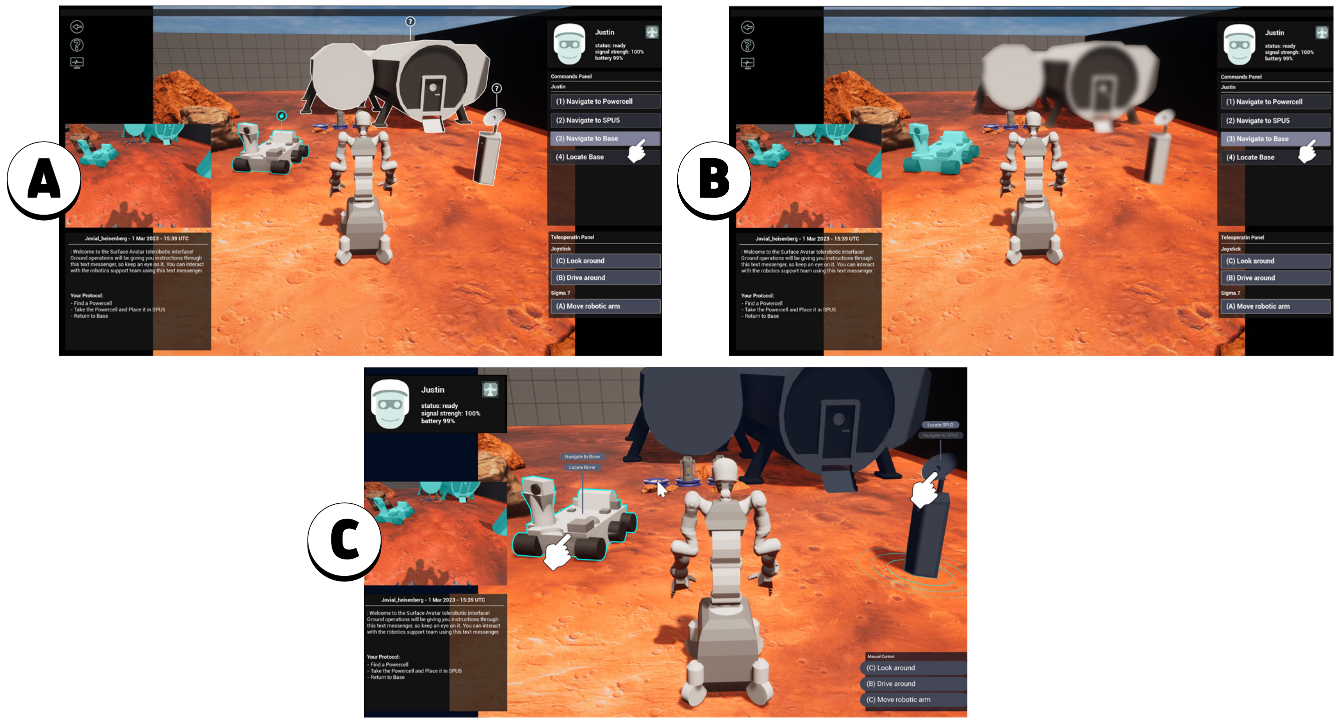
PLANNING WAIT



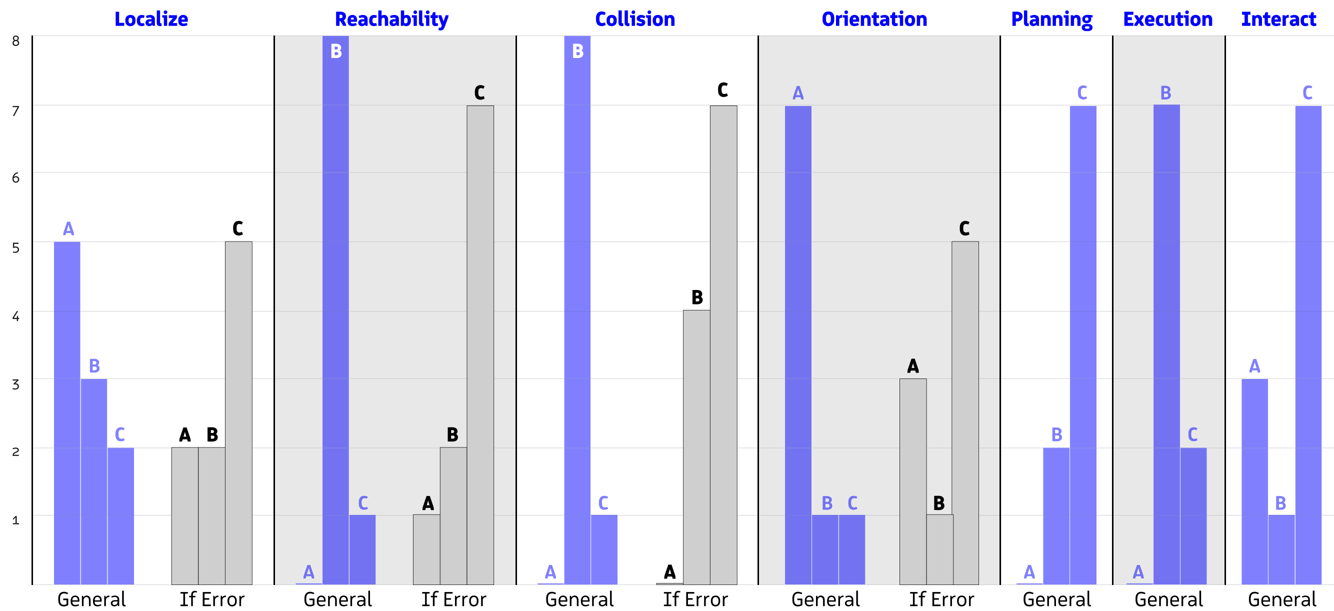
EXECUTION WAIT



INTERACTABLE?



RESULTS - FEATURE PREFERNCE



Plan Prototype Testing

OVERARCHING
RESEARCH PLAN

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTRODUCTION
Hello, my name is Lilly! During this Session, I will ask you to interact with two prototypes, complete questionnaires, and share your experience, persistent difficulties, and potential improvements.

This interview is open-ended. If you feel that some parts need more research or there is something you would like to mention, please feel free to do so. I am not here to prove something, I am only here to learn from you and your experience.

SIGN AGREEMENTS
Before we start:
Is it ok, if I record this Interview/Test: ☐ Audio ☐ Video ☐ Both ☐ None

☐ Give User confidentiality agreements

SESSION SETUP
1. Introduction
2. Confidentiality agreement
3. Prototype 1: 3rd Person
   3.1 Talk out loud
   3.2 After each error PrEmo
4. NASA TLX 1
5. Interview
   5.1 General experience (UEQ?)
   5.2 Persistent difficulties
   5.3 Improvement suggestions
6. ABC Game elements Questionnaire
7. Prototype 2: Typical issues debug Page
   7.1 Talk out loud
   7.2 After each error PrEmo
8. NASA TLX 2
9. Interview
   9.1 General experience (UEQ?)
   9.2 Persistent difficulties
   9.3 Improvement suggestions
10. Comparison Study?

2

PROTOTYPE 1
3RD PERSON

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.05 - 26.05

In the first prototype on my laptop, you will have the opportunity to interact with a simulation of the robot's UI system. Please note that the interaction may not be 1:1 as you are accustomed to. The simulation provides a 3rd person perspective, which can be switched to 1st person by pressing 'L' (and again to go back to 3rd person). Instead of using the mouse to click on buttons in the UI, please use the number keys indicated next to the buttons in the interface. This allows you to control the robot while using the mouse to change the camera view (the final prototype will allow you to use the mouse again). You can use the WASD keys to manually control the robot and C to switch to the head. In this scenario, you do not need to worry about localisation as the robot will perfectly locate all objects in the environment. I will function as GO throughout the session, and the protocol of what you need to do has already been sent to you in the chat. (Give them overview of UI specs)

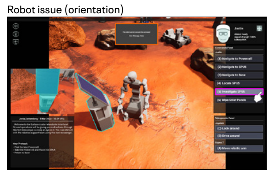
Now, I kindly ask you to engage in a 'Talk out loud' technique while interacting with the prototype. Please verbalize your thoughts, explaining why you are taking certain actions and expressing any challenges or observations you encounter. Additionally, I will provide you with PrEmos, a tool you might be familiar with from the last session. The PrEmos sheet will be given to you, and you can use it to indicate how you feel during the interaction. Please refer to the numbers on the sheet corresponding to the facial expressions that best represent your emotions. Remember, the numbers do not indicate a rating; simply choose the one that accurately reflects your feelings.

Robot issue (reachability)

What caused the problem?

What would you do?

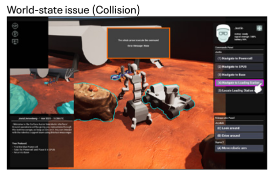
Robot issue (orientation)



What caused the problem?

What would you do?

World-state issue (Collision)



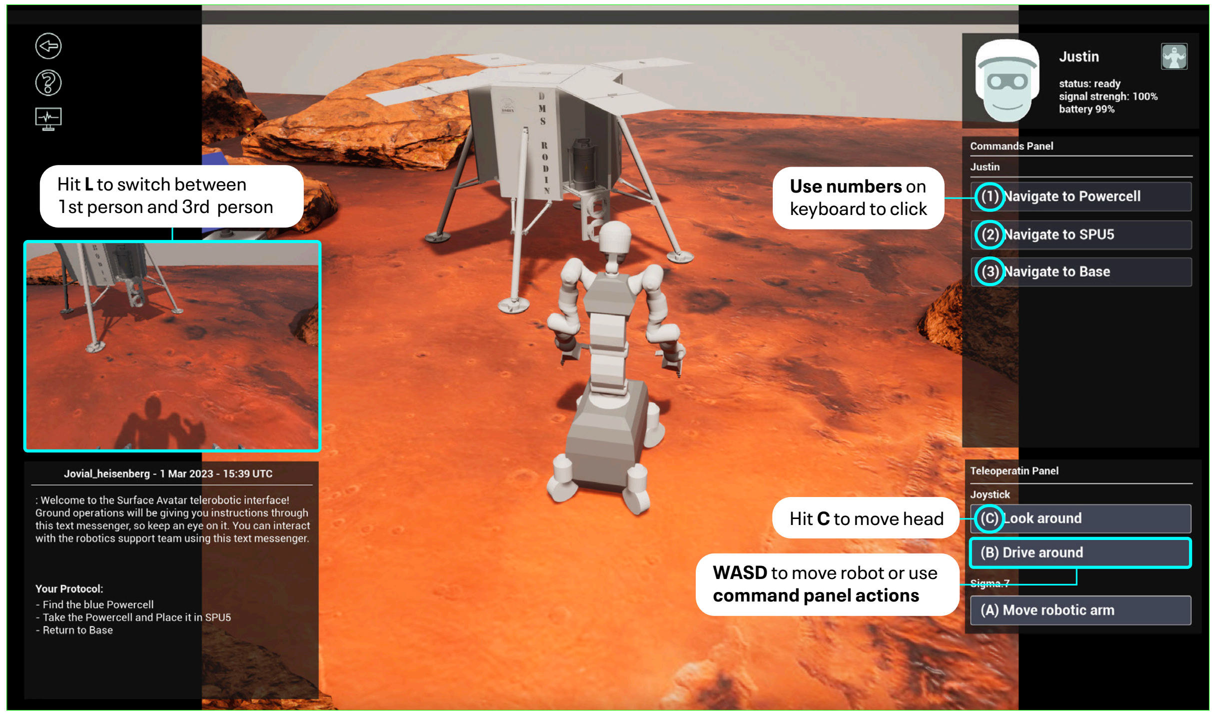
What caused the problem?

What would you do?

3

PROTOTYPE – 1
UI CONTROLS

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03



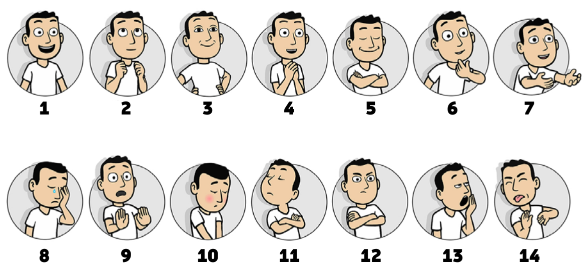
There is no need to localize

4

PREMO
HOW DO YOU FEEL ABOUT THIS SITUATION?

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

PrEmo measures distinct (pleasant and unpleasant) emotions in a non-verbal manner that is validated cross-culturally. Please look at the images below and use the numbers to write into the scenarios and indicate how you feel about certain aspect in them. Feel free to use as many as you need, there is no limit



5

NASA TASK LOAD INDEX
3RD PERSON PROTOTYPE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.
Please try to answer spontaneously!

Mental Demand — How mentally demanding was the task?
Very Low — Very High

Physical Demand — How physically demanding was the task?
Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?
Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?
Very Low — Very High

6

EXPERIENCE EVALUATION
3RD PERSON PROTOTYPE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

7

FEATURE QUESTIONNAIRE
3RD PERSON PROTOTYPE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

You will be presented with different visualizations of additional features for the system. Please compare and select the one you prefer based on your personal preference and explain your choice.

(1) **Localize**: Which of the three presented feature visualisations would you prefer?

To indicate how well the object is located
- ◯ **A:** Show localize percentage
- ◯ **B:** Project view
- ◯ **C:** Highlight April Tags

To show up in case of a localize error
- ◯ **A:** Show localize percentage
- ◯ **B:** Project view
- ◯ **C:** Highlight April Tags

(2) **Reachability**: Which of the three presented feature visualisations would you prefer?

To indicate how reachable an object is
- ◯ **A:** Thinking bubble
- ◯ **B:** Reachability area of robot
- ◯ **C:** Ask for help

To show up in case of a reachability error
- ◯ **A:** Thinking bubble
- ◯ **B:** Reachability area of robot
- ◯ **C:** Ask for help

(3) **Collision**: Which of the three presented feature visualisations would you prefer?

To indicate potential collision
- ◯ **A:** Thinking bubble
- ◯ **B:** Highlight colliding objects/body parts
- ◯ **C:** Ask for help

To show up in case of a collision error
- ◯ **A:** Thinking bubble
- ◯ **B:** Highlight colliding objects/body parts
- ◯ **C:** Ask for help

(4) **Orientation**: Which of the three presented feature visualisations would you prefer?

To indicate how well the robot is oriented
- ◯ **A:** Highlight unaligned body parts
- ◯ **B:** Thinking bubble
- ◯ **C:** Ask for help

To show up in case of a orientation error
- ◯ **A:** Highlight unaligned body parts
- ◯ **B:** Thinking bubble
- ◯ **C:** Ask for help

(5) **Planning wait**: Which of the three presented feature visualisations would you prefer?

To show during the planning time
- ◯ **A:** Loading circle with icon
- ◯ **B:** Loading screen with hints
- ◯ **C:** Progress bar with steps

(6) **Execution wait**: Which of the three presented feature visualisations would you prefer?

To show during the execution time
- ◯ **A:** Loading circle with icon
- ◯ **B:** Progress bar with steps
- ◯ **C:** Highlight robot during execution + loading beam

(7) **Interact-able**: Which of the three presented feature visualisations would you prefer?

To indicate an object as interact-able
- ◯ **A:** Difference in colour + icon
- ◯ **B:** Blur out interact-able but not localized
- ◯ **C:** Show interaction on object + cursor change

9

PROTOTYPE – 1
DISCUSSION GUIDE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTERVIEW QUESTIONS

Could you please tell us a little bit about yourself?

- Your position at DLR
- Other experience with robots?
- You tasks
- Have you interacted with the Robot and the current system already?

Were there any specific moments during the interaction that stood out to you positively or negatively? Can you describe those moments in more detail?

- Did you encounter any challenges or frustrations while interacting with the prototype? If yes, could you elaborate on those?
- Were there any aspects of the prototype that you found particularly intuitive or user-friendly?

Did you find the 3rd person perspective helpful in mitigating errors or providing guidance during the interaction? If so, in what specific ways did it assist you?

- Were there any instances where you felt the need for additional help or guidance while using the 3rd person perspective? If yes, can you describe those situations and how you think they could be addressed?
- Can you provide examples of moments where the perspective enhanced or hindered your understanding?
- Can you recall any specific instances where the 3rd person perspective guided you to perform actions without the need for external help or instructions? How did you feel about this level of guidance and independence?

Did the 3rd person perspective facilitate an easy and low cognitive load interaction? Or did it require additional mental effort or complexity? Please explain your experience.

How engaging did you find the interaction with the 3rd person perspective?

- Did it hold your attention and maintain your interest throughout the session?
- Were there any aspects that could be improved to enhance engagement?

Were there any specific features or functionalities of the prototype that stood out to you? If so, why?

Can you think of any additional features, functionalities, or improvements that you believe would enhance the prototype's error handling capabilities?

Is there anything else you would like to share about your experience with the prototype or any other insights you believe would be valuable for us to know?

PROTOTYPE 2
DEBUG PAGE

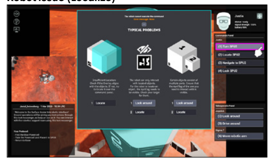
Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.05 - 26.05

In this part of the evaluation, we will focus on understanding how the system assists you in the moment of an error. You will be presented with an initial situation and an error message. Your task is to interact with the system by either clicking on elements in the user interface or describing what actions you would take to resolve the error. Please note that the user interface has limited functionality in this prototype.

Again, I kindly ask you to engage in a 'Talk out loud' technique while interacting with the prototype. Please verbalize your thoughts, explaining why you are taking certain actions and expressing any challenges or observations you encounter. Additionally, I will provide you with PrEmos, a tool you might be familiar with from the last session. Remember, the numbers do not indicate a rating; simply choose the one that accurately reflects your feelings.

I am particularly interested in your thoughts and actions related to the error. You can expand the error message but before you do so, please share your initial perception of what the error might be. After expanding the error message, I would like to know your updated understanding of the error cause. Additionally, for each assumption you make about the error cause, please describe the strategy you would employ to resolve the issue. In other words, what steps would you take to fix the issue? (Give them the Prototype)

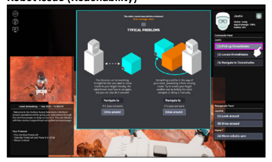
Robot issue (Localize)



What caused the problem?

What would you do?

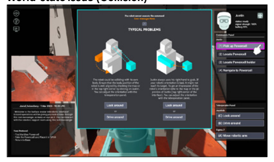
Robot issue (Reachability)



What caused the problem?

What would you do?

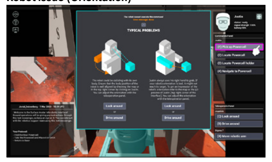
World-state issue (Collision)



What caused the problem?

What would you do?

Robot issue (Orientation)



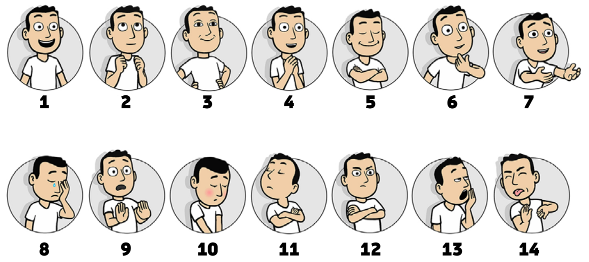
What caused the problem?

What would you do?

PREMO
HOW DO YOU FEEL ABOUT THIS SITUATION?

