MSc Thesis

Design of an AR-IoT Tool for Future Human Space Exploration

F.A.A.S.D.T. Rometsch 4277953

July 20, 2020

TU Delft

MSc Thesis
Design of an AR-IoT Tool for Future Human Space Exploration

by

F.A.A.S.D.T. Rometsch

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Monday August 17, 2020 at 10:00.

Student number:	4277953	
Project duration:	October 1, 2019 – July 20, 2020	
Thesis committee:	Dr. J. Guo	Delft University of Technology, Chair
	Dr. Ir. J. C. F. de Winter	Delft University of Technology, Supervisor
	Ir. R. Noomen	Delft University of Technology
	Dr. A. Cowley	European Space Agency

An electronic version of this thesis is available at http://repository.tudelft.nl/.

Front image: Apollo 17 Astronaut Harrison Schmitt below Tracy's Rock, retrieved from
http://www.apolloarchive.com/apollo_gallery.html

TU Delft

To David Bowie

Abstract

Humans are embarking on a new era of space exploration with the plan of sending crewed spacecraft beyond Low Earth Orbit (LEO), to the Moon, Mars and beyond. NASA is committed to land astronauts on the lunar surface by the year 2024. The goal of NASA's lunar exploration program - Artemis, a collaboration with commercial and international partners including the European Space Agency (ESA), is to establish sustainable exploration by the end of this decade. The plan is to use what is learned on and around the Moon to take the next giant leap, namely sending astronauts to Mars. Activities planned during the Artemis missions, especially in the early phases, involve finding critical resources needed for long-term exploration, and acquiring more knowledge on Moon, Earth and the universe by carrying out experiments. All these activities will involve extensive lunar geological field work, which is orders of magnitude more complex than field geology on Earth. Extravehicular activities (EVAs) will become increasingly more complicated than the tasks executed during the early Artemis missions and generally during human spaceflight missions so far. EVA systems and crewmember skills that currently do not exist will be required. This plan entails many challenges as real-time support from ground control cannot be provided to astronauts who thus need to become more autonomous. Hence, modern human-machine interfaces have to be designed to support astronauts during their deep space missions.

Augmented reality (AR) and the internet-of-things (IoT) are changing the way industries work, especially AR has found application for space applications, specifically for procedural work. Nevertheless, only one AR space-related study focused on the use of AR for future human planetary exploration, namely on navigation and traverse planning. While cuff-checklists guided Apollo astronauts on the Moon, wrist displays and tablets represent the standard tools during today's astronaut analog EVA missions. However, these are often operationally unfeasible as crew has to handle several tools simultaneously and/or repeatedly look at the display and thus gets distracted from the surroundings leading to a potential loss of situational awareness, affecting their safety. Based on the research that is currently being performed on IoT technologies in combination with AR for visualisation and enhanced situational awareness purposes, the benefits obtained through the use of these technologies applied to future human extraterrestrial surface EVAs, more specifically geological site inspections, were explored in this research.

The AR-IoT surface exploration tool developed for this research introduces a new approach for astronauts to carry out geological site inspections. The tool enables hands-free operations such as data logging, detailed photo-documentation, taking site coordinates, descriptions of sites through the presence of a verbal "field notebook", as well as mapping and highlighting features during a traverse by creating waypoints, while providing crew with suit diagnostics.

A user-centered design method was adopted to design the AR-IoT tool deployed on the Microsoft HoloLens. This highly iterative design process involved two to three expert reviews for each of the first three concepts, and three heuristic evaluations for the fourth concept until the subsequent generation of the first prototype. Key usability and user interaction aspects, pertinent capabilities determining the adoption of innovative interfaces, essential insights into future human-machine interaction and design requirements for AR meant for EVA astronauts were gathered through semi-structured interviews held with four ESA astronauts and astronaut geological field activities experts. The interviews together with the qualitative and quantitative data collected through the questionnaires were then used to assess the usability of the AR-IoT tool. Moreover, these data provided with additional knowledge on user-centered design AR studies in general but also and particularly on user-centered AR space-related studies.

Valuable insights into interface design and user interaction aspects were gained. Results from the qualitative content analysis of the interviews stressed the importance of user satisfaction (32% of 139 quotes) as a usability aspect. Key design factors identified were: displaying solely important information in the field-of-view while adjusting it to the user's visual acuity, easy usage, simplicity, helpfulness and extensibility. User interaction was the second most mentioned (24% of 139 quotes) aspect. While multimodal interaction was considered feasible, no conclusions could be drawn on the most suitable combination of inputs. Nonetheless, most experts defined voice the most intuitive input. Based on the positive feedback from ESA astronauts and other experts, the AR-IoT proof of concept proved to be a potentially usable tool for future geological site inspection activities. The AR-IoT tool is therefore a promising asset for analogue training missions, such as Caves & Pangaea, and in the future for lunar geological field work. While limitations in both research and design are outlined, a set of recommendations aimed at warranting future testing and development of a more advanced AR-IoT tool for astronaut geological field activities is provided.

Preface

These nine months as a visiting researcher at the European Astronaut Centre (EAC), in Cologne, have been one of the most memorable and valuable experiences I have made so far. As an Aerospace Engineer and MSc student in Space Robotics, as I like to define the specialization of my Double M.Sc. Degree, being able to conduct research at the European Space Agency (ESA) and more specifically at the European Astronaut Centre (EAC), represented a dream come true for me. The conducted research is about the usability of an Augmented Reality (AR) - Internet-of-Things (IoT) tool for future human lunar and planetary surface exploration activities. I selected this topic because of my profound interest in the development of innovative technologies, in particular applications for future human spaceflight missions.

Here at the EAC, I have been part of the initiative Spaceship EAC, a think-tank composed of international students and researchers working on the development of innovative technologies for future space exploration, focusing in particular on the Moon.

I am extremely grateful for this experience as I have been able to evolve professionally in the space robotics sector. In that regard, I would like to thank my supervisor Dr. Aidan Cowley, Science Advisor and team-lead of Spaceship EAC, for granting me with the opportunity of being part of this extremely diverse and inspiring team and for his time and support during my time here.