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

PrEmo measures distinct (pleasant and unpleasant) emotions in a non-verbal manner that is validated cross-culturally. Please look at the images below and use the numbers to write into the scenarios and indicate how you feel about certain aspect in them. Feel free to use as many as you need, there is no limit



EXPERIENCE EVALUATION
DEBUG PAGE PROTOTYPE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression. **Please try to answer spontaneously!**

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

NASA TASK LOAD INDEX
DEBUG PAGE PROTOTYPE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.
Please try to answer spontaneously!

Mental Demand — How mentally demanding was the task?
Very Low — Very High

Physical Demand — How physically demanding was the task?
Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?
Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?
Very Low — Very High

PROTOTYPE – 2

DISCUSSION GUIDE

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

INTERVIEW QUESTIONS

Could you please tell us a little bit about yourself?

- Your position at DLR
- Other experience with robots?
- You tasks
- Have you interacted with the Robot and the current system already?

Were there any specific moments during the interaction that stood out to you positively or negatively? Can you describe those moments in more detail?

- Did you encounter any challenges or frustrations while interacting with the prototype? If yes, could you elaborate on those?
- Were there any aspects of the prototype that you found particularly intuitive or user-friendly?

How intuitive was the system in guiding you towards resolving the error?

- Did you feel adequately supported and guided throughout the error scenario, or did you find yourself seeking additional help or information?
- Were there any instances where you felt the system could have provided clearer instructions or more explicit guidance?
- Did you feel confident in your ability to effectively address the error with the support provided by the prototype?

How useful and informative did you find the additional information provided when expanding the error message?

- Did the expanded error page clarify your understanding of the error cause? If not, what information would you have liked to see?

Did the error handling prototype enhance your understanding of common errors and their potential causes?

- Based on your experience with the prototype, do you feel more equipped to handle similar errors in real-world scenarios?

Were there any particular aspects of the error message or the user interface that were unclear or confusing to you?

- Did you face any challenges in identifying or understanding the cause of the error?

How engaging did you find the interaction with the Debug page?

- Did it hold your attention and maintain your interest throughout the session?
- Were there any aspects that could be improved to enhance engagement?

Were there any specific features or functionalities of the prototype that stood out to you? If so, why?

- Is there any additional information or guidance that you would like to see in the error message to help you resolve the issue more effectively?
- Are there any features or functionalities you feel are missing from the user interface that could enhance your ability to address errors?

Is there anything else you would like to share about your experience with the prototype or any other insights you believe would be valuable for us to know?

MORE TIME?

-> Comparative Study

PROTOTYPE 1

PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Reachability

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Orientation

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Collision

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

PROTOTYPE 2

PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Reachability

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Orientation

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Collision

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Localize

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
Daniel														
Nesrine														
Nicole														
Neal														
Florian														
Xuwei														
Adrian														
Anne														
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

16

1ST PERSON

PREMO

Research Plan Thesis DLR (2023)
Liliane (Lilly) Filthaut | 5609402
On Site Studies 22.02 - 08.03

Reachability

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Orientation

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Collision

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Participant	Joy	Admiration	Pride	Hope	Satisfaction	Fascination	Desire	Sadness	Fear	Shame	Contempt	Anger	Boredom	Disgust
	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Raw Data Protoype Testing

UEQ – 1ST PERSON

Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
5	3	5	2	3	5	5	5	3	4	5	3	6	4	3	4	4	3	3	5	3	4	3	3	3	4	CP01
5	3	1	2	3	5	5	4	3	4	2	3	4	5	5	5	3	2	4	4	4	5	4	4	2	5	CP02
5	6	3	3	2	5	7	5	4	3	6	3	3	5	5	5	3	3	4	5	2	6	2	3	1	6	CP03
6	3	6	5	5	2	3	4	3	5	3	6	2	3	6	3	5	7	7	3	6	1	3	3	4	4	CP04
6	5	2	2	3	6	5	6	3	4	6	2	5	6	5	6	2	2	2	6	2	5	4	2	2	7	CP05
4	3	2	4	4	3	6	3	2	3	1	3	2	4	4	4	4	4	6	3	5	2	4	4	4	4	CP06
6	3	6	4	4	3	3	5	3	4	2	6	3	2	5	3	5	7	6	2	7	1	6	6	5	3	CP07
4	2	5	6	4	3	3	3	2	4	2	5	3	3	4	3	2	4	5	3	4	6	3	4	3	6	CP08

UEQ – 3RD PERSON

Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
7	7	2	1	1	7	7	3	1	1	6	1	6	7	6	6	4	2	1	6	2	6	2	2	2	6	P01
6	7	3	1	1	6	7	4	1	4	6	2	6	7	5	5	2	2	2	5	2	6	2	2	2	6	P04
6	6	2	3	2	4	5	5	2	4	6	2	6	5	4	6	3	2	2	6	3	5	3	2	4	5	P02
6	6	3	2	2	6	5	6	2	3	5	2	6	6	5	6	3	2	2	6	3	5	2	2	2	5	P08
7	7	1	1	1	7	7	5	2	2	7	1	6	7	6	7	2	1	1	6	1	7	1	1	1	7	P06
6	6	2	1	1	6	7	6	2	2	6	2	6	6	6	7	2	2	2	6	1	7	2	1	1	6	P03
7	7	1	1	1	7	7	6	2	1	6	1	6	7	6	7	2	1	1	6	2	7	1	1	1	6	P07
5	5	3	4	3	6	6	5	3	4	5	2	5	5	4	5	3	3	3	5	3	4	5	4	3	5	P05

UEQ – DEBUG

Items																										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
5	6	4	2	1	3	4	4	2	4	7	1	7	7	4	6	4	2	2	6	1	7	1	3	2	4	P01
7	6	2	1	1	6	6	5	2	3	7	1	6	6	5	6	2	2	2	6	2	6	2	2	2	6	P04
5	7	3	1	2	5	6	7	5	3	7	1	6	6	5	5	2	2	2	6	2	6	2	2	2	4	P02
5	6	2	2	2	5	5	6	2	2	6	3	6	6	6	6	2	3	3	6	3	5	2	2	2	5	P08
7	7	1	1	1	7	7	7	2	1	7	1	7	7	7	7	1	1	1	7	1	7	1	1	1	7	P06
6	6	1	1	1	7	7	6	1	2	7	1	6	6	6	6	1	1	1	7	2	7	1	1	1	6	P03
7	7	1	1	1	7	7	6	1	1	7	1	6	7	6	6	1	1	1	7	1	7	2	2	2	6	P07
5	5	3	3	3	6	6	6	3	3	5	3	6	5	4	5	3	3	4	5	3	5	3	3	3	4	P05

NASA TLX – 1ST PERSON

	B	C	D	E	F	G
User #	Mental	Physical	Temporal	Performance	Effort	Frustration
0/Ex	55	10	75	25	80	40
CP01	45	10	45	75	45	70
CP02	55	5	35	40	50	25
CP07	65	10	50	75	70	85
CP04	50	1	50	25	60	35
CP08	35	5	25	40	20	65
CP03	50	20	50	40	55	65
CP05	25	5	20	15	75	15
CP06	55	5	40	65	50	55

Raw/Unweighted		STAWN
Overall	41,06	22,20492274
Diagnostic Subscores		STAWN
Mental	47,59	11,7260394
Physical	7,63	5,429951657
Temporal	43,21	11,02199506
Performance	46,60	21,05610066
Effort	53,46	15,79903082
Frustration	51,73	22,767507

NASA TLX – 3RD PERSON

	B	C	D	E	F	G
User #	Mental	Physical	Temporal	Performance	Effort	Frustration
0/Ex	20	5	10	35	15	10
P01	20	5	10	35	15	10
P03	15	5	25	25	15	5
P04	50	20	15	25	35	40
P08	15	10	10	15	20	15
P02	60	5	20	50	30	40
P05	50	2	25	30	40	50
P06	35	5	25	25	35	40
P07	25	10	10	40	35	35

Raw/Unweighted		STAWN
Overall	24,52	14,55448039
Diagnostic Subscores		STAWN
Mental	33,75	16,53594569
Physical	7,75	5,285593628
Temporal	24,06	6,614378278
Performance	30,63	10,13579671
Effort	28,13	9,333240327
Frustration	29,38	15,69982086

NASA TLX – DEBUG

	B	C	D	E	F	G
User #	Mental	Physical	Temporal	Performance	Effort	Frustration
0/Ex	55	10	75	25	80	40
P01	25	5	10	25	15	20
P03	10	5	10	15	25	15
P04	70	5	50	20	45	50
P08	40	20	20	25	20	15
P02	25	5	20	30	20	15
P05	25	5	25	20	20	15
P06	55	15	30	20	40	15
P07	35	10	20	30	30	10

Raw/Unweighted		
Overall	22,81	
Diagnostic Subscores		Standard deviation
Mental	35,63	17,92998536
Physical	8,75	5,448623679
Temporal	23,13	11,97327754
Performance	23,13	4,960783708
Effort	26,88	9,980449639
Frustration	19,38	11,84205958

3RD PERSON VS 1ST PERSON – TIME SPENT AND ATTEMPTS

Third person attempts to solve						
3rd Attempt	Powercell: Reachability	help*	SPU5: Orientation	help*	Base: Collision	help*
P01	2	0	3	0	1	0
P04	2	0	3	0	2	0
P03	2	0	6	1	1	0
P08	3	0	5	1	2	0
P06	2	0	4	0	2	0
P02	2	0	2	1	2	0
P07	2	0	8	1	2	0
P05	2	0	2	0	2	0
Mean	2,125	0	4,125	4	1,75	0
STABWN	0,330718914		1,964529206		0,433012702	

with/without help difference	
Mean	SPU5: Orientation
With help	5,25
Without help	3

Debug Page: attempts to solve**				
Debug Attempt	Localize	Reachability	Collision	Orientation
P01	1	1	1	1
P04	1	1	1	1
P03	2	1	1	1
P08	1	1	1	1
P06	1	1	1	1
P02	2	1	1	1
P07	1	1	1	1
P05	1	1	1	1
Mean	1,25	1	1	1
STABWN	0,4330127	0	0	0

Third person time spent until solved in seconds						
Time in s	Powercell: Reachability	help*	SPU5: Orientation	help*	Base: Collision	help*
P01	95	0	144	0	50	0
P04	80	0	313	0	44	0
P03	83	0	148	1	15	0
P08	265	0	468	1	114	0
P06	84	0	156	0	40	0
P02	160	0	143	1	57	0
P07	76	0	230	1	30	0
P05	122	0	116	0	36	0
Mean	120,625	0	214,75	4	48,25	0
STABWN	60,67523692		112,823701		27,56242914	

with/without help difference	
Mean	SPU5: Orientation
With help	247,25
Without help	182,25

Debug Page: time spent until solved in seconds**				
Time in s	Localize	Reachability	Collision	Orientation
P01	75	20	41	32
P04	41	37	74	104
P03	66	24	24	21
P08	108	42	143	120
P06	35	18	46	47
P02	55	19	29	35
P07	32	19	50	38
P05	46	38	35	81
Mean	57,25	27,125	55,25	59,75
STABWN	23,7052209	9,43977622	36,1308386	34,567868

**No situation with external help

First person attempts to solve						
1st Attempt	Powercell: Reachability	help*	SPU5: Orientation	help*	Base: Collision	help*
CP01	2	0	4	1	2	0
CP02	1	0	5	1	2	0
CP03	1	0	2	0	2	0
CP04	1	0	3	1	2	0
CP05	1	0	5	1	2	0
CP06	3	0	2	1	2	0
CP07	2	0	4	1	2	0
CP08	2	0	4	1	2	0
Mean	1,625	0	3,625	7	2	0
STABWN	0,695970545		1,111024302		0	

with/without help difference	
Mean	SPU5: Orientation
With help	3,857142857
Without help	2

First person time spent until solved in seconds						
Time in s	Powercell: Reachability	help*	SPU5: Orientation	help*	Base: Collision	help*
CP01	30	0	215	1	41	0
CP02	64	0	204	1	62	0
CP03	130	0	113	0	105	0
CP04	68	0	213	1	51	0
CP05	162	0	339	1	65	0
CP06	90	0	131	1	94	0
CP07	25	0	143	1	38	0
CP08	168	0	176	1	29	0
Mean	92,125	0	191,75	7	60,625	0
STABWN	52,22173278		66,46568664		25,23359616	

with/without help difference	
Mean	SPU5: Orientation
With help	203
Without help	113

PARTICIPANT STATISTICS – PROTOYPE TESTING

Name	Part of Team	Usage Experience	Training Experience	Astronaut Contacts
01	Yes	Multiple times	Observed once	A few: Spoken to and observed
02	No	Tester once, couple days ago	Check out once	None
03	No	3 Months ago once	Observed two	None
04	Yes	Many times Tester	Few Checkouts	A few: Spoken to and observed
05	No	None	Observed once	None
06	Yes	Several Years: Many times tester and test observer	Observed multiple	Many Years: Spoken to, observed, trained
07	No	6 Months ago once	Observed once	None
08	Yes	5 Years with Surface Avatar	Holds Checkouts and observed many Trainings	Many Years: Spoken to, observed, trained

PARTICIPANT STATISTICS – BASELINE TESTING

Name	Part of Team	Usage Experience	Training Experience	Astronaut Contacts
01	Yes	Multiple times	Observed once	A few: Spoken to and observed
02	Yes	Many times Tester	Few Checkouts	None
03	No	None	Observed once	None
04	No	3 Months ago once	Observed once	None
05	Yes	3 Months ago once	Few Checkouts	None
06	No	Several Years: Many times	Observed multiple	A few: Observed
07	No	None	None	None
08	No	Multiple times	Observed multiple	A few: Spoken to and observed

Raw Data Questionnaire A/B Comparison

3RD PERSON VS 1ST PERSON – “OK“ ERROR RESOLUTION STRATEGIES

Localisation					
Localisation 1: Rodin		Localisation 2: Base		Localisation 3: SPU3	
Third person	First person	Third person	First person	Third person	First person
90,5	95,2	90,5	85,7	95,2	90,5

Reachability					
Reachability 1: Grab Powercell		Reachability 2: Place Loadingstation		Reachability 3: Open Door	
Third person	First person	Third person	First person	Third person	First person
90,5	66,7	81	57,1	85,7	81

Collison					
Collision 1: Drone Placement		Collision 2: Stuck between		Collision 3: Multi Powercell	
Third person	First person	Third person	First person	Third person	First person
85,7	66,7	81	47,6	62	47,6

Orientation					
Orientation 1: Grab Powercell		Orientation 2: Droneblade Floor		Orientation 3: SPU5	
Third person	First person	Third person	First person	Third person	First person
47,6	4,8	66,7	28,6	52,4	19

	Localize: SPU2		
	No help	Typical Debug	Motion Planning
4,8	14,3	47,6	4,8
23,8			
9,5	0	0	
	Reachability: Drone blade		
57,1	No help	Typical Debug	Motion Planning
38,1	47,6	66,7	71,4
9,5	0	0	
	Collision: Stones		
76,2	No help	Typical Debug	Motion Planning
76,2	76,2	66,7	38,1
0	0	0	
	Orientation: Rodin		
14,3	No help	Typical Debug	Motion Planning
4,8	9,55	62	4,8
4,75	0	0	

3RD PERSON VS 1ST PERSON – “BAD“ ERROR RESOLUTION STRATEGIES

Localisation					
Localisation 1: Rodin		Localisation 2: Base		Localisation 3: SPU3	
Third person	First person	Third person	First person	Third person	First person
9,5	0	4,5	9,5	0	4,8

Reachability					
Reachability 1: Grab Powercell		Reachability 2: Place Loadingstation		Reachability 3: Open Door	
Third person	First person	Third person	First person	Third person	First person
4,8	14,3	4,8	23,8	14,3	9,5

Collison					
Collision 1: Drone Placement		Collision 2: Stuck between		Collision 3: Multi Powercell	
Third person	First person	Third person	First person	Third person	First person
9,5	23,8	14,3	38,1	38,1	42,9

Orientation					
Orientation 1: Grab Powercell		Orientation 2: Droneblade Floor		Orientation 3: SPU5	
Third person	First person	Third person	First person	Third person	First person
14,3	28,6	9,5	47,6	4,8	19

	Localize: SPU2		
	No help	Typical Debug	Motion Planning
19	19	4,8	14,3
19			
0	0	0	
	Reachability: Drone blade		
9,5	No help	Typical Debug	Motion Planning
14,3	11,9	4,8	9,5
2,4	0	0	
	Collision: Stones		
14,3	No help	Typical Debug	Motion Planning
23,8	19,05	19	47,6
4,75	0	0	
	Orientation: Rodin		
14,3	No help	Typical Debug	Motion Planning
47,6	30,95	9,5	47,6
16,65	0	0	

3RD PERSON VS 1ST PERSON – CORRECT ERROR IDENTIFICATION / “OK“ ANSWERS

Reachability					
Reachability 1: Grab Powercell		Reachability 2: Place Loadingstation		Reachability 3: Open Door	
Third person	First person	Third person	First person	Third person	First person
95	85,7	85	52,4	90	76,2

Collison					
Collision 1: Drone Placement		Collision 2: Stuck between		Collision 3: Multi Powercell	
Third person	First person	Third person	First person	Third person	First person
75,5	47,7	85	47,7	65	28,6

Orientation					
Orientation 1: Grab Powercell		Orientation 2: Droneblade Floor		Orientation 3: SPU5	
Third person	First person	Third person	First person	Third person	First person
50	14,3	70	33,4	85	14,3

	Localize: SPU2		
	Mean: no help	Typical Debug	Motion Planning
	65,85	85,7	35
STABWN	0,85	0	0

	Reachability: Drone blade		
	Mean: no help	Typical Debug	Motion Planning
	33,95	95	85,7
STABWN	8,95	0	0

	Collision: Stones		
	Mean: no help	Typical Debug	Motion Planning
	14,5	81	20
STABWN	4,5	0	0

	Orientation: Rodin		
	Mean: no help	Typical Debug	Motion Planning
	9,8	89,5	28,6
STABWN	0,2	0	0

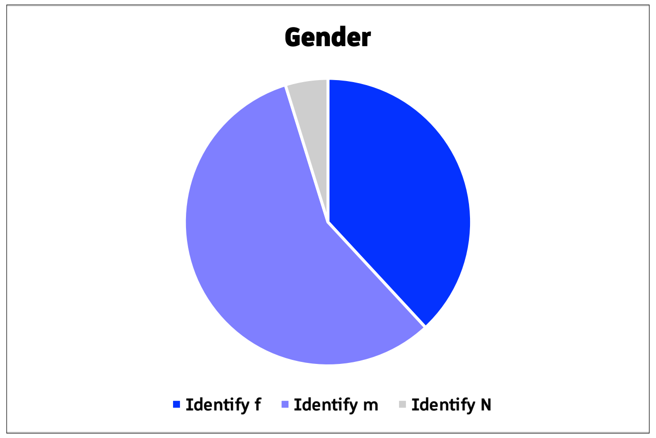
Localize: SPU2	
No help A	No help B
66,7	65

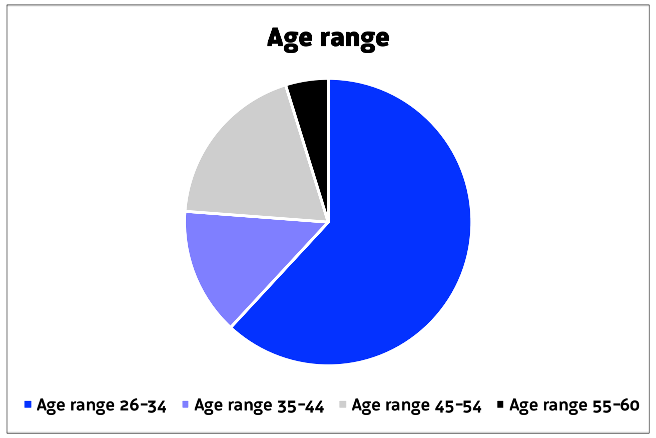
Reachability: Drone blade	
No help A	No help B
42,9	25