I would like to extend my gratitude to my supervisors Dr. Joost de Winter and Dr. Jian Guo for their guidance and advice in conducting this thesis.

I appreciate Dr. Andrea Casini, Anne Drepper and Eóin Tuohy for their constant support and guidance during my time as part of the Spaceship EAC team.

Many thanks to Leonardo Turchi for his useful insights on the Caves & Pangaea training and to my favourite product designer Emilia Rosselli del Turco.

I would like to thank Dr. Mikael Wolff and Marco Carrano for their valuable information on the MobiPV project and Hervé Stevenin for his useful insights on EVA training.

I would like to extend my gratitude to Martial Costantini, thank you so much for your support and for making sure that I was all set in terms of hardware and software.

Many thanks to all the test participants for making this thesis possible.

Special thanks to all my friends across the world for their unconditional support and the unforgettable fun time together.

Inoltre, ringrazio la mia mamma che ha sempre seguito le mie avventure e i miei progetti con entusiasmo e mi ha dato supporto in momenti difficili, è soprattutto grazie a lei che ho raggiunto questo traguardo. Grazie ovviamente anche al mio adorato cane Oscar.

Por ultimo, me gustaría agradecer a mi querido Arturo por su apoyo incondicional, consejos y por enseñarme español durante el confinamiento.

Flavie Rometsch
Cologne, July 2020

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
AIT	Assembly Integration and Testing
AIV	Assembly Integration and Verification
ALSEP	Apollo Lunar Surface Experiment Package
ALTEC	Aerospace Logistics Technology Engineering Company
API	Application Programming Interfaces
AR	Augmented Reality
ARAMIS	Augmented Reality Application for Maintenance, Inventory and Stowage
ARML	Augmented Reality Markup Language
ARPASS	Augmented Reality in Product Assurance & Safety Study
BASALT	Biologic Analog Science Associated with Lava Terrains
BH	Biosensor Harness
BLE	Bluetooth Low Energy
BVP	Blood Volume Pulse
CAD	Computer-Aided Design
CAI	Computer Assisted Instructions
CapCom	Capsule Communicator
ConOps	Concept of Operations
Covid-19	Corona Virus 2019
CSA	Canadian Space Agency
CPU	Central Processing Unit
DLR	German Aerospace Centre
DOF	Degrees of Freedom
DRATS	Desert Research And Technology Studies
EAC	European Astronaut Centre
EAMD	Exploration Analog and Mission Development
ECG	Electrocardiogram
EdcAR	Engineering data in cross platform AR
EEG	Electroencephalogram
EFB	Electronic Field Book
EKG	Electrocardiography
ESA	European Space Agency
ESOC	European Space Operations Centre
EV	Extravehicular
EVA	Extravehicular Activity
FDIR	Fault Detection, Isolation, and Recovery
FOV	Field-of-View
GPIO	General Purpose Input Output
GPS	Global Positioning System
GSR	Galvanic Skin Response
GSTP	General Support Technology Program
GUI	Graphical User Interface
HCD	Human Centered Design
HCI	Human Computer Interaction

HIDH	Human Integration Design Handbook
HIDP	Human Integration Design Processes
HLS	Human Landing System
HMD	Head Mounted Display
HMI	Human Machine Interaction
HPU	Holographic Processing Unit
HUD	Head-Up Display
ID	Identifier
IMS	Inventory Management System
IMU	Inertial Measurement Unit
IoT	Internet-of-Things
IoST	Internet of Space Things/CubeSats
IQR	Interquartile Range
ISO	Inventory and Stowage Officers
ISS	International Space Station
IV	Intravehicular
IVA	Intravehicular Activity
JAXA	Japanese Aerospace Exploration Agency
JPL	Jet Propulsion Laboratory
JSC	Johnson Space Center
JSON	Java Script Object Notation
KML	Keyhole Markup Language
LEO	Low Earth Orbit
LM	Lander Module
LoRa	Long Range
LPWA	Low Power Wide Area
LRV	Lunar Roving Vehicle
MARVIN	Multi-purpose Astronaut and Remote Vehicle INteraction
MCC	Mission Control Centre
MCH	Cooper-Harper Scale
Mdn	Median
METERON	Multi-purpose End To End Robotic Operation Network
MQTT	Message Queuing Telemetry Transport
MobiPV	Mobile Procedure Viewer
MOE	METERON Operations Environment
MOM	Message-oriented Middleware
MOON	asseMbly Oriented authOring augmeNted reality
MOPS	Multi-purpose Operations Software
MR	Mixed Reality
MRS	Meteron Robotic System
MRTK	Mixed Reality Toolkit
MS	Microsoft
MSC	Mission Support Centre
M2M	Machine-to-Machine
NASA	National Aeronautics and Space Administration
NEEMO	NASA Extreme Environment Mission Operations
NFC	Near Field Communication
NFT	Natural Feature Tracking
ODF	Operation Data File
OW	Overall Workload

Pangaea-X	Pangaea-eXtension
PLRP	Pavilion Lake Research Project
PLSS	Portable Life Support System
POI	Point of Interest
QCA	Qualitative Content Analysis
QUIS	Questionnaire for User Interaction Satisfaction
R	Respondent
RFID	Radio Frequency IDentification
ROS	Robotics Operating System
RSME	Rating Scale Mental Effort
RSS	Received Signal Strength
SAR	Spatial Augmented Reality
SCICOM	Science Communicator
SDK	Software Development Kit
SEXTANT	Surface Exploration Traverse Analysis and Navigational Tool
SGUS	Smart Glasses Usability Scale
SOA	Service Oriented Architecture
SSP	Space Station Program
SSQ	Simulator Sickness Questionnaire
SSR	Science Support Room
SPSS	Statistical Package for the Social Sciences
SUS	System Usability Scale
SWANS	Sensor Wireless Actuator Networks in Space
SWAT	Subjective Workload Assessment Technique
TA	Thematic Analysis
TAMARA	Technology Acceptance Model for AR/WT
TAS	Thales Alenia Space
TLX	Task Load Index
TM	Telemetry
TM	Transfer Mechanism
TOP	Targets of Opportunity
TRL	Technology Readiness Level
UI	User Interface
VR	Virtual Reality
VTT	Technical Research Centre of Finland
WEAR	WEarable Augmented Reality
WEKIT	Wearable Experience for Knowledge Intensive Training
WIM	World in Miniature
WSN	Wireless Sensor Network
XML	Extensible Markup Language