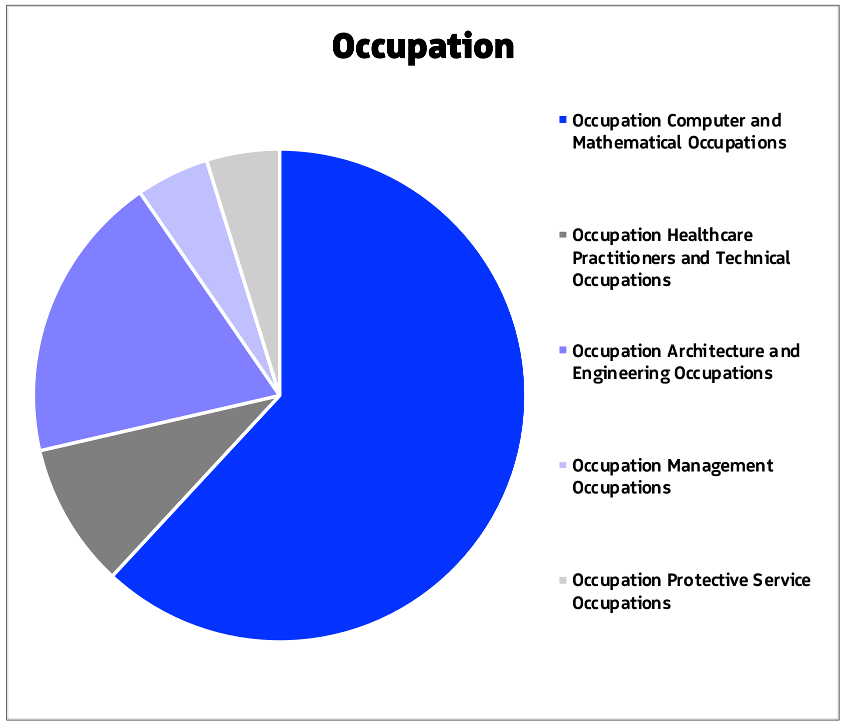
Collision: Stones	
No help A	No help B
10	19

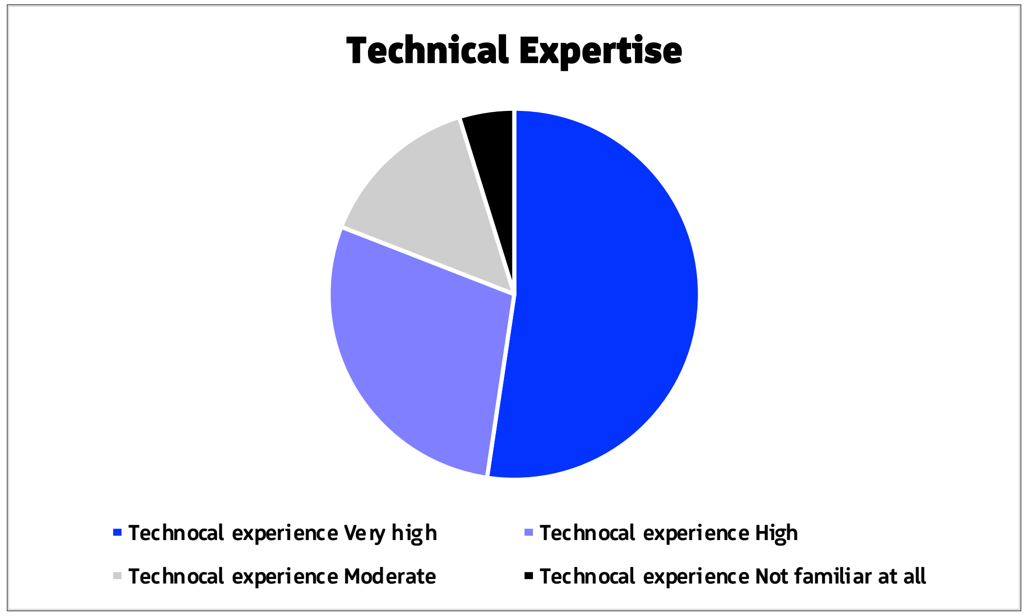
Orientation: Rodin	
No help A	No help B
9,6	10

PARTICIPANT STATISTICS QUESTIONNAIRE A (THIRD PERSON)

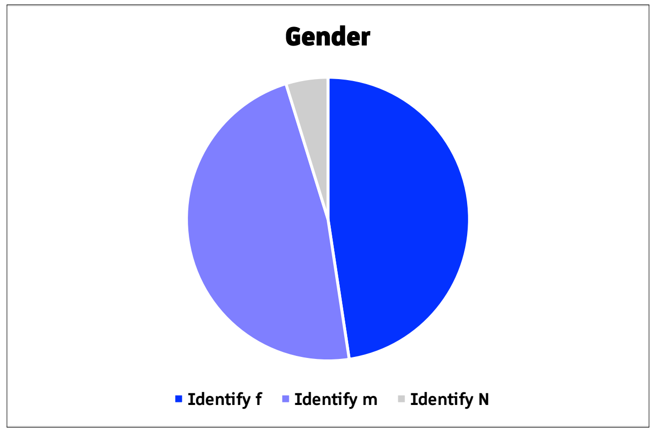


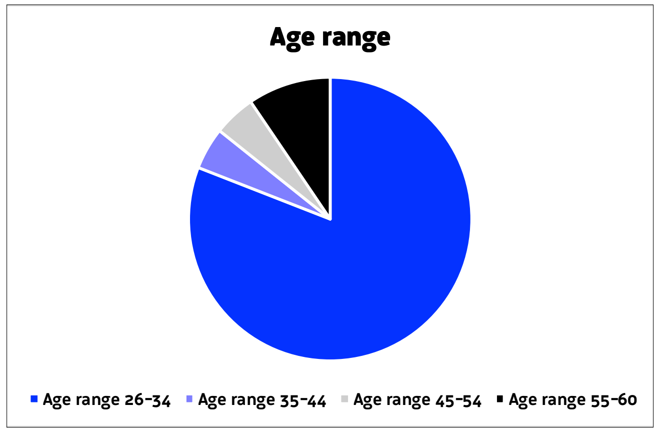


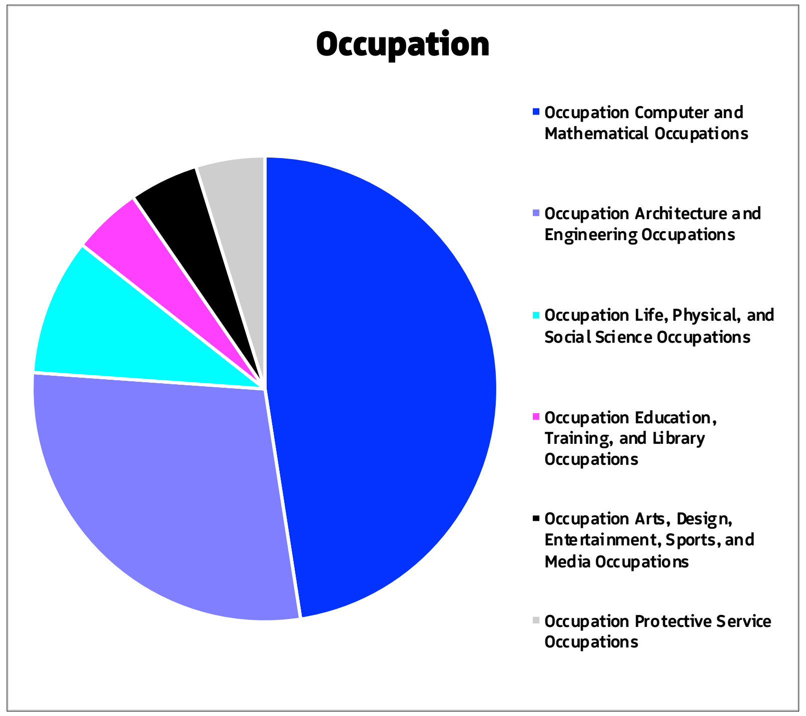


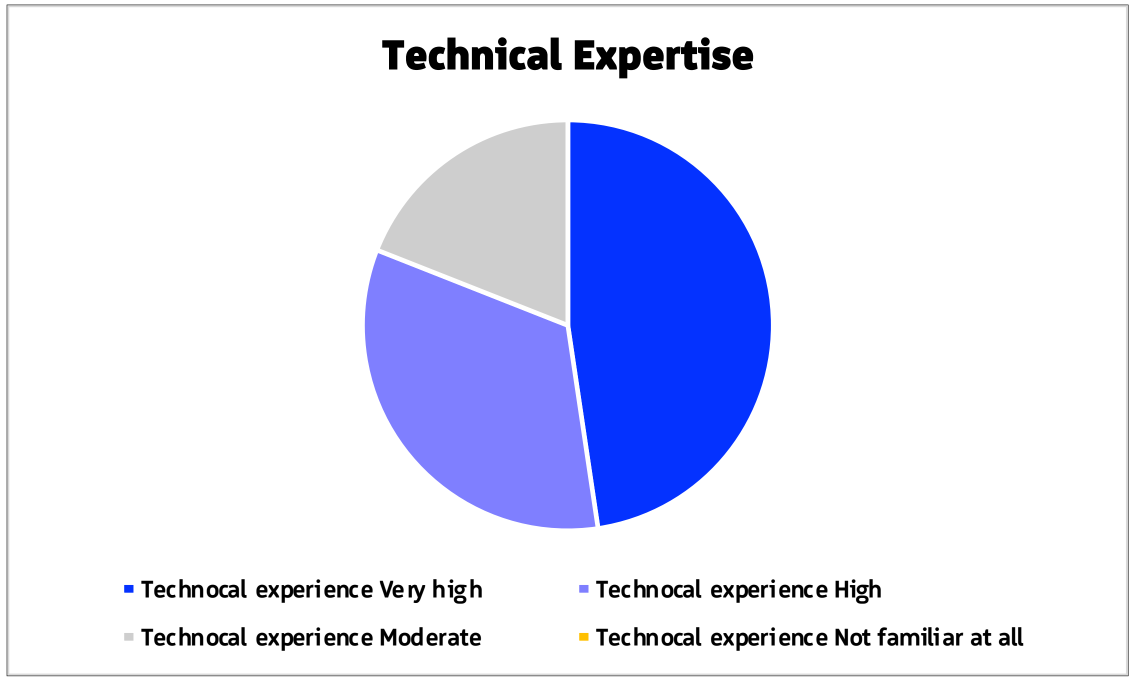


PARTICIPANT STATISTICS QUESTIONNAIRE B (FIRST PERSON)