1 Introduction

Humans are embarking on a new era of space exploration with the plan of sending crewed spacecraft beyond Low Earth Orbit (LEO) to the Moon, Mars and beyond. NASA is committed to land astronauts on the Moon by 2024 [179]. The goal of NASA's lunar exploration program, Artemis to which the European Space Agency (ESA) and several other international space agencies such as the Japanese Space Agency, JAXA, the Canadian Space Agency, CSA, and the Russian Space Agency, ROSCOSMOS, are contributing, is to establish sustainable missions by 2028. In fact, these missions do not have the aim of "putting boots on the Moon" again but to "go forward to the Moon" and use what is learned on and around the Moon to take the next giant leap, namely sending astronauts to Mars [179].

Activities planned during the missions of the Artemis program, especially in the early phases, involve finding and using water and other critical resources needed for long-term exploration, investigating the Moon's mysteries and acquiring more knowledge on Earth and the universe by carrying out experiments [179]. All these activities involve lunar field geology which is orders of magnitude more complex than field geology on Earth. In contrast to a field geologist on Earth, a lunar field geologist must always be aware that time is limited, consumables are limited, fatigue can be fatal, and that, generally, going back to a location is unlikely [81]. All of these components add to the normal intellectual workload of doing field work. Furthermore, tasks take longer in a spacesuit, if they can be completed at all. Writing notes as well as drawing diagrams by hand is impossible while wearing pressurized gloves. Planning, automated location, voice communications and the documentation of sampling activities, as well as the whole field experience are of even greater importance than usual [81].

Extravehicular activities (EVAs) will become increasingly more complicated than the tasks executed during the early Artemis missions and generally during human spaceflight missions so far. In fact, EVAs will involve building crew bases, laboratories and facilities [48]. EVA systems and crewmember skills that currently do not exist will be required. When communication latencies will reach up to 40 minutes, as will be the case for Martian operational scenarios, astronauts will no longer be able to rely on ground support to know where their tools are located or what their location is, as it was the case during the Apollo missions [37].

To give an example of EVA crew and CapCom (Capsule Communicator) communication regarding their tool location, an excerpt of voice transcript of the Apollo 17 mission is reported below [167]. During this mission, astronaut Eugene A. Cernan was the CDR (Commander) while Robert A. Parker (Bob) was the CC (Capsule Communicator, CapCom).

> ***146:30:19*** *Cernan: Say, Bob, where can I get a new set of bags?*
> ***146:30:23*** *Parker: Okay, you want new bags...They'll be under Jack's seat.*

In fact, MCC (Mission Control Centre) could monitor EVA crew activities through the Lunar Surface TV which was either connected to the LM (Lander Module) by cable or mounted on the ground-controlled television assembly on the LRV (Lunar Roving Vehicle) and could be controlled remotely [142].

CapCom, the Mission Control Centre (MCC) in Houston, played a major role when it comes to instructional guidance, tracking, logging, monitoring, advising and scheduling. From the voice transcripts of Apollo 17, it could be concluded that EVA crew interacted with CapCom during their EVA when [37]:

- logging (reading out information)
- asking about the location of materials
- reporting descriptions on the surroundings or on the equipment status
- suggesting, requesting or documenting EVA equipment settings and usage

CapCom managed the EVA actively by [37]:

- communicating elapsed time, remaining time at a specific site, warnings when turning around
- communicating revised plans for reprioritizing, skipping or substituting work
- communicating navigation advice and identifying sites the crew is seeing

Nevertheless, some issues in communication with CapCom occurred with respect to verbal exchange, tracking and planning. Identified issues were [37]:

- Crew mishearing or not listening to CapCom or other crewmembers
- Crew not understanding who is talking to whom
- Communication breakdowns
- Unnecessary remarks from CapCom leading to disruption
- Non-optimal scheduling decisions made by CapCom, as EVA crew was seeing opportunities CapCom could not see
- Misunderstandings between rover indicator and checklist

Moreover, to carry out any kind of operation during future human spaceflight missions, procedures being displayed on cuff-checklists (see Figure 1.1 [219]) will no longer be sufficient as means of crew guidance and support as they were for crew during the Apollo missions [118]. More autonomy will be required from EVA crew to manage both systems performance and mission execution. The crew residing in the crew base will play an increasingly important role in assisting crew during EVAs, taking on activities previously performed by ground control. Hence, astronauts in the crew base need the tools and the skills to perform the EVA monitoring. Ensuring the health of the crew, efficient workflows and critical events will all be additional challenges to face.

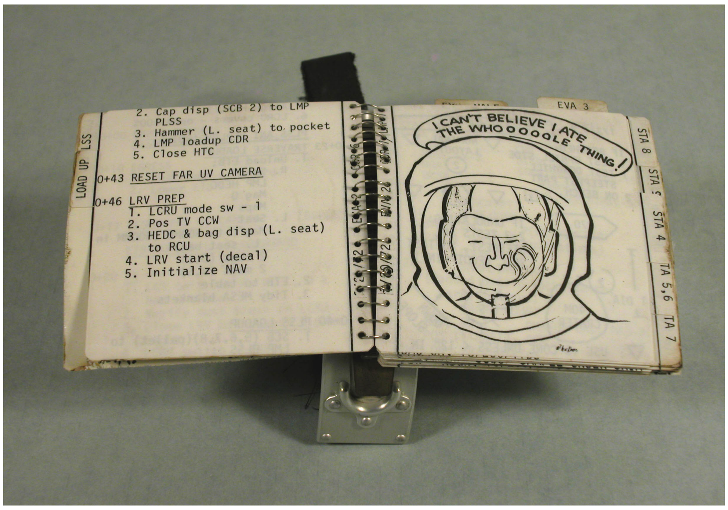


Figure 1.1: Cuff-Checklist EVA 2 & 3 Apollo 16 [219].