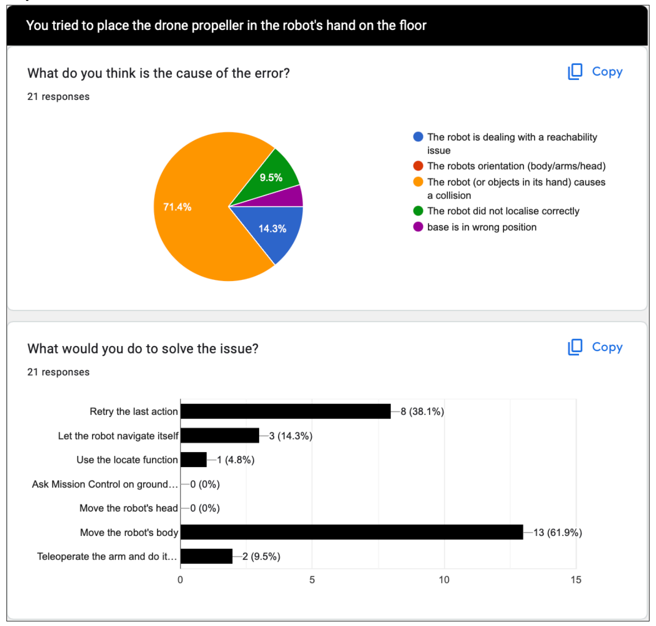


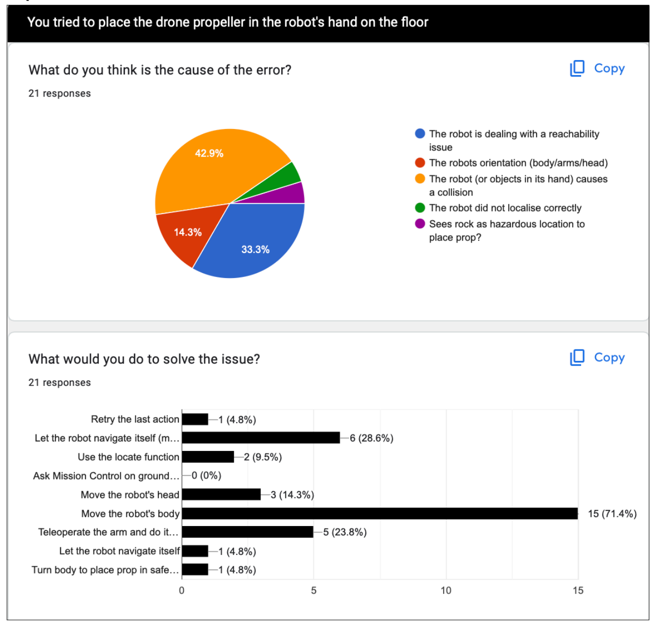


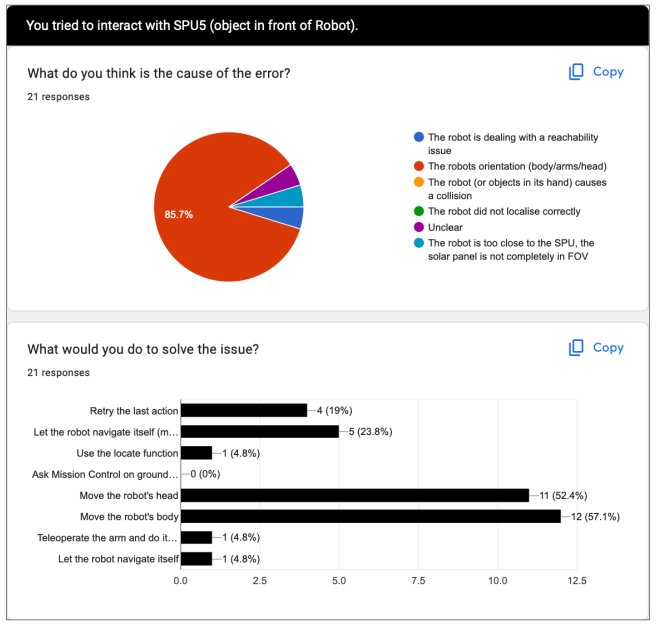
QUESTIONNAIRE RESULTS FROM A VS. B

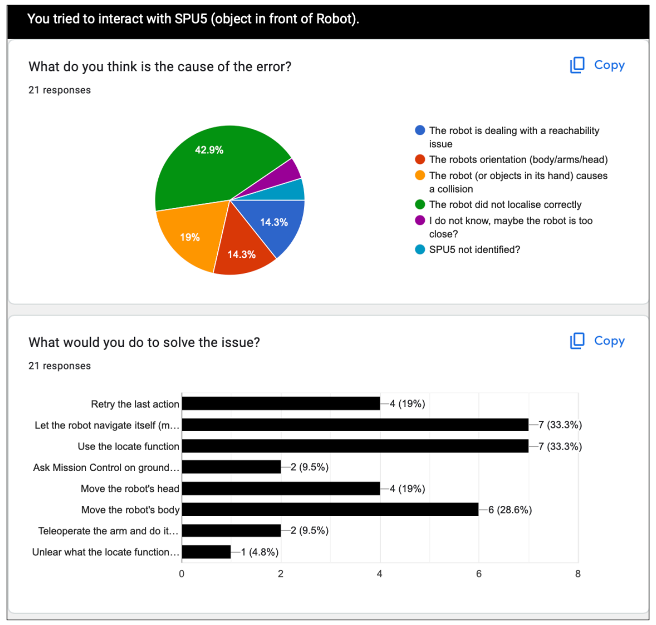
QUESTIONNAIRE A

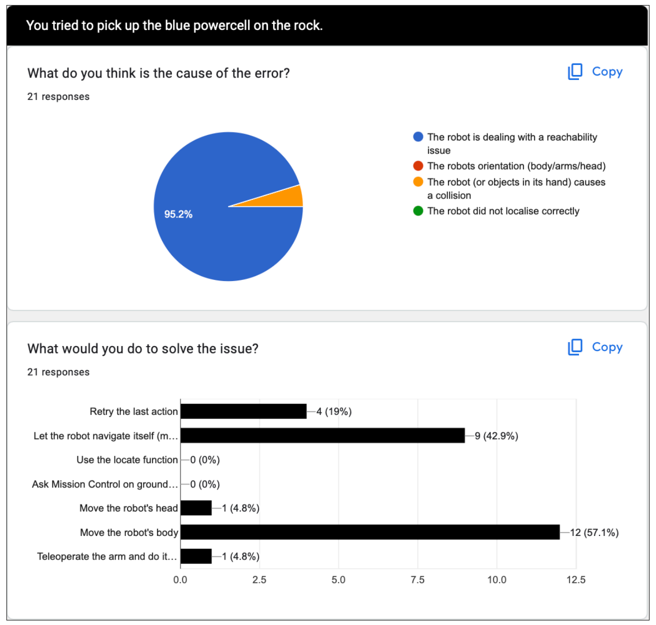
QUESTIONNAIRE B

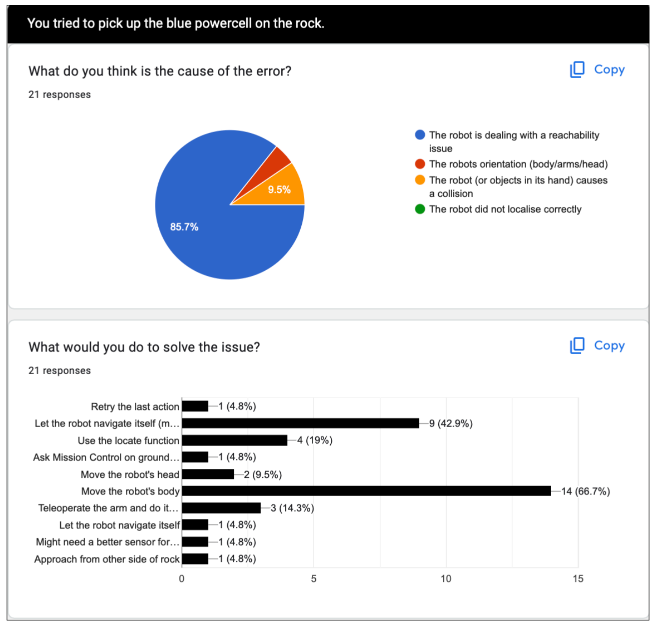




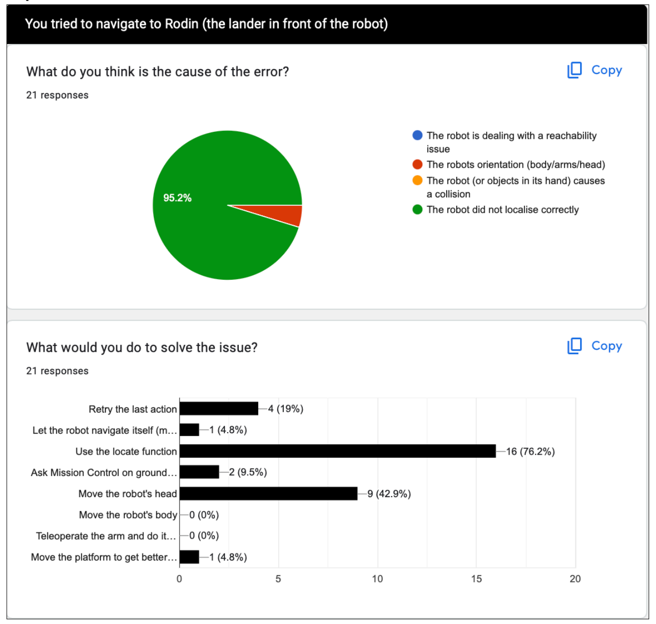


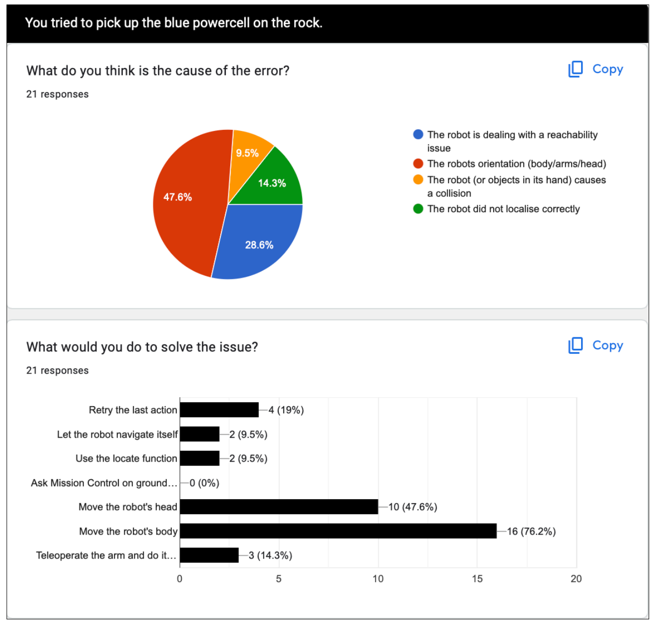


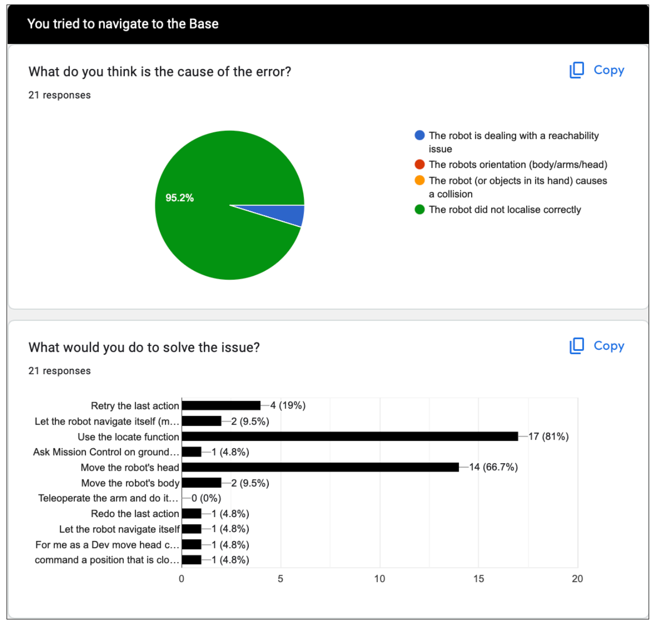




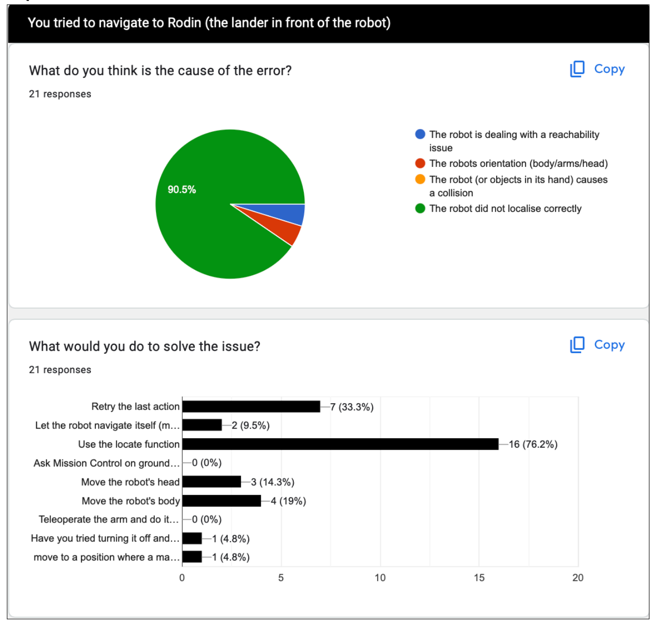
QUESTIONNAIRE A

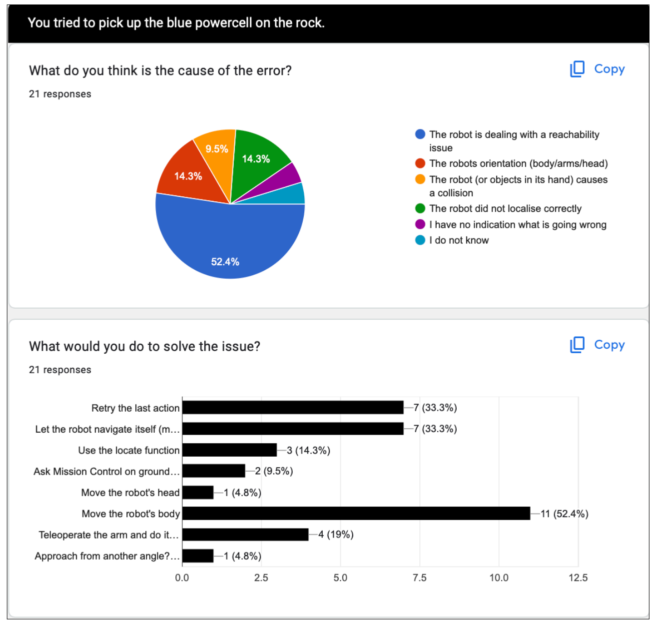


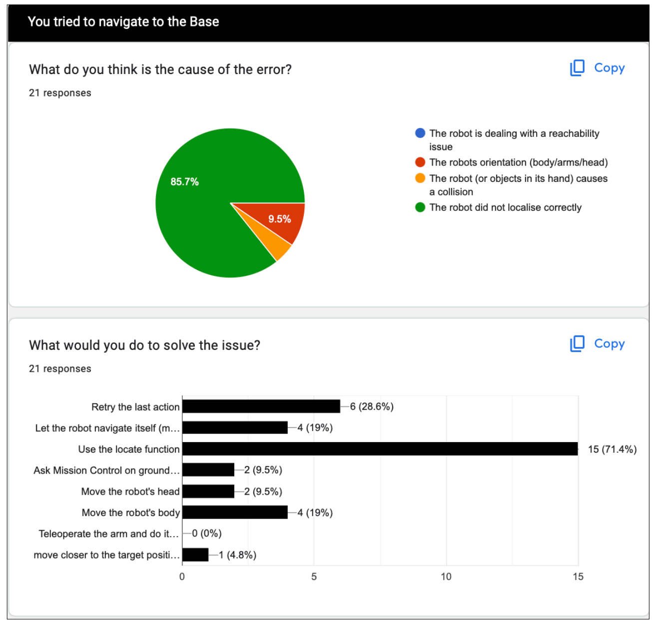




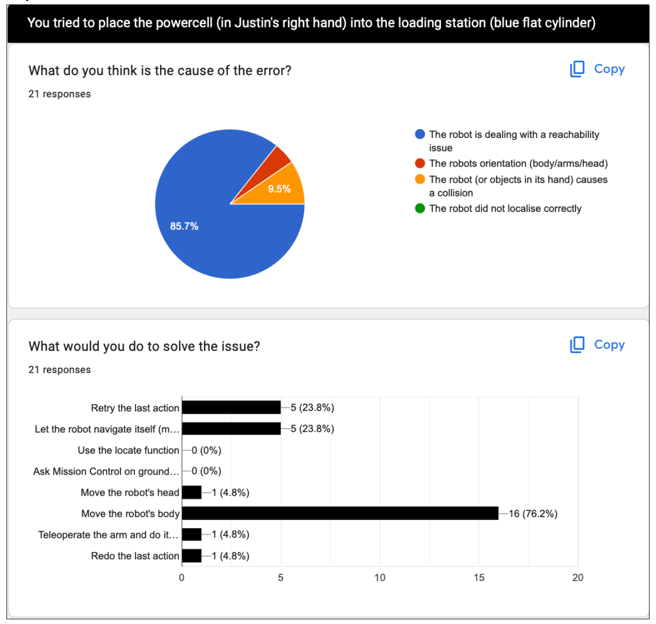
QUESTIONNAIRE B

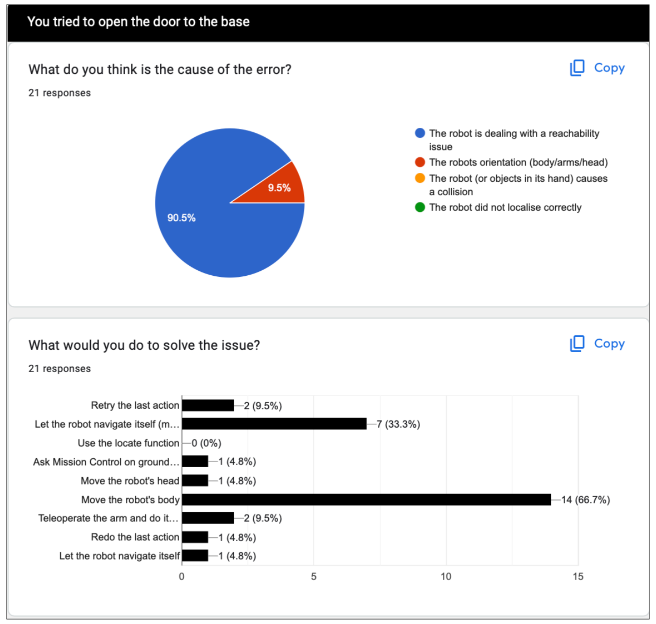


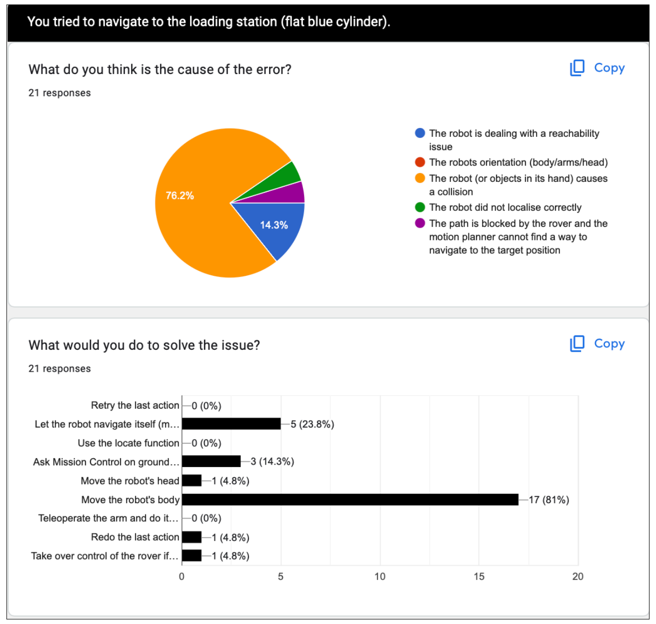




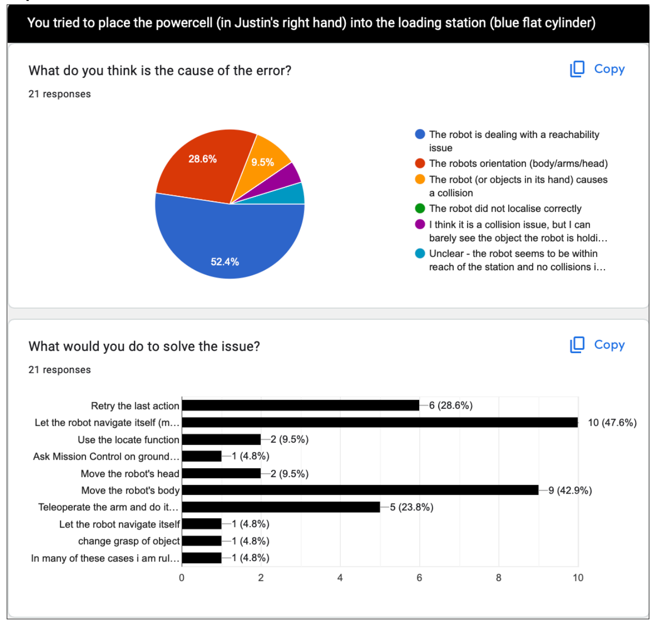
QUESTIONNAIRE A

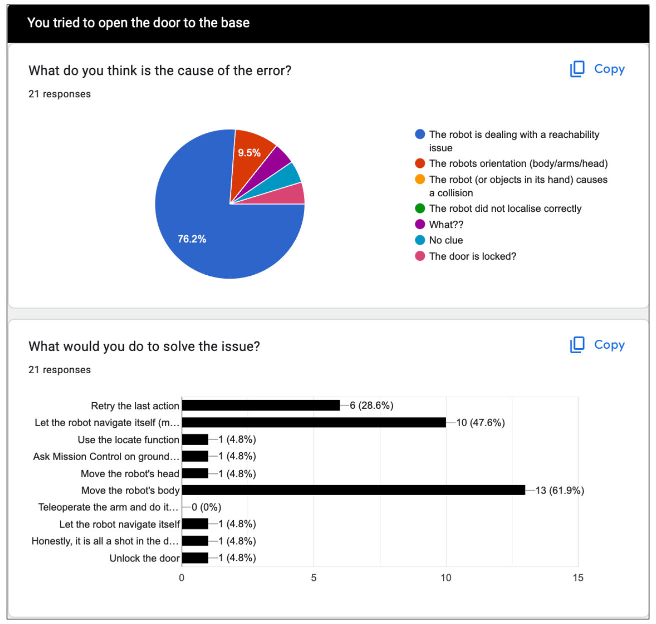


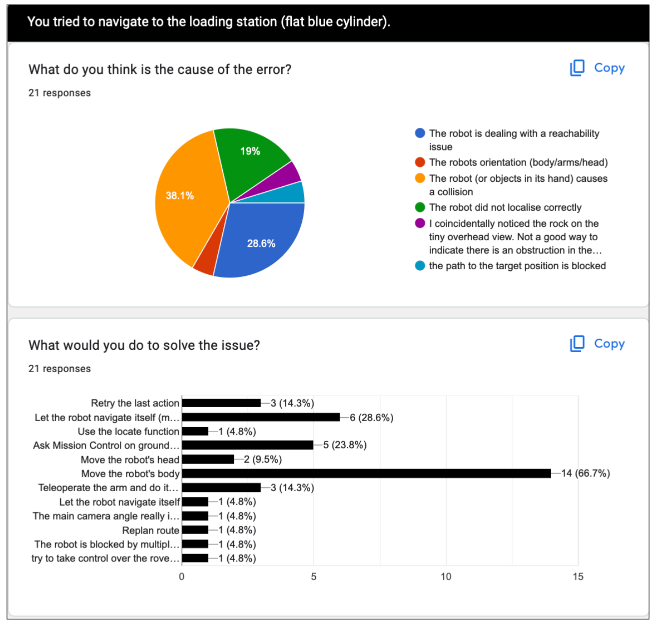




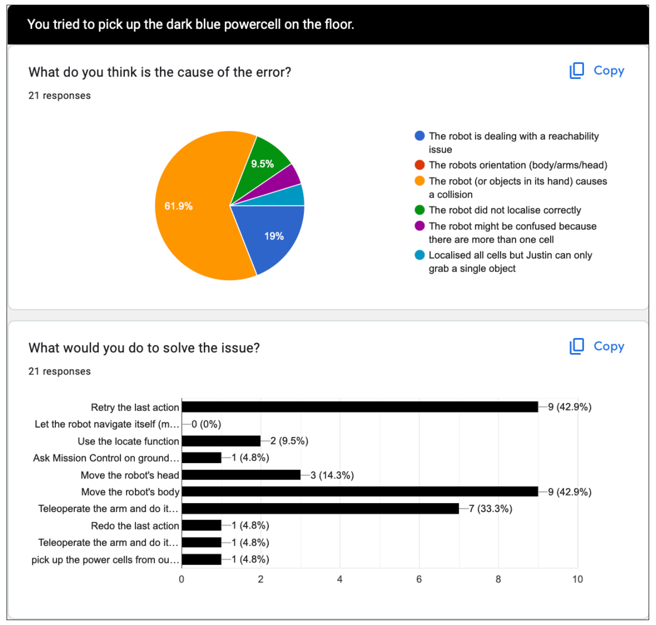
QUESTIONNAIRE B

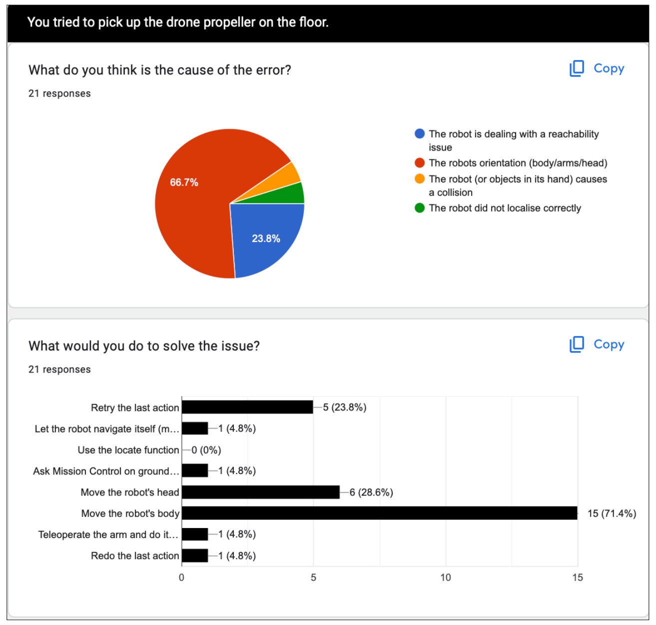


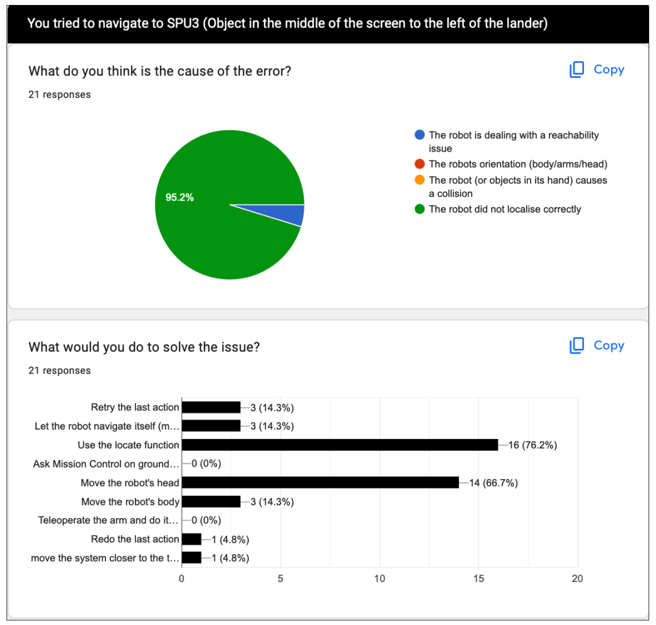




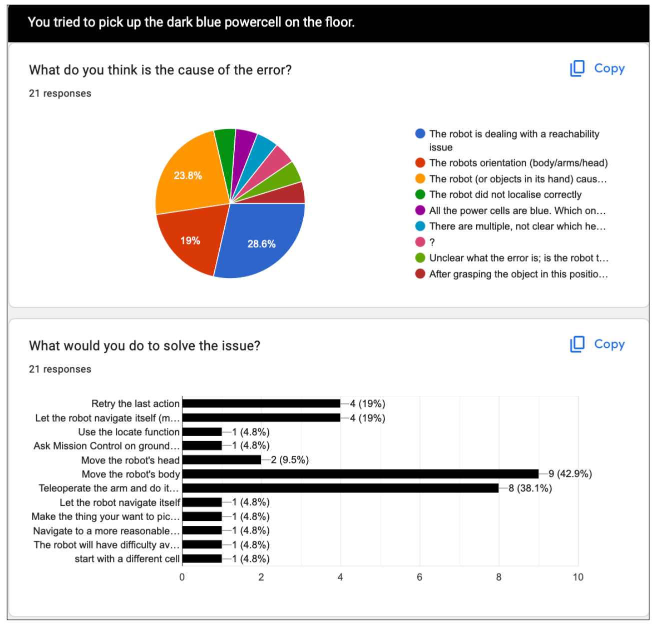
QUESTIONNAIRE A

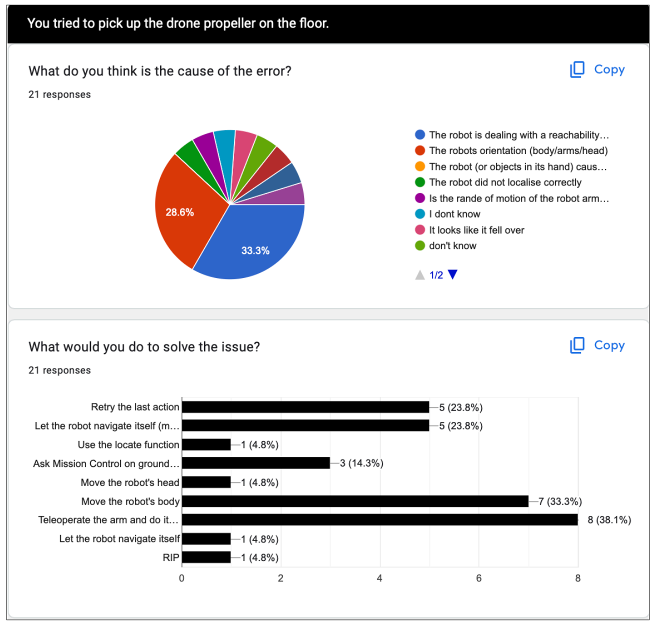


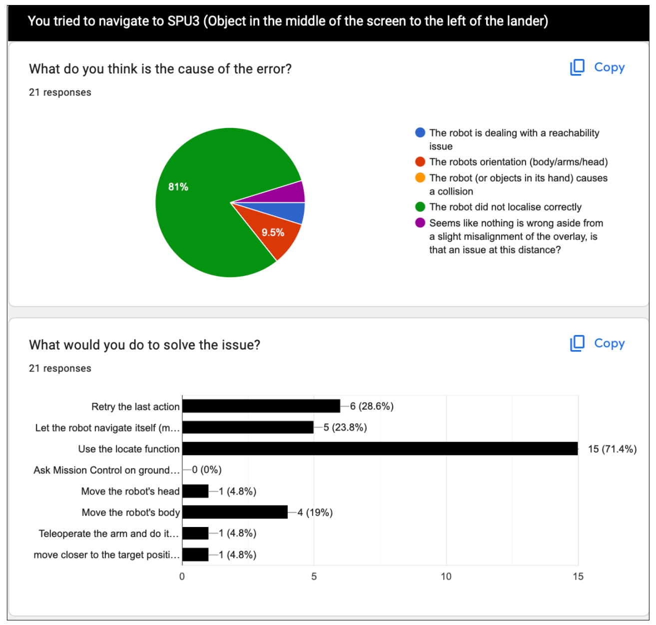




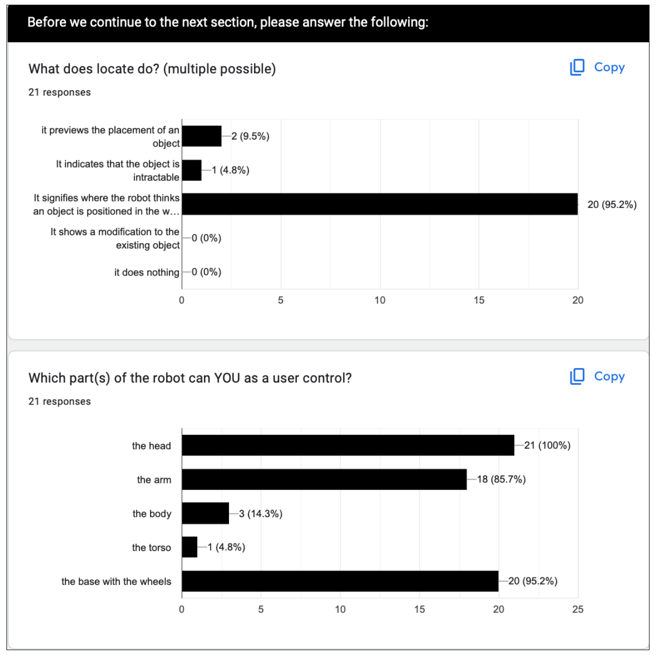
QUESTIONNAIRE B

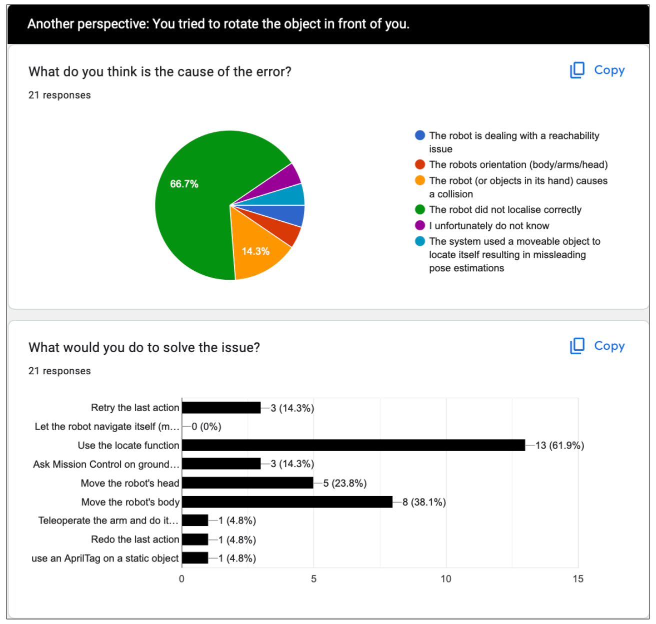


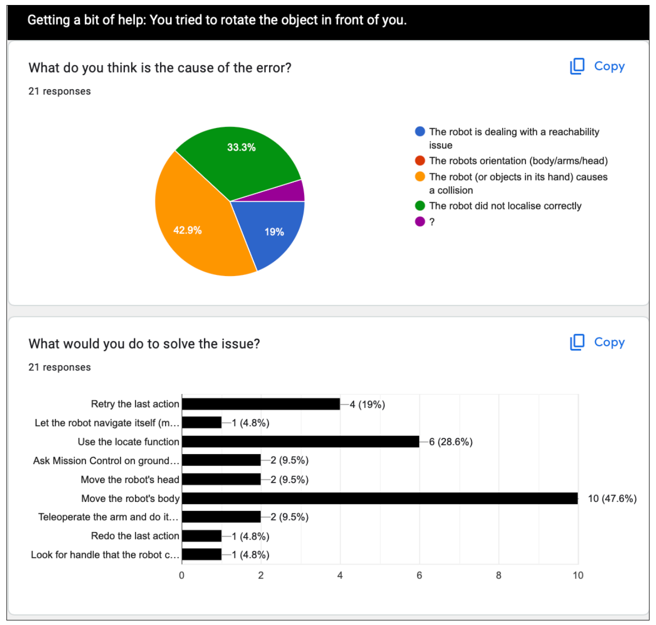




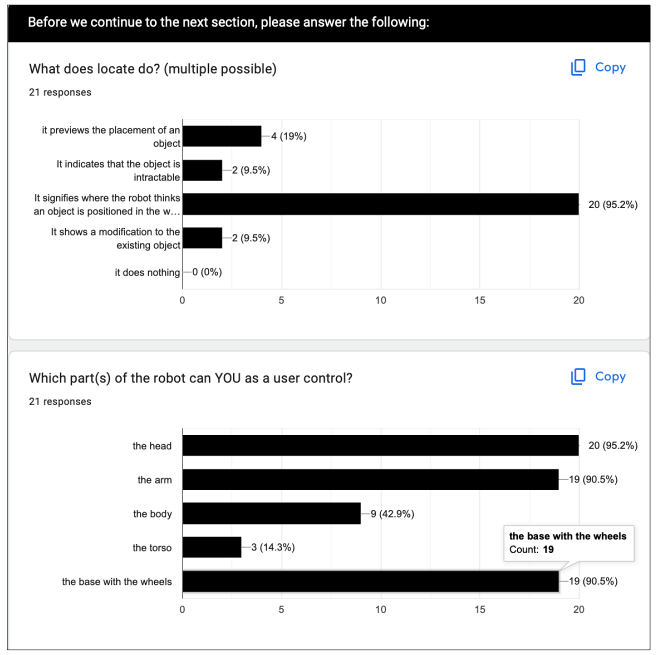
QUESTIONNAIRE A

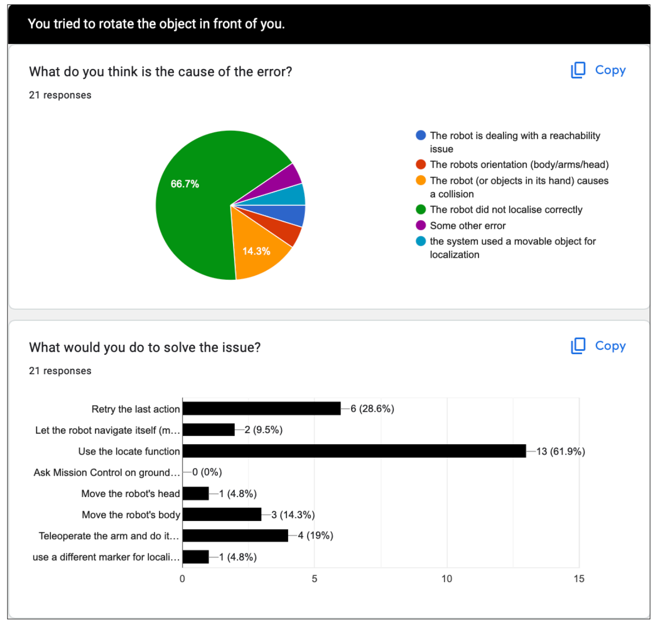


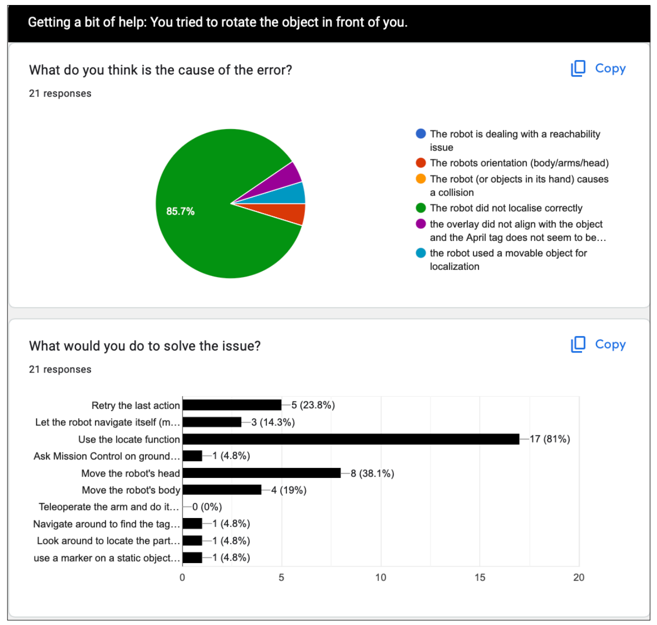




QUESTIONNAIRE B

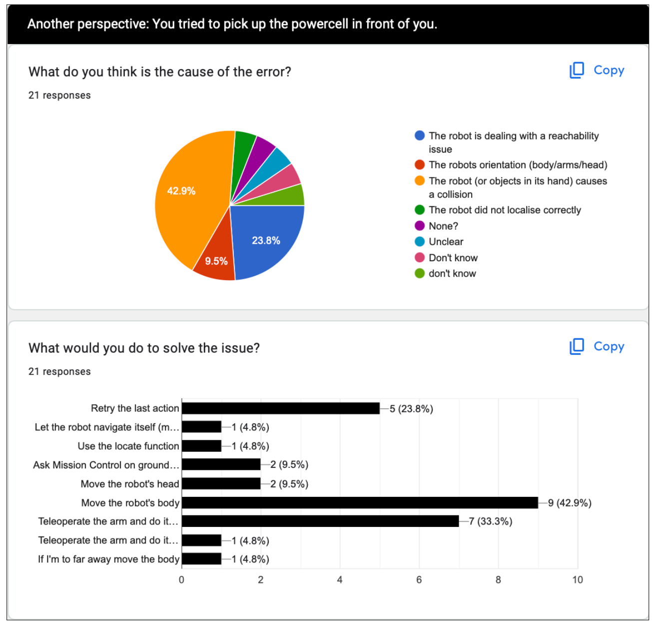


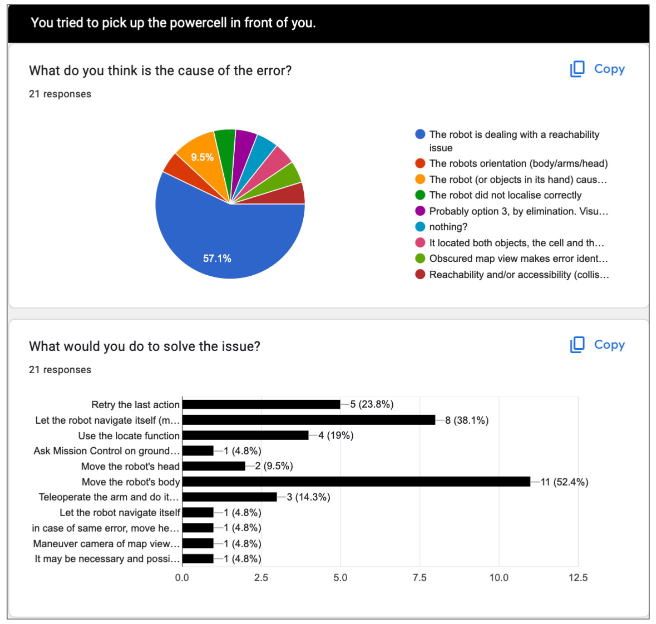


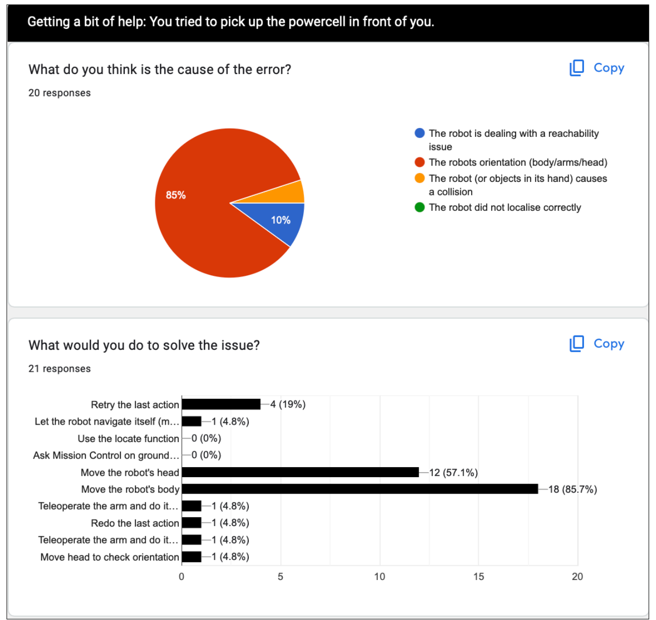


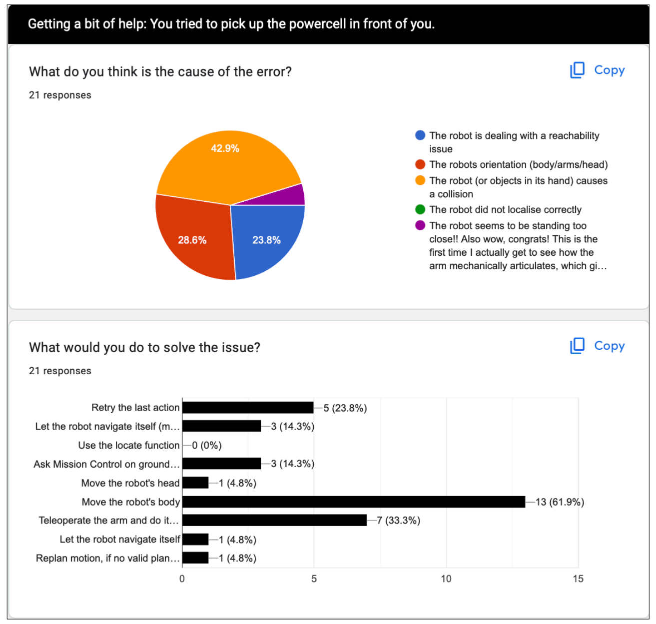
QUESTIONNAIRE A

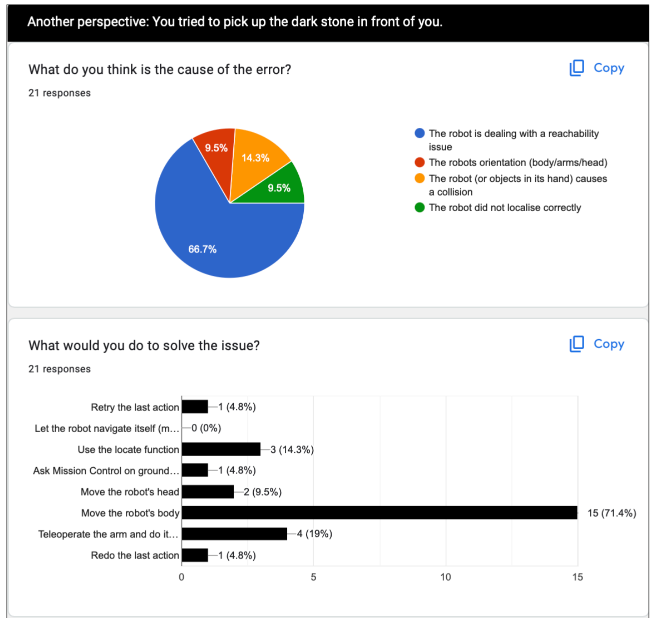
QUESTIONNAIRE B

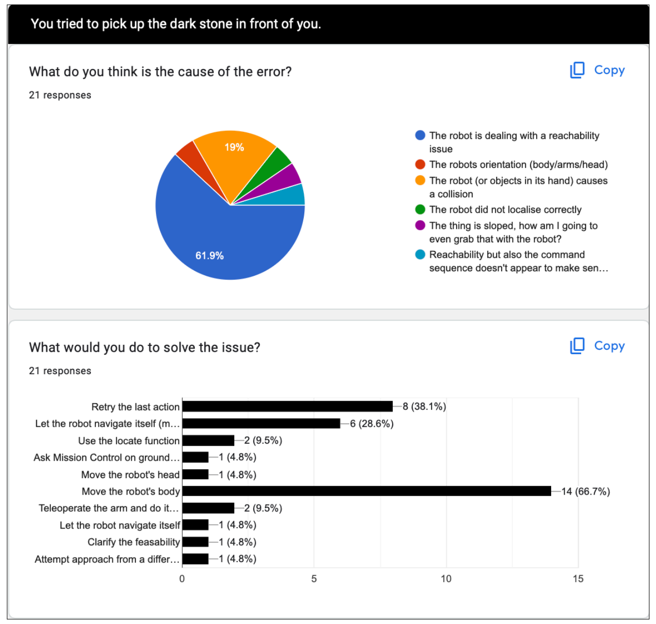






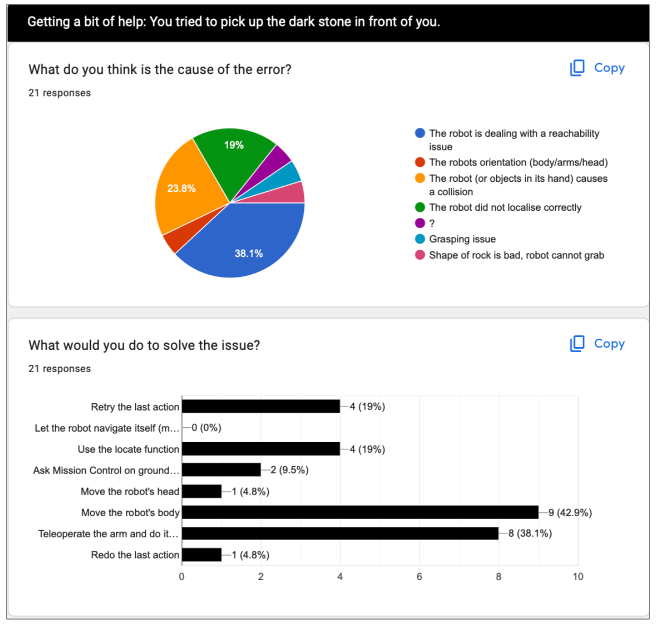


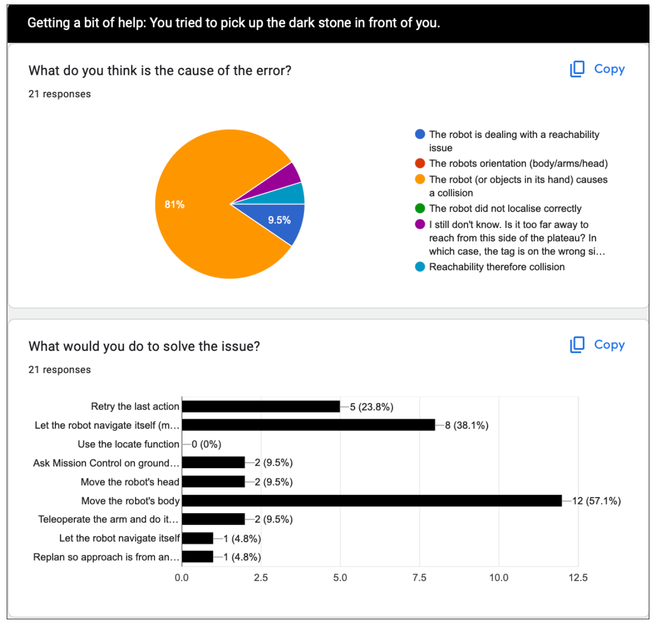


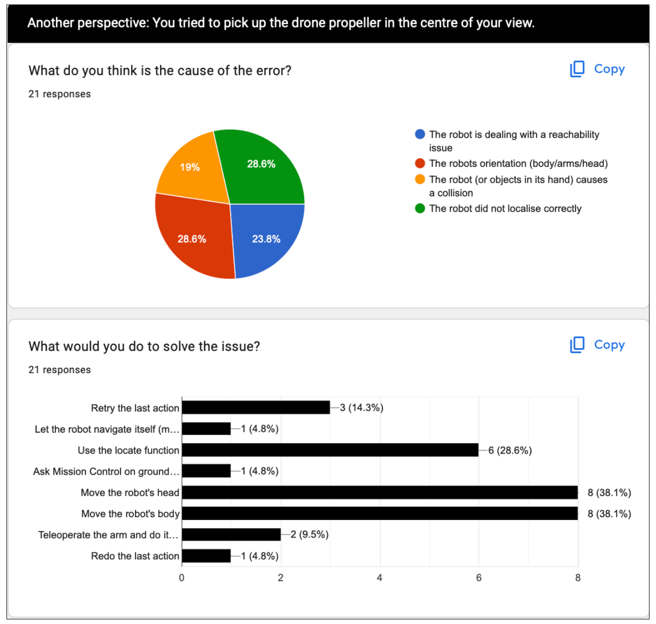


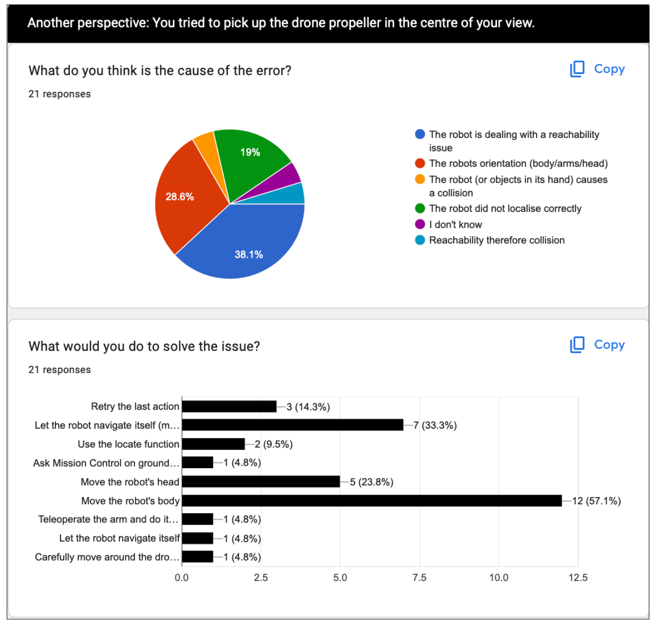
QUESTIONNAIRE A

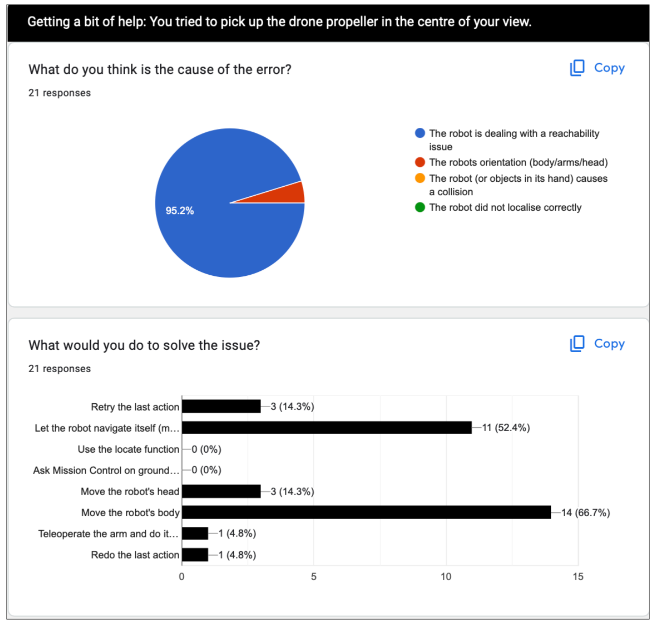
QUESTIONNAIRE B

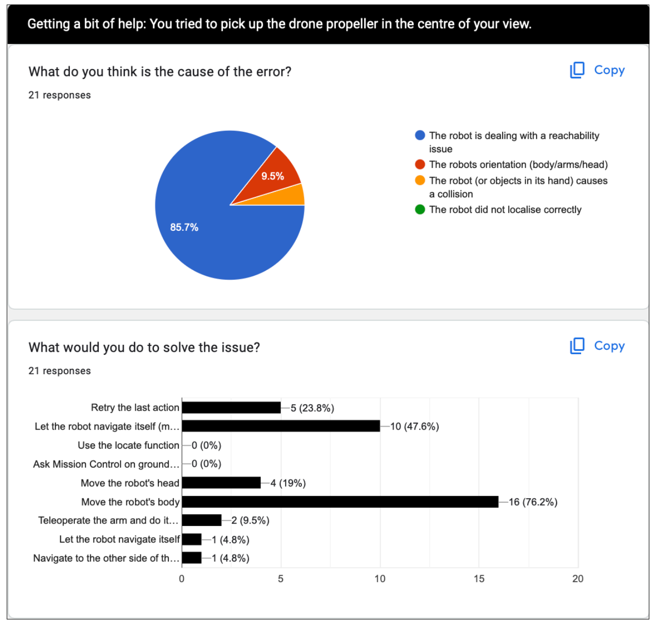












Consent Form for Session Participation

You are being invited to participate in a Master thesis research about using Rollin Justin. This study is being done by Liliane Filthaut from the TU Delft, Industrial Design Engineering in cooperation with the DLR.

The purpose of this research study is to understand how people use and perceive the interaction with Rollin Justin when handling errors and will take you approximately 1h to complete. The data will be used for analysing and designing improvements and suggestions for the current interaction, specifically in the case of planning errors. In this research, you will be asked to interact with and provide feedback on two prototypes aimed at improving error handling in a robot system, focusing on aspects such as user experience, difficulties encountered, suggested improvements, engagement, cognitive load, intuitiveness, and effectiveness in guiding error resolution.

1. By signing this consent form, I volunteer to participate in a research project conducted by the researcher from TU Delft. I understand that the project is designed to gather information about Justin in the context of teleoperation for academic and design purposes that will benefit the DLR.
2. My participation in this project is voluntary. I understand that I will not be paid for my participation. I may withdraw and discontinue participation at any time without penalty. If I decline to participate or withdraw from the study, no repercussions will be taken.
3. I will answer questions from the researchers. If, however, I feel uncomfortable in any way during the session, I have the right to decline to answer any question or to end the session.
4. During the session the researcher might take written notes. An audio and video tape of the interview will be recorded. If I don't want to be recorded, I will inform the researchers and if a middle ground cannot be made, I will not be able to participate in the study.
5. I understand that the researcher will not identify me by name in any reports using information obtained from this interview, and that my confidentiality as a participant in this study will remain secure. Subsequent uses of records and data will be subject to standard data use policies which protect the anonymity of individuals and institutions.
6. I have read and understand the explanation provided to me. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this study.
7. I have been given a copy of this consent form.

Date: ……/……/…….

Participant Name: …………………………… Researcher Present: ……………………………

Participant Signature: ………………………… Researcher Signature: …………………………

Plan Overview

A Data Management Plan created using DMPonline

Title: Designing supervised autonomy for astronaut-robot coaction in space

Creator: Liliane Filthaut

Affiliation: Delft University of Technology

Template: TU Delft Data Management Plan template (2021)

Project abstract:

This research is done in cooperation with the DLR (https://www.dlr.de/EN/Home/home_node.html). There is no funding.

For the purpose of designing a better user experience when encountering planning errors while interacting with a teleoperated robot (from DLR), roughly 10 people who tested the robot before will take part in research sessions to understand the current usage pattern and find potential areas of improvement. Participants are employees of DLR. Sessions will contain standard questionnaires, interview questions, cognitive walkthroughs and usage of the current system. Moreover, in a later stage, the same or similar participants (again 10) will get to try out prototypes that are directed at improving the user experience when handling errors.

ID: 119675

Start date: 16-02-2023

End date: 14-07-2023

Last modified: 29-03-2023

Designing supervised autonomy for astronaut-robot coaction in space

0. Administrative questions

1. Name of data management support staff consulted during the preparation of this plan.

The DPM for my project was discussed with my chair and mentor.
And my faculty data steward, Jeff Love, has reviewed this DMP on 29.03.23.

2. Date of consultation with support staff.

2023-03-29

I. Data description and collection or re-use of existing data

3. Provide a general description of the type of data you will be working with, including any re-used data:

Type of data	**File format(s)**	**How will data be collected (for re-used data: source and terms of use)?**	**Purpose of processing**	**Storage location**	**Who will have access to the data**
Anonymised Qualitative interview data	m4a, txt	Interviews, English and German	To deepen insights from the below questionnaires and explain user choices when interacting with the system	OneDrive	The company and the project team (the chair and mentor: Dr. Murray-Rust, D and Maria Luce Lupetti).
Anonymised Quantitative questionnaire	.csv file	Questionnaires Quantitative: UEQ & Godspeed Qualitative: PrEmo	To find quantified patterns in how people perceive and relate to the robot and teleoperation system.	OneDrive	The company and the project team (the chair and mentor: Dr. Murray-Rust, D and Maria Luce Lupetti).
Demographics (**usage experience:** amount and last usage | **Astronaut encounters:** amount | **Training:** participation/observations amount | **Part of team:** yes/no)	csv	Interviews, English and German	To see the effect of the collected demographics on the answers	OneDrive	The company and the project team (the chair and mentor: Dr. Murray-Rust, D and Maria Luce Lupetti).
Anonymised data on usage of system before adapting it concerning: navigation, object interaction and error handling task analysis	m4a, mov, csv, txt	Usage of existing robot system doing typical experiment or cognitive walkthrough of the typical experiment if prior knowledge and system unavailable	To find quantified patterns in handling errors in the system, understand users' perspective/awareness of the system and find opportunities for improvement	OneDrive	The company and the project team (the chair and mentor: Dr. Murray-Rust, D and Maria Luce Lupetti).
Anonymised data on usage of the system after adapting it concerning adaptation-specific changes.	m4a, mov, csv, txt	Usage of existing robot system with additional changes doing typical experiments or cognitive walkthrough of the typical experiment	To find quantified patterns in handling errors in the adapted system and evaluate different adaptation of the system	OneDrive	The company and the project team (the chair and mentor: Dr. Murray-Rust, D and Maria Luce Lupetti).

4. How much data storage will you require during the project lifetime?

- < 250 GB

II. Documentation and data quality

5. What documentation will accompany data?

- Methodology of data collection
- README file or other documentation explaining how data is organised

III. Storage and backup during research process

6. Where will the data (and code, if applicable) be stored and backed-up during the project lifetime?

- Another storage system - please explain below, including provided security measures
- OneDrive

OneDrive, provided by TU Delft

IV. Legal and ethical requirements, codes of conduct

7. Does your research involve human subjects or 3rd party datasets collected from human participants?

- Yes

8A. Will you work with personal data? (information about an identified or identifiable natural person)

If you are not sure which option to select, ask your Faculty Data Steward for advice. You can also check with the privacy website or contact the privacy team: privacy-tud@tudelft.nl

- Yes

8B. Will you work with any other types of confidential or classified data or code as listed below? (tick all that apply)