Therefore, modern human-machine interfaces have to be designed to support astronauts during their deep space mission operations. The goal is that these systems allow crew to effectively carry out operations, while keeping cognitive overload during off-nominal or critical events to the minimum and situational awareness to the maximum [32].

At the foundation of human spaceflight operations lies procedural work, interestingly, it was found that **half of the accidents reported between 1996 and 2005 were caused by human errors and a significant portion resulted from the incorrect execution of procedures [15]**.

The benefits of augmented reality (AR) especially for procedural tasks have been recognized across many industries. AR for assembly research, namely manufacturing, assembly, maintenance & repair (for surveys see [86], [215], [243]) has been heavily investigated as these activities in industry require collaboration, interaction, delivery of real-time information to release quick and accurate products. So far, procedural instructions used were mostly paper instructions which are time consuming [209] to look up and often show unnecessary information [110], [137], [215]. AR studies investigating the usability of the technology for procedural work have shown that the use of AR leads to lower task completion times for assembly, higher accuracy and lower mental effort [9], [18], [211], [222]. Furthermore, users defined AR as intuitive [209–211]. Despite the fact that research and development of AR for space applications is at an early stage, several studies and projects investigating the benefits of AR for space applications have been performed as well. The majority of these projects focused on AR for procedural work on the International Space Station (ISS), namely mobiPV [3], which was tested on different mobile devices including the Microsoft HoloLens, namely MobiPV4HoloLens [126], WEKIT [34], [128], EdcAR [52], [124], [128], WEAR [146], [221], ARAMIS [72], Sidekick [175] and T2 Augmented Reality [230]. Several pilot studies had a similar focus, namely the one by Braly, Nuernberger, and Young Kim [15] and the one by Markov-Vetter and Staadt [160]. Exploiting the proven benefits of AR for procedural work, more specifically lower task completion times and higher accuracy could enhance human spaceflight operations majorly. Nevertheless, other applications of AR were investigated as well. AR for rover operations was investigated by projects such as WEKIT [139], in fact this project aimed at different use cases, and OnSight [175]. Whereas Furuya et al. [79] investigated the benefits of AR for stowage operations and logistics, the study by Karasinski et al. [99] investigated the use of a tool combining AR and the Internet-of-Things (IoT) for just-in-time training. Finally, an interesting study involved the development of an AR tool, called Holo-SEXTANT, aimed at helping extravehicular (EV) crewmembers while navigating a planned traverse during their extravehicular activities. SEXTANT stands for Surface Exploration Traverse Analysis and Navigational Tool [108] and is a tool which was integrated in the Microsoft HoloLens and tested during NASA's Biologic Analog Science Associated with Lava Terrains (BASALT) exploration field campaign which took place in November 2017, at Hawai'i Volcanoes National Park [55]. This study concluded that the tool is significantly enhancing EVAs [123].

Designing user interfaces that are effective is a major challenge when it comes to emerging technologies that present new ways of user interaction and perception and do not have established design guidelines [100]. Augmented reality is one of these emerging technologies as it has the potential of changing the way humans perceive the world. A literature survey performed in 2004 showed that user-based studies have been rarely utilized for augmented reality user interface design as engineering challenges were the main focus. Swan and Gabbard [97] identified the need to develop AR systems from a technology-centric medium to a user-centric medium. From the survey it could be concluded that of 1104 articles on augmented reality, only roughly 3% addressed aspects of Human Computer Interaction (HCI) and only roughly 2% described a formal user-based study [97]. Swan and Gabbard [97] state that to create effective user interfaces for emerging technologies, such as augmented reality it is crucial to understand perceptual and cognitive characteristics of the users, hence develop AR based on performed user studies [97]. A more recent study which analysed AR studies between 2005-2014 concluded that overall on average less than 10% of all the AR papers were user study papers [7]. Additionally, analysis showed that the majority of the studies included little field testing and few heuristic evaluations. That survey however, included only one AR space-related application study, namely the one by Markov-Vetter and Staadt [160]. It should be noted that most AR studies for space applications focused on the user and gathered feedback from the users through direct observations, surveys and questionnaires and in some cases performed usability tests (see Table D.1 in Appendix D). However, the methodology used for data collection and analysis for these evaluations is rarely clearly outlined. Additional knowledge on user-centered design AR studies in general but also and particularly on user-centered AR space-related application studies are therefore required.

Relevant recommendations for AR applications extracted from several of the above mentioned studies include: the importance of text besides graphics [78], the possibility to visualize telemetry from surrounding devices and the ability to directly interact with the values and parameters [158]. Furthermore, more visual guidance is recommended to cope with the narrow field-of-view (FOV) of currently available off-the-shelf AR devices [128]. Additionally, it is recommended that real-time information should be provided including alarms and warnings. Finally, Wolff et al. [158] stated that IoT can lead to major improvements to sense and control surrounding devices while artificial intelligence (AI) could enable these devices to be smarter. The Holo-SEXTANT study included several interesting recommendations for future surface exploration applications as well, namely an AR assistant able to aid during sample identification, providing suit diagnostics, aiding with tasks, taking notes, communicating with intravehicular (IV) crew or mission control. Moreover, such an EVA tool could have integrated intelligence going beyond simple voice commands. Finally, an artificially intelligent assistant could be included to aid in decision-making processes [123].

While some AR space-related studies recommended combining AR and IoT to enhance sensing and control of surrounding devices other non-space related studies already demonstrated an increase in work-productivity and situational awareness through the use of an AR-IoT tool.