If you are not sure which option to select, ask your Faculty Data Steward for advice.

- No, I will not work with any confidential or classified data/code

9. How will ownership of the data and intellectual property rights to the data be managed?

For projects involving commercially-sensitive research or research involving third parties, seek advice of your Faculty Contract Manager when answering this question. If this is not the case, you can use the example below.

The datasets underlying the published papers will be publicly released following the TU Delft Research Data Framework. They will be released publicly no later than at the time of publication of corresponding research papers.
The data we collect and the results from it will be co-owned by myself, the project team: my chair Dr. Murray-Rust, D., mentor Maria Luce Lupetti, my company mentor Daniel Leidner and the DLR Surface avatar project.

10. Which personal data will you process? Tick all that apply

- Other types of personal data - please explain below
- Signed consent forms

I am collecting audio recordings during interviews, transcripts and observation notes of system usage.

11. Please list the categories of data subjects

Experts that are part of the Surface Avatar team at DLR.
Employees of DLR that have used the system in the past or are willing to use the system in a session.
Students at TU Delft for prototype testing.

Only if available, but unlikely: ESA Astronauts.

12. Will you be sharing personal data with individuals/organisations outside of the EEA (European Economic Area)?

- No

15. What is the legal ground for personal data processing?

- Informed consent

16. Please describe the informed consent procedure you will follow:

People I interview and participate in sessions using the robot system will sign an adapted form of the consent form provided by TU Delft.
before the start of the sessions.
In case of an online session, consent forms will be signed digitally and before the start of the session obtained verbally via an introduction statement.

17. Where will you store the signed consent forms?

- Same storage solutions as explained in question 6

18. Does the processing of the personal data result in a high risk to the data subjects?

If the processing of the personal data results in a high risk to the data subjects, it is required to perform a Data Protection Impact Assessment (DPIA). In order to determine if there is a high risk for the data subjects, please check if any of the options below that are applicable to the processing of the personal data during your research (check all that apply).
If two or more of the options listed below apply, you will have to complete the DPIA. Please get in touch with the privacy team: privacy-tud@tudelft.nl to receive support with DPIA.
If only one of the options listed below applies, your project might need a DPIA. Please get in touch with the privacy team: privacy-tud@tudelft.nl to get advice as to whether DPIA is necessary.
If you have any additional comments, please add them in the box below.

- None of the above applies

22. What will happen with personal research data after the end of the research project?

- Personal research data will be destroyed after the end of the research project

V. Data sharing and long-term preservation

27. Apart from personal data mentioned in question 22, will any other data be publicly shared?

- All other non-personal data (and code) produced in the project

- All other non-personal data (and code) underlying published articles / reports / theses

29. How will you share research data (and code), including the one mentioned in question 22?

- I will upload the data to another data repository (please provide details below)

Supplemental anonymized materials will be included as appendixes to the thesis published in the education repository.

31. When will the data (or code) be shared?

- As soon as corresponding results (papers, theses, reports) are published

VI. Data management responsibilities and resources

33. Is TU Delft the lead institution for this project?

- Yes, leading the collaboration - please provide details of the type of collaboration and the involved parties below

The DLR (German Aerospace Center), specifically the Surface Avatar team (https://www.dlr.de/rm/en/desktopdefault.aspx/tabid-18508/) collaborate in this research and are providers of the investigated robot system.

34. If you leave TU Delft (or are unavailable), who is going to be responsible for the data resulting from this project?

My chair Dr. Murray-Rust, D., mentor Maria Luce Lupetti as well as my cooperation partner Daniel Leidner from the DLR

35. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

None - I will handle all data management on my own using resources provided by the university.

Delft University of Technology
HUMAN RESEARCH ETHICS
CHECKLIST FOR HUMAN RESEARCH
(Version January 2022)

IMPORTANT NOTES ON PREPARING THIS CHECKLIST

1. An HREC application should be submitted for every research study that involves human participants (as Research Subjects) carried out by TU Delft researchers
2. Your HREC application should be submitted and approved **before** potential participants are approached to take part in your study
3. All submissions from Master's Students for their research thesis need approval from the relevant Responsible Researcher
4. The Responsible Researcher must indicate their approval of the completeness and quality of the submission by signing and dating this form OR by providing approval to the corresponding researcher via email (included as a PDF with the full HREC submission)
5. There are various aspects of human research compliance which fall outside of the remit of the HREC, but which must be in place to obtain HREC approval. These often require input from internal or external experts such as Faculty Data Stewards, Faculty HSE advisors, the TU Delft Privacy Team or external Medical research partners.
6. You can find detailed guidance on completing your HREC application here
7. Please note that incomplete submissions (whether in terms of documentation or the information provided therein) will be returned for completion **prior to any assessment**
8. If you have any feedback on any aspect of the HREC approval tools and/or process you can leave your comments here

I. Applicant Information

PROJECT TITLE:	**Designing supervised autonomy for astronaut-robot coaction in space**
Research period: *Over what period of time will this specific part of the research take place*	**15.02 – 14.07 2023**
Faculty:	**IDE/IO**
Department:	**HICD / DA**
Type of the research project: *(Bachelor's, Master's, DreamTeam, PhD, PostDoc, Senior Researcher, Organisational etc.)*	**Master's**
Funder of research: *(EU, NWO, TUD, other – in which case please elaborate)*	
Name of Corresponding Researcher: *(If different from the Responsible Researcher)*	**Liliane Filthaut**
E-mail Corresponding Researcher: *(If different from the Responsible Researcher)*	
Position of Corresponding Researcher: *(Masters, DreamTeam, PhD, PostDoc, Assistant/ Associate/ Full Professor)*	**Masters**
Name of Responsible Researcher: ***Note:** all student work must have a named Responsible Researcher to approve, sign and submit this application*	**Dr. Murray-Rust, D.**
E-mail of Responsible Researcher: *Please ensure that an institutional email address (**no Gmail, Yahoo, etc.**) is used for all project documentation/ communications including Informed Consent materials*	
Position of Responsible Researcher : *(PhD, PostDoc, Associate/ Assistant/ Full Professor)*	**Associate Professor**

II. Research Overview

***NOTE:** You can find more guidance on completing this checklist here*

a) Please summarise your research very briefly (100-200 words)

What are you looking into, who is involved, how many participants there will be, how they will be recruited and what are they expected to do?

Add your text here – (please avoid jargon and abbrevations)
This research is in cooperation with the DLR (https://www.dlr.de/EN/Home/home_node.html). There is no funding. For the purpose of designing a better user experience when encountering planning errors while interacting with a teleoperated robot (from DLR), roughly ten people who tested the robot before will take part in research sessions to understand the current usage pattern and find potential areas of improvement. Participants are employees of DLR. Sessions will contain standard questionnaires, interview questions, cognitive walkthroughs and usage of the current system. Moreover, in a later stage, the same or similar participants (again 10) will get to try prototypes that are directed at improving the user experience when handling errors.

b) **If your application is an additional project** related to an existing approved HREC submission, please provide a brief explanation including the existing relevant HREC submission number/s.

Add your text here – (please avoid jargon and abbrevations)

III. Risk Assessment and Mitigation Plan

***NOTE:** You can find more guidance on completing this checklist here*

Please complete the following table in full for all points to which your answer is "yes". Bear in mind that the vast majority of projects involving human participants as Research Subjects also involve the collection of **Personally Identifiable Information (PII)** and/or **Personally Identifiable Research Data (PIRD)** which may pose potential risks to participants as detailed in Section G: Data Processing and Privacy below.

To ensure alighment between your risk assessment, data management and what you agree with your Research Subjects you can use the last two columns in the table below to refer to specific points in your Data Management Plan (DMP) and Informed Consent Form (ICF) – **but this is not compulsory**.

It's worth noting that **you're much more likely to need to resubmit your application if you neglect to identify potential risks**, than if you identify a potential risk and demonstrate how you will mitigate it. If necessary, the HREC will always work with you and colleagues in the Privacy Team and Data Management Services to see how, if at all possible, your research can be conducted.

			If YES please complete the Risk Assessment and Mitigation Plan columns below.		*Please provide the relevant reference #*	
ISSUE	**Yes**	**No**	**RISK ASSESSMENT – what risks could arise?** *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	**MITIGATION PLAN – what mitigating steps will you take?** *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	**DMP**	**ICF**
A: Partners and collaboration						
1. Will the research be carried out in collaboration with additional organisational partners such as: • One or more collaborating research and/or commercial organisations • Either a research, or a work experience internship provider[1] *[1] If yes, please include the graduation agreement in this application*	X		**The DLR is a work experience internship provider** There may be concerns about the **confidentiality** of the data collected during the research. Moreover, there is a possibility that DLR may have its own agenda, which may **conflict** with the student's research **interests**.	**Confidentiality:** Any information shared with the researchers has already been published. Any insights gathered during the cooperation are planned to be published as well, there is no embargo in place. The research focus is moreover not on the employees but the robot usage, which is anonymized and participants are represented by numbers. However, In case that confidential data will still be collected during the cooperation the confidential appendix of the thesis report will be used. The company has their own networks and standards for data transfer to which the researcher will be trained and adhere to. This includes using only specific channels for data transfer like Gigamove or		

			If YES please complete the Risk Assessment and Mitigation Plan columns below.		*Please provide the relevant reference #*	
ISSUE	**Yes**	**No**	**RISK ASSESSMENT – what risks could arise?** *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	**MITIGATION PLAN – what mitigating steps will you take?** *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	**DMP**	**ICF**
				Mattermost and using only devices provided by the cooperation partner for storing confidential information. **Conflict of interest:** Any considerations are agreed upon in the graduation contract. To mitigate conflicts, any research at the facilities of the cooperation partner will be approved beforehand by the supervisor of the cooperation partner and weekly meetings are set in place so that all parties are always informed on further decisions and plans.		
2. Is this research dependent on a Data Transfer or Processing Agreement with a collaborating partner or third party supplier? *If yes please provide a copy of the signed DTA/DPA*		X				
3. Has this research been approved by another (external) research ethics committee (e.g.: HREC and/or MREC/METC)? *If yes, please provide a copy of the approval (if possible) and summarise any key points in your Risk Management section below*		X				
B: Location						
4. Will the research take place in a country or countries, other than the Netherlands, within the EU?	X		**Yes, Germany and Netherlands** **Cultural and language differences:** Conducting research in a foreign country may involve working with participants who speak a different language or have different cultural backgrounds. These differences may affect the validity and reliability of the research results. **Data protection:** Conducting research in a foreign country may raise concerns about data protection and privacy laws, particularly if personal data is being collected or transferred across borders. **Ethical considerations:** Conducting research in a foreign country may require adherence to different ethical standards than those of the Netherlands. **Logistical challenges:** Conducting research in a foreign country may involve logistical challenges, such as obtaining necessary permits or visas, finding	**Cultural and language differences:** The researcher's German nationality and upbringing allows them to speak the same language if participants prefer to participate in their mother-tongue. Moreover, any cultural differences are already part of the researcher's culture and should therefore not cause any complications. **Data protection:** No personal data will be collected. The company has their own networks and standards for data transfer to which the researcher will be trained and adhere to. This includes using only specific channels for data transfer like Gigamove or Mattermost and using only Devices provided by the cooperation partner for storing confidential information. **Ethical considerations:** Concerning human rights, data protection, principles for social sciences, research		

ISSUE	Yes	No	*If YES please complete the Risk Assessment and Mitigation Plan columns below.* RISK ASSESSMENT – what risks could arise? *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	MITIGATION PLAN – what mitigating steps will you take? *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	*Please provide the relevant reference #* DMP	ICF
			suitable research participants, and navigating unfamiliar environments.	practices, and ethics, the Netherlands and Germany share the same ethical standards. See: https://satoriproject.eu/media/D3.2-Int-differences-in-ethical-standards.pdf **Logistical challenges:** Considering the german nationality of the researcher, the listed logical challenges do not apply in this case.		
5. Will the research take place in a country or countries outside the EU?		X				
6. Will the research take place in a place/region or of higher risk – including known dangerous locations (in any country) or locations with non-democratic regimes?		X				
C: Participants						
7. Will the study involve participants who **may** be vulnerable and possibly (legally) unable to give informed consent? (e.g., children below the legal age for giving consent, people with learning difficulties, people living in care or nursing homes,).		X				
8. Will the study involve participants who **may** be vulnerable under specific circumstances and in specific contexts, such as victims and witnesses of violence, including domestic violence; sex workers; members of minority groups, refugees, irregular migrants or dissidents?		X				
9. Are the participants, outside the context of the research, in a dependent or subordinate position to the investigator (such as own children, own students or employees of either TU Delft and/or a collaborating partner organisation)? *It is essential that you safeguard against possible adverse consequences of this situation (such as allowing a student's failure to participate to your satisfaction to affect your evaluation of their coursework).*		X				
10. Is there a high possibility of re-identification for your participants? (e.g., do they have a very specialist job of which there are only a small number in a given country, are they members of a small community, or employees from a partner company collaborating in the research? Or are they one of only a handful of (expert) participants in the study?		X				
D: Recruiting Participants						
11. Will your participants be recruited through your own, professional, channels such as conference attendance lists, or through specific network/s such as self-help groups		X				

ISSUE	Yes	No	*If YES please complete the Risk Assessment and Mitigation Plan columns below.* RISK ASSESSMENT – what risks could arise? *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	MITIGATION PLAN – what mitigating steps will you take? *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	*Please provide the relevant reference #* DMP	ICF
12. Will the participants be recruited or accessed in the longer term by a (legal or customary) gatekeeper? (e.g., an adult professional working with children; a community leader or family member who has this customary role – within or outside the EU; the data producer of a long-term cohort study)		X				
13. Will you be recruiting your participants through a crowd-sourcing service and/or involve a third party data-gathering service, such as a survey platform?		X				
14. Will you be offering any financial, or other, remuneration to participants, and might this induce or bias participation?		X				
E: Subject Matter *Research related to medical questions/health may require special attention. See also the website of the CCMO before contacting the HREC.*						
15. Will your research involve any of the following: • Medical research and/or clinical trials • Invasive sampling and/or medical imaging • Medical and *In Vitro Diagnostic Medical* Devices Research		X				
16. Will drugs, placebos, or other substances (e.g., drinks, foods, food or drink constituents, dietary supplements) be administered to the study participants? *If yes see here to determine whether medical ethical approval is required*		X				
17. Will blood or tissue samples be obtained from participants? *If yes see here to determine whether medical ethical approval is required*		X				
18. Does the study risk causing psychological stress or anxiety beyond that normally encountered by the participants in their life outside research?		X				
19. Will the study involve discussion of personal sensitive data which could put participants at increased legal, financial, reputational, security or other risk? (e.g., financial data, location data, data relating to children or other vulnerable groups) *Definitions of sensitive personal data, and special cases are provided on the TUD Privacy Team website.*		X				
20. Will the study involve disclosing commercially or professionally sensitive, or confidential information? (e.g., relating to decision-making processes or business strategies which might, for example, be of interest to competitors)		X				
21. Has your study been identified by the TU Delft Privacy Team as requiring a Data Processing Impact Assessment (DPIA)? *If yes please attach the advice/ approval from the Privacy Team to this application*		X				
22. Does your research investigate causes or areas of conflict?		X				

ISSUE	Yes	No	*If YES please complete the Risk Assessment and Mitigation Plan columns below.* RISK ASSESSMENT – what risks could arise? *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	MITIGATION PLAN – what mitigating steps will you take? *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	*Please provide the relevant reference #* DMP	ICF
If yes please confirm that your fieldwork has been discussed with the appropriate safety/security advisors and approved by your Department/Faculty.						
23. Does your research involve observing illegal activities or data processed or provided by authorities responsible for preventing, investigating, detecting or prosecuting criminal offences *If so please confirm that your work has been discussed with the appropriate legal advisors and approved by your Department/Faculty.*		X				
F: Research Methods						
24. Will it be necessary for participants to take part in the study without their knowledge and consent at the time? (e.g., covert observation of people in non-public places).		X				
25. Will the study involve actively deceiving the participants? (For example, will participants be deliberately falsely informed, will information be withheld from them or will they be misled in such a way that they are likely to object or show unease when debriefed about the study).		X				
26. Is pain or more than mild discomfort likely to result from the study? And/or could your research activity cause an accident involving (non-) participants?		X				
27. Will the experiment involve the use of devices that are not 'CE' certified? *Only, if 'yes': continue with the following questions:*		X				
• Was the device built in-house?						
• Was it inspected by a safety expert at TU Delft? *If yes, please provide a signed device report*						
• If it was not built in-house and not CE-certified, was it inspected by some other, qualified authority in safety and approved? *If yes, please provide records of the inspection*						
28. Will your research involve face-to-face encounters with your participants and if so how will you assess and address Covid considerations?	X		**Infection increase:** Face-to-face encounters increase the risk of transmitting COVID-19 between participants and researchers. **Compliance with local regulations:** COVID-19 regulations and restrictions vary by location, Germany might have other rules. **Psychological and emotional well-being:** Participants may be experiencing additional stress or	Face-to-face encounters with the participants would happen regardless of the research, as they are coworkers or peers. **Infection increase can be mitigated by:** If necessary wearing masks, practising social distancing, and at all times maintaining good hygiene practices during all sessions. In case one of the parties was in contact with an infected person, the sessions need to be rescheduled or held online.		

ISSUE	Yes	No	*If YES please complete the Risk Assessment and Mitigation Plan columns below.* RISK ASSESSMENT – what risks could arise? *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	MITIGATION PLAN – what mitigating steps will you take? *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	*Please provide the relevant reference #* DMP	ICF
			anxiety due to the pandemic, and face-to-face encounters may exacerbate these feelings.	**Local regulations mitigations:** following local quarantine or testing requirements, and adhering to capacity limits for indoor spaces. **Psychological and emotional well-being mitigation:** It is important to create a supportive and safe environment, therefore offering the option to hold online sessions, wear masks or keep more distance will be offered to all participants. This will be agreed upon in advance. In the introduction statement, these options will be offered again to ensure that all participants are informed.		
29. Will your research involve **either:** a) "big data", combined datasets, new data-gathering or new data-merging techniques which might lead to re-identification of your participants **and/or** b) artificial intelligence or algorithm training where, for example biased datasets could lead to biased outcomes?		X				
G: Data Processing and Privacy						
30. Will the research involve collecting, processing and/or storing any directly identifiable PII (Personally Identifiable Information) including name or email address that will be used for administrative purposes only? (eg: obtaining Informed Consent or disbursing remuneration)	X		**Data breaches:** If the PII data is not securely stored and protected, there is a risk of a data breach, which can result in unauthorized access to sensitive personal information. **Misuse of personal information:** If PII data is obtained for administrative purposes only but is used for other purposes without the participant's consent, such as marketing or advertising, this can be a violation of the participant's privacy rights.	**Data breaches:** Communication between the researchers, client mentor, and participants will be conducted via the client's email or education email. Participants will be assigned a unique participant number upon enrollment to ensure that their personally identifiable information remains confidential and anonymous. No additional personally identifiable information will be collected or used during the course of the study. All study data will be securely stored in an electronic format and password-protected to prevent unauthorized access. Upon completion of the study, all data will be securely destroyed. The anonymized results of the study will only be used for academic purposes **Misuse of personal information:** All participants need to be informed about the use of their PII data, obtain		

			If YES please complete the Risk Assessment and Mitigation Plan columns below.		*Please provide the relevant reference #*	
ISSUE	**Yes**	**No**	**RISK ASSESSMENT – what risks could arise?** *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	**MITIGATION PLAN – what mitigating steps will you take?** *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	**DMP**	**ICF**
				their informed consent, and provide options for opting out of data collection or use. All collected data will solely be used for the purposes described in the agreements made with the participants.		
31. Will the research involve collecting, processing and/or storing any directly or indirectly identifiable PIRD (Personally Identifiable Research Data) including videos, pictures, IP address, gender, age etc and **what other Personal Research Data** (including personal or professional views) will you be collecting?	X		Professional knowledge on how the system works, professional views on astronauts using the system and what they personally observed when others interact with the system (no pictures of people, no gender and age). Moreover audio of sessions and interviews. **Risks:** **Breach of confidentiality:** If the data is not properly secured or handled, there is a risk of harm or embarrassment to the participants. **Data security risks:** There is always a risk of data security breaches when collecting, processing, and storing data, regardless of the content.	**Breach of confidentiality:** Participants' email addresses and signed consent forms will be stored separately from all other data obtained during the study to maintain confidentiality and prevent any linkage to individual participants. Upon completion of the study, all data will be securely destroyed to ensure that participant confidentiality is maintained. **Data security risks:** Participant ID numbers will be used to collect and store all raw data, which will be assigned to participants upon signing the consent form. All data will be securely stored in an electronic format and password-protected to prevent unauthorized access. The anonymized results of the study will only be used for academic purposes and		
32. Will this research involve collecting data from the internet, social media and/or publicly available datasets which have been originally contributed by human participants		X				
33. Will your research findings be published in one or more forms in the public domain, as e.g., Masters thesis, journal publication, conference presentation or wider public dissemination?	X		Master thesis with potential for journal publication. **Breach of confidentiality:** If the research data contains sensitive or confidential information, there is a risk that it could be disclosed through the publication process, even if the data is anonymized. **Reputational risks:** If the research results are controversial or are interpreted in a negative way, there is a risk that this could harm the reputation of the participants, the institution, or the researchers involved.	**Breach of confidentiality:** Before publication, the cooperation partner can decide if any of the information should be deemed confidential. In this case, the confidential appendix would be used. **Reputational risks:** To mitigate all participants will be anonymized by removing personal identifiers (both direct and indirect) that may lead to an individual being identified within the publication. **Intellectual property risks:** Only already published information will be used and referenced.		

			If YES please complete the Risk Assessment and Mitigation Plan columns below.		*Please provide the relevant reference #*	
ISSUE	**Yes**	**No**	**RISK ASSESSMENT – what risks could arise?** *Please ensure that you list ALL of the actual risks that could potentially arise – do not simply state whether you consider any such risks are important!*	**MITIGATION PLAN – what mitigating steps will you take?** *Please ensure that you summarise what actual mitigation measures you will take for each potential risk identified – do not simply state that you will e.g. comply with regulations.*	**DMP**	**ICF**
			Intellectual property risks: If the research involves the use of copyrighted materials, there is a risk of infringing on the rights of the copyright owner.			
34. Will your research data be archived for re-use and/or teaching in an open, private or semi-open archive?		X				

Date 31-Mar-2023
Contact person Dr. Cath Cotton, Policy Advisor
Academic Integrity
E-mail c.m.cotton@tudelft.nl

TU Delft Delft University of Technology

Human Research Ethics
Committee TU Delft
(http://hrec.tudelft.nl)

Visiting address
Jaffalaan 5 (building 31)
2628 BX Delft

Postal address
P.O. Box 5015 2600 GA Delft
The Netherlands

Ethics Approval Application: Designing supervised autonomy for astronaut- robot coaction in space
Applicant: Filthaut, Lilly

Dear Lilly Filthaut,

It is a pleasure to inform you that your application mentioned above has been approved.

Thanks very much for your submission to the HREC which has been approved. We do additionally note/advise the following:

1) Follow-up prototype testing can be addressed using an HREC amendment form, if/when it goes ahead.

In addition to any specific conditions or notes, the HREC provides the following standard advice to all applicants:

- In light of recent tax changes, we advise that you confirm any proposed remuneration of research subjects with your faculty contract manager before going ahead.
- Please make sure when you carry out your research that you confirm contemporary covid protocols with your faculty HSE advisor, and that ongoing covid risks and precautions are flagged in the informed consent with particular attention to this where there are physically vulnerable (eg: elderly or with underlying conditions) participants involved.
- Our default advice is not to publish transcripts or transcript summaries, but to retain these privately for specific purposes/checking; and if they are to be made public then only if fully anonymised and the transcript/summary itself approved by participants for specific purpose.
- Where there are collaborating (including funding) partners, appropriate formal agreements including clarity on responsibilities, including data ownership, responsibilities and access, should be in place and that relevant aspects of such agreements (such as access to raw or other data) are clear in the Informed Consent.

Good luck with your research!