Combining AR and the IoT can significantly improve the user's experience, in fact AR can enable the interaction with IoT objects in the surrounding environment [189]. AR-IoT tools for different applications such as farming [189], retail [46] and remanufacturing [151] have already been developed. These demonstrated that using AR-IoT technology was less error prone and more promising than traditional visualization methods. Despite the great potential of AR in combination with IoT, there are only a few studies combining AR and IoT for space applications which have been carried out. Karasinski et al. [99] investigated the use of an AR-IoT tool for just-in-time training and concluded that cognitive load and task completion time was reduced when using the tool. During the combined EdcAR [124] and WEKIT [128] study, the combination of the two technologies had different purposes. EdcAR simply demonstrated and tested IoT features in AR [124]. While WEKIT included a wearable IoT sensor network in form of a biosensor harness which had to be worn by the test subjects such that potential "stress" indicators could be tracked to get more insights later on during the result analysis phase [139].

It can be concluded that most studies and projects on AR space applications focused on AR for procedural work on the ISS, while only one focused on AR for planetary surface exploration. Hence little knowledge on usability aspects of an AR tool for EVAs is available. Moreover, a human spaceflight tool that combines AR and IoT technologies to increase crew situational awareness has not been developed yet. Based on the recommendations from the Holo-SEXTANT study, namely assistance in: providing suit diagnostics, aiding with tasks while on the field as well as taking notes, and based on the forecast future missions to the Moon, Mars and beyond, where surface operations will be crucial for mission success and communication delays will be a major problem, an AR-IoT tool for lunar and planetary surface activities namely geological site inspection activities is seen as crucial and has not been developed so far. A study that investigates and assesses the usability of such a tool is therefore necessary.

Research Questions and Objective
This leads to the formulation of the following research question:

> *"is an augmented reality user interface a **usable concept** to tackle long communication latencies, hence increase autonomy of astronaut crew during future human extraterrestrial surface activities on the Moon and eventually on Mars?"*

Usable refers to the usability of a tool, it analyses the way the tool will be used and whether it allows the user to do so in an effective, simple and pleasurable way. To answer the main research question whether the tool is usable one needs to assess whether the target group, which in this case are astronauts and astronaut geological field activities experts, perceive the tool as useful, helpful and operationally feasible. In fact, this will be crucial to understand whether the target group would consider adopting the tool in the first place. This leads to the formulation of the following subquestions:

*“is the application perceived as **useful and helpful** for future geological site inspection activities by astronauts and astronaut geological field activities experts when compared to the current tools and media used?”*

*“is the application perceived as **operationally feasible** for future geological site inspection activities by astronauts and astronaut geological field activities experts when compared to the current tools and media used?”*

To be able to design a tool that is usable and relevant to the activities it is designed for, one needs to understand the type of activities that one is designing for and the type of information that is useful for the user and to be provided by the user to accomplish the given tasks. This leads to the formulation of the following subquestions:

*“what are **relevant activities** to be performed during astronaut geological field activities and which are feasible to be designed for in the user interface?"*

*"what **type of information** is required by the user to accomplish the given geological site evaluation task?"*

To ensure and increase the usability of a designed tool, the designer needs to make sure that design principles specifically relevant to the developed user interface are not violated. For this one needs to define the design principles and assess whether these are fulfilled. This leads to the formulation of the following subquestions:

*“what are the identified **design principles** that apply to the developed user interface?”*

“are the identified user interface design principles fulfilled?”

Based on the research performed on the Apollo missions and EVAs in general, it became evident that workload is high during these operations. Crew workload should be minimized based on NASA’s guidelines [173], moreover reduced workload can lead to increased user satisfaction and therefore needs to be assessed. This leads to the formulation of the following subquestion:

*"is the mental **workload** reduced when the AR-IoT tool is used compared to the current tools and media?"*

A major difference between an EVA and operations inside the ISS is the kind of interaction an astronaut can have with an AR system. Based on research it was concluded that voice commanding could present interference problems [154] while hand dexterity, mobility and tactility are limited [147]. This leads to the formulation of the following subquestion:

*"what **kind of interaction** with the AR-IoT tool is feasible for astronaut crew during a lunar or planetary EVA?"*

Finally, the definition of a concept of operations (ConOps) is considered crucial to get an overall picture of the operation from the point of view of the users who will operate the newly developed system [173]. Moreover, a concept of operations is considered important to provide the initially planned controlled experiments with sufficient fidelity to current and possibly future operational environments [121]. While performing controlled experiments was finally not possible due to Covid-19 restrictions on human subject testing, these experiments are still highly recommended for future work, and thus the development of a feasible concept of operation remained essential and relevant. This leads to the formulation of the following subquestion:

*“what is a relevant and feasible **concept of operations (ConOps)** to operate and test the technology developed?”*

To answer all the above mentioned research questions, the following research objective was formulated:

*"**develop and investigate the usability of an Augmented Reality (AR) User Interface (UI)** for future human lunar and planetary surface exploration extravehicular activities (EVAs) and its potential in providing support to the astronaut crew when real-time ground support is unavailable, by carrying out **expert review sessions, heuristic evaluations and expert evaluations through a design requirements compliance questionnaire in combination with semi-structured interviews**"*

Thesis Structure

This thesis is structured as follows: first background information on astronaut geological field activities as well as analogue training missions is presented in chapter 2. A thorough investigation of these activities was crucial to understand the context of use, requirements and the needs as well as to get insights on the type of tasks usually carried out and tools employed. The methodology used for this research project is explained in chapter 3, including the hardware and the methodology used to develop and implement the AR-IoT application as well as the methodology used to evaluate the AR-IoT tool. In chapter 4, the innovative AR-IoT tool developed to assess whether it represents a usable concept to aid astronaut crew during future lunar and planetary surface exploration is elucidated and shown. After that, the conceived operational scenario developed to get an overall picture of the operation from the perspective of the users who will operate the tool as well as to test the tool is defined in chapter 5. The results from the tool evaluation, namely the expert reviews, heuristic evaluations, questionnaires and semi-structured interviews are elucidated in chapter 6. In chapter 7, the results are analysed and discussed, research and design limitations are elaborated upon and recommendations for future work are provided. Finally, in chapter 8 the conclusions from this research are presented.