Sincerely,

Project Brief

DESIGN FOR our future

TUDelft

IDE Master Graduation

Project team, Procedural checks and personal Project brief

This document contains the agreements made between student and supervisory team about the student's IDE Master Graduation Project. This document can also include the involvement of an external organisation, however, it does not cover any legal employment relationship that the student and the client (might) agree upon. Next to that, this document facilitates the required procedural checks. In this document:

- The student defines the team, what he/she is going to do/deliver and how that will come about.
- SSC E&SA (Shared Service Center, Education & Student Affairs) reports on the student's registration and study progress.
- IDE's Board of Examiners confirms if the student is allowed to start the Graduation Project.

USE ADOBE ACROBAT READER TO OPEN, EDIT AND SAVE THIS DOCUMENT

Download again and reopen in case you tried other software, such as Preview (Mac) or a webbrowser.

STUDENT DATA & MASTER PROGRAMME

Save this form according the format "IDE Master Graduation Project Brief_familyname_firstname_studentnumber_dd-mm-yyyy".
Complete all blue parts of the form and include the approved Project Brief in your Graduation Report as Appendix 1 !

family name: Filthaut
initials: LF — given name: Liliane
student number:
street & no.:
zipcode & city:
country:
phone:
email:

Your master programme (only select the options that apply to you):
IDE master(s): IPD — DfI — SPD
2nd non-IDE master: none
individual programme: - - (give date of approval)
honours programme: Honours Programme Master
specialisation / annotation: Medisign — Tech. in Sustainable Design — Entrepeneurship

SUPERVISORY TEAM **

Fill in the required data for the supervisory team members. Please check the instructions on the right !

** chair: Dr. Dave Murray-Rust — dept. / section: HICD
** mentor: Maria Luce Lupetti — dept. / section: DA
2nd mentor: Dr.-Ing. Daniel Leidner
organisation: German Aerospace Center (DLR)
city: Munich — country: Germany

comments (optional)

Chair should request the IDE Board of Examiners for approval of a non-IDE mentor, including a motivation letter and c.v..

Second mentor only applies in case the assignment is hosted by an external organisation.

Ensure a heterogeneous team. In case you wish to include two team members from the same section, please explain why.

TUDelft

Procedural Checks - IDE Master Graduation

APPROVAL PROJECT BRIEF

To be filled in by the chair of the supervisory team.

chair: Dr. Dave Murray-Rust — date: 16 - 02 - 2023 — signature: Dave Murray-Rust (Digitally signed by Dave Murray-Rust Date: 2023.02.16 16:48:44 +01'00')

CHECK STUDY PROGRESS

To be filled in by the SSC E&SA (Shared Service Center, Education & Student Affairs), after approval of the project brief by the Chair. The study progress will be checked for a 2nd time just before the green light meeting.

Master electives no. of EC accumulated in total: 24 EC — YES: all 1st year master courses passed

Of which, taking the conditional requirements into account, can be part of the exam programme: 24 EC — NO: missing 1st year master courses are:

List of electives obtained before the third semester without approval of the BoE

name: Robin den Braber — date: 21 - 02 - 2023 — signature: Robin den Braber (Digitaal ondertekend door Robin den Braber Datum: 2023.02.21 13:58:15 +01'00')

FORMAL APPROVAL GRADUATION PROJECT

To be filled in by the Board of Examiners of IDE TU Delft. Please check the supervisory team and study the parts of the brief marked **. Next, please assess, (dis)approve and sign this Project Brief, by using the criteria below.

- Does the project fit within the (MSc)-programme of the student (taking into account, if described, the activities done next to the obligatory MSc specific courses)?
- Is the level of the project challenging enough for a MSc IDE graduating student?
- Is the project expected to be doable within 100 working days/20 weeks ?
- Does the composition of the supervisory team comply with the regulations and fit the assignment ?

Content: APPROVED — NOT APPROVED
Procedure: APPROVED — NOT APPROVED

comments

name: Monique von Morgen — date: 06 - 03 - 2023 — signature:

Initials & Name: LF Filthaut 6241 — Student number: 5609402

Title of Project: Designing supervised autonomy for astronaut-robot coaction in space

Designing supervised autonomy for astronaut-robot coaction in space — project title

Please state the title of your graduation project (above) and the start date and end date (below). Keep the title compact and simple. Do not use abbreviations. The remainder of this document allows you to define and clarify your graduation project.

start date 16 - 02 - 2023 — 14 - 06 - 2023 end date

INTRODUCTION **

Please describe, the context of your project, and address the main stakeholders (interests) within this context in a concise yet complete manner. Who are involved, what do they value and how do they currently operate within the given context? What are the main opportunities and limitations you are currently aware of (cultural- and social norms, resources (time, money,...), technology, ...).

The conditions in planetary exploration are hazardous and draining for astronauts. Therefore, robots are being developed so they can do physical tasks on the surface of other planets for them. However, these robots need a certain amount of teleoperation, as they are not fully autonomous yet. Considering the time delay one would have if such robots would be controlled from the earth, the current approach is to let astronauts control those robots from orbit, consequently decreasing the delay.

Figure 1 describes the parties involved in this process: Astronauts, the robot(s), researchers working on the robot, engineers developing the robot, mission control during teleportation and the two main agencies involved DLR and ESA.
The process works as follows: An astronaut can control a robot remotely from orbit to fulfil missions. For example, X. In the case of this thesis, the robot in question is Justin Rollin, a semi-autonomous humanoid service robot.
This robot can be controlled via three modes (figure 2, (1), (2) and (3)). The main focus of thesis concerns the teleoperation via the screen, figure 2, (1).
The touch interface allows the astronaut to select objects that the robot sees (like the mug in figure 2 (1)) and then select from actions that the robot could perform (like picking up the mug). The only thing the astronaut has to do in this case is select an action and wait for the robot to fulfil it (figure 2, (1)). It might take the robot some time to figure out how to perform the command, the astronaut has to wait during this time. If the action succeeds, the process restarts and the astronaut can select the next action for the robot. However, the robot does not always figure out how to fulfil a commanded action.

Here lays the main opportunity and assignments for the thesis. The interface currently only notifies the astronaut that the mission could not be fulfilled. It neither explains why nor does it provide solution directions. In case of failure, the astronaut can switch to the other modes of teleoperation (figure 2, (2) and (3)) or start communication with mission control (figure 1). Switching to the other modes of control and therefore manually operating the robot only works on short time delays and contacting earth would result in up to 40 minutes com-time dealy. Ideally, the astronaut should mainly interact with the touch interface. The researcher's goal is to decrease the mental load astronauts are already experiencing due to the conditions in space and make the teleoperation of Justin Rollin as user-friendly as possible. In the future, not just one, but multiple robots should be controlled by this interface.

The detection of execution errors is a current limitation of this project (this is ongoing research, currently the astronaut gets only notified over planning errors, which means that the robot simply does nothing and says it could not figure out a solution). Moreover, the environmental conditions that make teleoperations harsh for both the astronaut and the robot, as well as the unavoidable time delay which, however, has less of an impact on semi-autonomous control are part of the limitations. Additionally, the waiting time for the astronaut when the robot is "thinking" and the long intervals between which this system is used are limiting factors. The latter requires the system to be easy to understand. While astronauts get training for about 1hour prior to flight, usually at least 6 months pass until the system gets used again. Even tough, astronauts receive a 15 minute quick introduction in flight, this is all the astroinauts know about the software and one can therefore assume that their knowledge of the system is very basic.

space available for images / figures on next page

introduction (continued): space for images

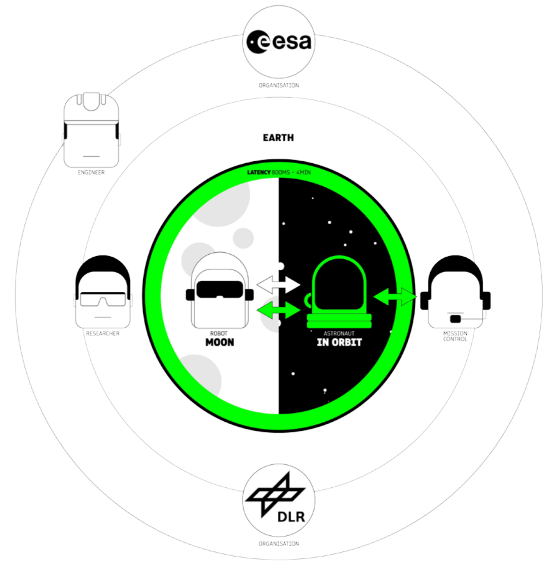


image / figure 1: Stakeholdermap including the stakeholders position

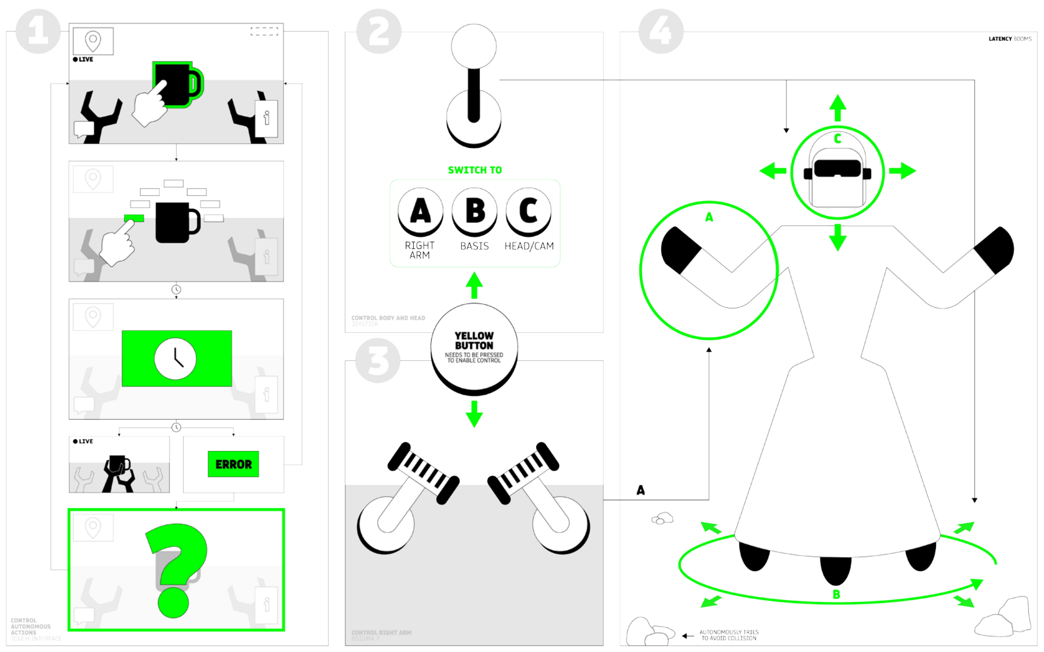


image / figure 2: (1) Interface process, (2) Joystick and mode buttons, (3) Sigma.7 control and (4) effect on robot

PROBLEM DEFINITION **

Limit and define the scope and solution space of your project to one that is manageable within one Master Graduation Project of 30 EC (= 20 full time weeks or 100 working days) and clearly indicate what issue(s) should be addressed in this project.

The main problem of this project is the difficulty that astronauts experience in using and controlling the robot Rollin Justin, particularly in the event of errors. The scope of the project will focus on the touch interface and error handling, to improve the user experience and minimize the need for other controls.

A central objective of the thesis is to develop simple and clear communication within the interface that can help astronauts understand and handle errors in the robot, without feeling confused or overwhelmed. This will involve researching the current user experience of the astronaut, and using that information to design a more effective and user-friendly interface. Additionally, the project will aim to improve the transparency of the robot's actions and limitations, making it easier for the astronaut to comprehend why the robot is unable to complete a task, as the actions that the astronaut commands the robot may be simple for humans but complex for the robot to execute. For example, errors might occur due to factors like distance or orientation towards an object.

Another important consideration is that astronauts receive training on the teleoperation of the robot, but a significant amount of time may pass before they use it. Therefore, the interface should be designed so that it can be easily understood without requiring lots of prior knowledge or frequent refresher training.

While some existing solutions or explanations for certain errors may not be possible, the goal is still to make it as easy as possible for the astronaut to understand and handle the robot's issues, and help them find a solution on their own without needing to contact support.

ASSIGNMENT **

State in 2 or 3 sentences what you are going to research, design, create and / or generate, that will solve (part of) the issue(s) pointed out in "problem definition". Then illustrate this assignment by indicating what kind of solution you expect and / or aim to deliver, for instance: a product, a product-service combination, a strategy illustrated through product or product-service combination ideas, ... In case of a Specialisation and/or Annotation, make sure the assignment reflects this/these.

the aim is to develop an easy way of communicating the error of the robot and ideally propose solution directions to the astronaut so that they can take steps to fix the issue and easily restart the process, likely in the form of a UI intervention. To arrive at this goal, the research will utilise think-aloud protocols, cognitive walkthroughs, task analysis, co-creation and potentially machine learning and data processing techniques.

1) Research will concentrate on the user experience of astronauts when interacting with the robot, with a focus on understanding and handling failures.

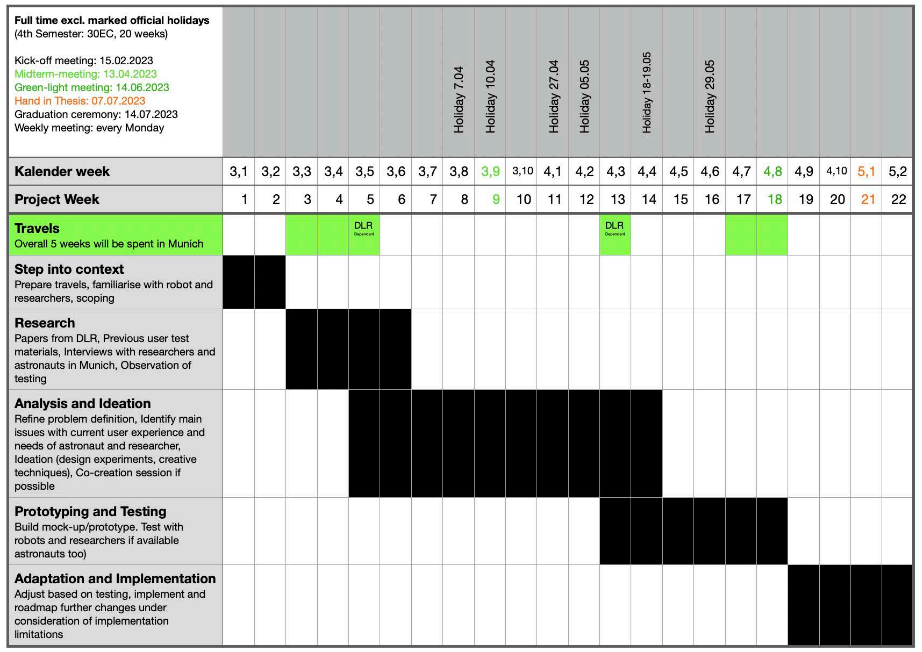
The initial phase will involve developing a contextual understanding of the current user experience as well as the technical aspects of the system. Depending on the available data, machine learning and data processing techniques may be employed to identify common error patterns. To ensure that the final design is easily understandable (also for amateurs), testing will also be done with the general public. For testing with astronauts, a think-aloud protocol will be proposed as a research method. This should help in understanding the cognitive processes that users go through when interacting with the robot and identify areas of confusion or difficulty. For testing with the general public, task analysis can be used to identify subtasks in the robot's planning process. Additionally, cognitive walkthroughs can be conducted with participants to evaluate the planning process, both at the beginning and end of the project. Furthermore, involving the general public in the testing of initial ideas and co-creation methods can provide valuable insights into the user experience.

2. Designing an intervention to guide the user when the robot fails to find a solution. Key elements: transparency (as far as possible), easily understandable communication, considering the astronaut's emotions, and making the operation of the robot more fluent. This intervention will likely be, depending on the research, a UI design.

PLANNING AND APPROACH **

Include a Gantt Chart (replace the example below - more examples can be found in Manual 2) that shows the different phases of your project, deliverables you have in mind, meetings, and how you plan to spend your time. Please note that all activities should fit within the given net time of 30 EC = 20 full time weeks or 100 working days, and your planning should include a kick-off meeting, mid-term meeting, green light meeting and graduation ceremony. Illustrate your Gantt Chart by, for instance, explaining your approach, and please indicate periods of part-time activities and/or periods of not spending time on your graduation project, if any, for instance because of holidays or parallel activities.

start date 16 - 2 - 2023 14 - 6 - 2023 end date



MOTIVATION AND PERSONAL AMBITIONS

Explain why you set up this project, what competences you want to prove and learn. For example: acquired competences from your MSc programme, the elective semester, extra-curricular activities (etc.) and point out the competences you have yet developed. Optionally, describe which personal learning ambitions you explicitly want to address in this project, on top of the learning objectives of the Graduation Project, such as: in depth knowledge a on specific subject, broadening your competences or experimenting with a specific tool and/or methodology, ... Stick to no more than five ambitions.

My personal interest in this project stems from an ongoing interest in science aside from design which has been present since early childhood. I position myself as a designer of making communication effortlessly and an intermediary between complex contexts and users. This was already a major part of my bachelor thesis on physics education and creating an interactive platform for more understandable communication of complex school subjects. (This thesis also drew the German aerospace centre's attention to me).
My bachelor thesis also showed me how many of the technologies we use daily originated from space travel like GPS or the downsizing of computer parts. Helping in the development of further space travels therefore appears very beneficial to me, even if my impact might be indirect or minimal. The thesis combines my interest in science (incl. topics like outer space), making complex contents easier and interactive technology design.

Moreover, most of my recent university projects deal with the topic of outer space. In Lighting Design, I explored lighting for future extraterrestrial habitats, which included intense research on outer-space requirements and conditions. In the course Design For Children's play, I examined how one can make outer space experiential through sound. Other projects of mine focused less on outer space but on communication. In Exploring Interaction I developed a tool to make otherwise intimidating communication easier for Makerspace users. On the other hand, in Interactive Technology Design, I (with my teammate) developed an interactive experience to make AI and its more hidden and complex consequences more tangible and easier understandable.

I plan to combine my knowledge of design and communication from my bachelor's and my acquired learning on design research from the last three semesters of my master's. The bachelor's taught me about gestalt principles, attention patterns, design standards and interface design. It further made me proficient in many design tools. My master's taught me how to derive existing user experiences and perspectives and build new, better-fitting ones, as well as many design research methods and creative techniques. By combining both perspectives, I hope to derive a meaningful and fitting solution for the problem at hand.

FINAL COMMENTS

In case your project brief needs final comments, please add any information you think is relevant.

Initials & Name LF Filthaut 6241 Student number 5609402

Title of Project Designing supervised autonomy for astronaut-robot coaction in space

H: More on Informed Consent and Data Management

NOTE: You can find guidance and templates for preparing your Informed Consent materials) here

Your research involves human participants as Research Subjects if you are recruiting them or actively involving or influencing, manipulating or directing them in any way in your research activities. This means you must seek informed consent and agree/ implement appropriate safeguards regardless of whether you are collecting any PIRD.

Where you are also collecting PIRD, and using Informed Consent as the legal basis for your research, you need to also make sure that your IC materials are clear on any related risks and the mitigating measures you will take – including through responsible data management.

Got a comment on this checklist or the HREC process? You can leave your comments here

IV. Signature/s

Please note that by signing this checklist list as the sole, or Responsible, researcher you are providing approval of the completeness and quality of the submission, as well as confirming alignment between GDPR, Data Management and Informed Consent requirements.

Name of Corresponding Researcher (if different from the Responsible Researcher) (print)
Liliane Filthaut

Signature of Corresponding Researcher:

Date: 22/03/2023

Name of Responsible Researcher (print)
Dr. Murray-Rust, D.

Signature (or upload consent by mail) Responsible Researcher:

Date: 22/03/2023