2
Astronaut Geological Field Activities

In this chapter, general information on astronaut field geology activities, past astronaut geological field activities as well as past and current analogue training missions are presented. An overview of the tools that are currently used and/or developed for astronaut geological field training is given. Moreover, recommendations regarding desired functionalities and capabilities of astronaut crew support tools are described.

2.1. Background Information on Field Geology

On account of the systematic and analytical studies of Earth's landforms and rock outcroppings performed during the late eighteenth century, indications of the great antiquity of our planet were gained, insights which were not quantified by laboratory science up until a century later [71].

Considering that natural sciences will most certainly be the principal scientific motivation to explore new worlds beyond Earth, it is of utmost importance to investigate field geology activities to render these as fruitful as possible as they will underpin successful scientific exploration of the Moon, Mars, and other targets such as asteroids.

As stated by Hodges and Schmitt [130], the primary goal of a field geologist is to reconstruct the geologic history of a delineated study area or outcrop based on observations done on site and samples collected such that they are informed by those observations. Geology is a field of science which entails on-the-fly interpretive synthesis of an aggregation of observations that develop rapidly while the research progresses. Field geologists hardly perceive a traverse as a sequence of destinations. In fact, they study the landscape, noting features that can provide with opportunities for unanticipated study as the features were not right away obvious or visible in the previously provided remote-sensing images. They ameliorate the geologic map and construct a better aggregation of geologic histories during their walk toward the next station. The analysis process is strongly iterative for a field geologist. In fact, the final traverse route is very often quite different from the initial traverse plan specified by the remote sensing data, and the number and locations of stations are very different from the ones forecast prior to field work. Flexecution is therefore the term that best describes the work of a field geologist. The term outlines a problem solving method which includes the continuous redefinition of goals while working toward achieving them, this is why on-the-fly re-planning is such a crucial part of executing a task like a field traverse [130].

The set of actions usually performed by a field geologist, once a target is selected to be studied in detail and hence becomes a station, includes first quickly but carefully laying out a plan of action for the study. Before conducting geological field work activities, it is crucial to understand the context of an area, by studying the broad-scale geology. On Earth, this is possible through remote sensing data, seismological data as well by analysing the existing geological maps from previous studies conducted in that region, whereas for extraterrestrial bodies maps and previous information is limited. Nonetheless, it is still possible to study the broader context of geology by identifying the outcrops of interest from remote sensing images. After that, the process involves a rapid survey from a distance and several different perspectives. Geologists normally take low-resolution photographs of the station from different perspectives, while making preliminary interpretations of the geologic relationships observed. Numerous field geologists also draw annotated sketches of station outcrops in their field notebooks. The field geologist then utilizes the information from the primary survey to identify specific parts of the station outcrop that deserve further detailed study, quantitative measurement and, in some cases, sampling [130].

2.2. Field Geology on the Moon

For a moment in history, in the years from 1969 to 1972, humans got a preview of field work on the Moon. During six successful human lunar surface exploration missions, the Apollo program procured groundbreaking knowledge on Moon geology and 381.7 kg of rocks [130].

Field exploration activities of the Moon during the Apollo missions, together with the following analysis and interpretation of the discoveries and the collected samples reinforce today's understanding of the origin and history of the Moon. And this understanding keeps on providing us with new and important insights into the early histories of Earth and other bodies in our Solar System, as well as the time during which life formed and started to evolve on Earth [80], [130]. Apollo lunar exploration activities led to the discovery of crucial and conceivably commercially viable lunar resources which could aid in satisfying future demand for space consumables, such as life support, power and propulsion. In fact, both ESA and NASA are currently planning activities and developing technologies for lunar sample return missions and in-situ resource utilisation experiments [63], [169]. These conclusions on the potential benefits gained from lunar surface resources come from lunar sample analyses and depend heavily on a foundation of geological field observations, documentation through photos, detailed locations, and the context corroborated by active and passive geophysical data collection as well as photo-geological mapping [130].

While developing the lunar geology concept for the Apollo missions it quickly became clear that several basic operational parameters constituting field geology on Earth had to be changed for lunar field work. In fact, limited consumables such as breathing oxygen, electrical power for the crew's spacesuit systems and all the safety factors would present major constraints to the flexecution strategy.

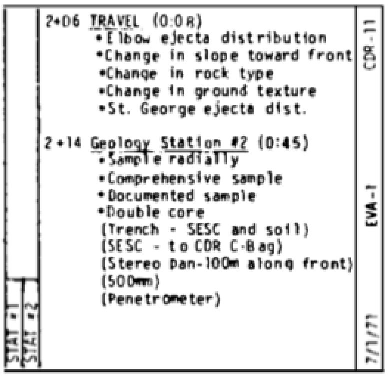
Finally, geological traverses were extremely judiciously planned during the Apollo missions, in fact only minor modifications occurred. Nevertheless, Apollo astronauts followed similar observational and sampling strategies as terrestrial geologists, despite the tighter time constraints [130].

During Apollo missions, astronaut crew assigned to geological field activities had the chance to use more "heads" than just their own, thanks to the presence of the "Science Backroom" which, having taken part in astronaut training and traverse planning, was assigned to monitoring the Apollo extravehicular activities (EVAs) and communicating questions or recommendations for surface observations to crew through the Flight Director [130]. It should be noted that 50 years ago, supporting technology that could replace the "Science Backroom" or "extra brains" was not available yet desirable as communication issues occurred. Nevertheless, the scientific investigations at every station as well as the observations in between were carried out with significant autonomy by the astronauts which were following simple instructions prepared prior to the expedition by science team members on Earth. These instructions were included in abbreviated form into a cuff-checklist attached to the astronauts' spacesuit. In Figure 2.1 instructions provided to the crew for EVA I along the rim of Hadley Rille as found in the Apollo 15 Lunar Surface Procedures document [130], [171] are shown. Examples of abbreviated forms on instructions and maps for cuff-checklists are shown in Figure 2.2 [171] and Figure 2.3 [171], respectively.

Station/ activity	Planned time at start (h)	Planned segment time (h)	Geological features	Observations and activities
LM	–	1:25	Smooth mare	• See lunar surface procedures document
Travel	1:25	0:17	Across typical smooth mare fill toward rim of Hadley Rille	• Observe and describe traverse over smooth mare fill material • Describe surface features and block distribution • Note any differences between mare and rille rim material
Checkpoint	1:42	0:02	Near Canyon Crater	–
Travel	1:44	0:07	Around Elbow Crater	• Observe low scarp around Elbow Crater • Observe any differences between rille rim material and mare material • Observe distribution of ejecta around Elbow Crater
1	1:51	0:15	Near southern part of Elbow Crater ejecta blanket	• Radial sampling of Elbow Crater • Pan
Travel	2:06	0:08	To Apennine Front slope north of St. George Crater	• Look for changes in lithology or ground texture as indications of base of front • Compare mare and rille rim material to Apennine Front • Observe character and distribution of St. George ejecta blanket
2	2:14	0:45	Near base of Apennine Front north of St. George Crater	• Radial sample of St. George Crater as slope permits • Comprehensive sample area at Apennine Front • Double core tube • 500-mm-lens camera photography—blocks on St. George rim and Hadley Rille • Stereo pan from high point—100 m base along front • Fill SESC at Apennine Front • Penetrometer
Travel	2:59	0:09	Across base of Apennine Front to edge of possible debris flow	• Observe Apennine material and relation to mare surface
3	3:08	0:14	At base of Apennine Front adjacent to possible debris flow	• Examine flow and compare to mare and front • Documented samples of Apennine Front and "flow" material • Observe and describe lateral changes in Apennine Front; compare to previous stop • Pan • Observe characteristics of EVA II route
Travel	3:22	0:28	From base of Apennine Front across mare to LM	• Observe characteristics and extent of possible debris flow • Observe area to be traversed on EVA II • Compare mare material to Apennine Front and rille rim • Observe possible ray material
LM	3:50	3:10	Smooth mare	• ALSEP deployment • EVA closeout

Notes: Reproduced from Jones and Glover (2009). Times are shown in hours:minutes form. Abbreviations: ALSEP—Apollo lunar surface experiments package; EVA—extravehicular activity; LM—lunar module; Pan—photographic panorama; SESC—special environmental sample container. During the actual EVA, station 3 was omitted for reasons described in the text.

Figure 2.1: Instructions provided to the crew for EVA I along the rim of Hadley Rille [171], [130].


2+06 TRAVEL (0:08)
•Elbow ejecta distribution
•Change in slope toward front
•Change in rock type
•Change in ground texture
•St. George ejecta dist.

2+14 Geology Station #2 (0:45)
•Sample radially
•Comprehensive sample
•Documented sample
•Double core
(Trench - SESC and soil)
(SESC - to CDR C-Bag)
(Stereo pan-100m along front)
(500mm)
(Penetrometer)

CDR-11
EVA-1
7/1/71
STAT #1
STAT #2

Figure 2.2: Cuff Checklist Apollo 15 - Instructions [171]

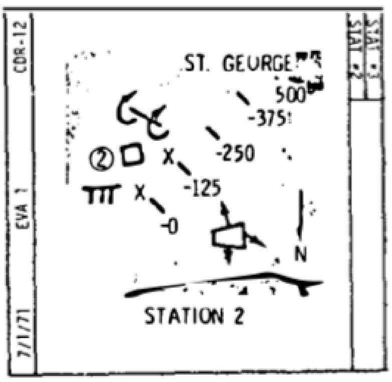


Figure 2.3: Cuff Checklist Apollo 15 - Map [171]

2.3. Future Extraterrestrial Field Geology

Field geology on the Moon is obviously a more challenging enterprise, physically and intellectually, compared to field geology on Earth. Therefore, it is crucial that technologies and strategies that increase overall efficiency as well as effectiveness are developed to ensure cost- and time-effective geologic field work on the Moon [81] and eventually on Mars and other targets in our Solar System.

In contrast to a field geologist on Earth, as mentioned earlier a lunar field geologist must always be aware that time is limited, consumables are limited, fatigue can be fatal, and that, generally, going back to a location is unlikely. Additionally, the lighting conditions on the lunar surface can represent major challenges for geological field work, particularly sampling, as at specific moments in time it is difficult to distinguish surface topography. Experience from the Apollo missions suggests that EVAs performed during the time around the lunar noon were particularly difficult caused by a lack of surface definition due to elimination of shadows [38]. All of these components add to the normal intellectual workload of doing field work. Furthermore, tasks take longer in a spacesuit, if they can be completed at all. Writing notes as well as drawing diagrams by hand is impossible while wearing pressurized gloves. Planning, automated location, voice communications and the documentation of sampling activities, as well as the whole field experience are of even greater importance than usual [81].

It is envisioned that future lunar exploration missions will provide with the opportunity to carry on and extend lunar human exploration including geological, geochemical, and geophysical exploration. The way future lunar field exploration will be conducted, the training activities as well as the planning of activities will derive from the lessons learned during the Apollo missions, new relevant technologies and ideas that have been developed over the past 50 years [130] including those that are currently being and will be developed in the next years.

To take full advantage of extraterrestrial geology when it again becomes practical during future lunar exploration activities, it is imperative for geologists on Earth to start learning how to incorporate advanced technologies into their research. In fact, as new models of "advanced" planetary field geology are being developed and iterated upon [130], and the probability that astronaut explorers will have more time in the field, it is crucial to develop an exploration strategy that expects greater astronaut pre-mission geologic experience and allows for greater astronaut research autonomy while recognizing the risks and resource limitations of lunar and planetary field geology.

Schmitt et al. [81] states that various types of equipment for mobility, sampling, sample documentation, communications, and positioning and navigation will be required to accomplish successful exploration missions. Additional surveying equipment and in-situ mineralogical, geochemical, and geophysical tools should be available and could significantly improve the scientific and operational return of extended exploration compared to the Apollo missions.

Some basic tools and some advanced tools specifically required for site evaluation activities are listed below:

- shoulder- or helmet-mounted, continuous or voice-activated, digital, stereo photo-documentation cameras [81];
- helmet visor or suit-mounted display with automatic position and elevation determination system integrated with a global satellite communications and navigation system or a local site-specific relay and triangulation system [81], [197];
- voice-activated, in-helmet display or visor-projected heads-up display (HUD) for displaying "cuff-checklists," exploration-related data, suit systems and consumables status as well as other relevant data [81], [197].

During future lunar exploration it is forecast that the Science Support Room (SSR), which has a different role compared to the "Science Backroom", will play an ever-increasing support function. The maintenance of coherent and directly accessible and addressable compilations of preceding observations, documentation, samples, and other data collection will have to be ensured. In fact, these will be deciding for planning the next day's activities and longer-term exploration on whichever given mission or sortie from a lunar or planetary base [81].

2.4. Innovative Technologies for Analogue Campaigns and Recommendations

The autonomy that the Apollo explorers could avail themselves of while carrying out field geology tasks on station was only enabled by attentive tutoring in field geology after being assigned to a mission [130].

Provided the current vision for advanced planetary field geology, as outlined by Hodges and Schmitt [130], is realized, future astronaut explorers will require even more intensive geological training as they will have even greater autonomy.

The Apollo J-mission training experience has been successful as it featured field tutorials in a simulated mission environment and not plainly classroom education [130]. Hodges and Schmitt [130] state that an even more proactive training regime would necessitate an accentuation on real geological problem solving at analog sites with astronauts. To test and improve emerging concepts and designs for advanced planetary field geologic studies, geologic studies of analog sites on Earth that employ new technology enabled strategies rather than traditional research methods represent ideal opportunities [130].

Several analogue studies have been carried out over the years, these among others include: the NASA Extreme Environment Mission Operations (NEEMO) analogue, NASA's Desert Research and Technology Studies (Desert RATS), NASA's Biologic Analog Science Associated with Lava Terrains (BASALT) and ESA's Caves & Pangaea analogue missions.

Several Pangaea as well as Caves training and test campaigns have taken place over the past years, in fact Caves training started in 2011 while Pangaea started in 2016. During the Pangaea-eXtension campaign, also known as Pangaea-X [135] in 2018, the specific goal of the campaign was to evaluate potential applications and developments for future mission scenarios and to increase operational relevance of training. Several different approaches for the operational scenarios were simulated, tested and analysed [134]. Different sites, with specific geologic and environmental characters were made available within the Lanzarote Geopark. A complex logistic framework was set up with the aim of maximising efficiency and allowing an interconnection between the different experiments [134].

One of the technologies tested during the Pangaea-X campaigns is the Electronic Field Book (EFB) for tablets Figure 2.4 [62]. ESA-EAC initiated the development of the EFB, with the aim of integrating geological reference, sampling protocols, traverse preparation, navigation and documentation, analytical tools as well as providing real-time awareness to support mission control personnel and scientists during operations [135].

Figure 2.4: Pangaea-X 2018 campaign in Lanzarote. Participants with Electronic Field Book (EFB) on tablet [62].

As reported by Bessone et al. [135] the EFB provides with the following functionalities:

- Pre-defined traverses, geological stops and sampling sites are accessible through the system with the possibility to retrieve associated reference and real-time information;
- Geological stops and sampling sites can be marked on the map;
- The location of all field elements can be identified real time;
- Geo-located relevant geological multimedial information and data analysis can be collected and stored;
- Information can be exchanged between field and ground segment;
- Type of information to be retrieved, collected and exchanged includes, but is not limited to: geolocation, rich text, geo-located photos, audio, videos and/or panviews, maps, procedures and geological databases.

Finally, the set of actions selected for a site evaluation activity are grouped into activity categories following the order mentioned hereafter [135]:

1. **Stop Description** which consists of describing the stop.
2. **Geolocation Site Screening** which consists of creating a stop on the EFB, taking photos and panoramic 360 photos.
3. **Analytical Screening** which consists of mineral screening and communicating the mineral spectra.
4. **Documenting Stop and Area** which consists of GPS photos and filling in other information in the EFB regarding coordinates, areas, description and performing sketches.

It can be noted that the actions mentioned above are in line with the ones described by Hodges and Schmitt [130] who reports that when a target is selected to be studied in detail and hence turns into a station, a good field geologist at first quickly but carefully lays out a plan of action for the study. After that, the process involves a quick survey from a distance and several different perspectives. Geologists normally take low-resolution photographs of the station from different perspectives, making preliminary interpretations of the geologic relationships observed. Numerous field geologists also draw annotated sketches of station outcrops in their field notebooks. The field geologist then utilizes the information from the primary survey to identify specific parts of the station outcrop that deserve further detailed study, quantitative measurement and, in some cases, sampling.

For BASALT analogue campaigns, unlike Pangaea analogues where the main goal is to evaluate potential applications and developments for future mission scenarios and to increase operational relevance of training, the scope is to develop, test, analyse and validate new exploration architectures, specifically for Martian exploration. Nevertheless, insights gained from these Mars-like mission scenarios can be adopted for Lunar-like exploration scenarios.

Various improvements have been identified during the BASALT-2 analogue mission in 2016 and are listed below [122] :

1. The ability to incorporate a hands-free (e.g. speech recognition) mean with which one could add geospatially linked electronic field notes to increase efficiency and reduce workload. More specifically, adding a feature such as speech recognition "Create EV1 waypoint A" which would directly create a waypoint called A at the current location of the extravehicular (EV) 1 crewmember. This would enable EV or other crewmembers to easily and promptly mark field locations of interest without EV crew having to take their hands off the tools they are using in the field at that specific instant in time.
2. Certain means to maintain a running tabulation of all notes linked to a certain candidate sample that is automatically organized and visible to both the intravehicular (IV) crew and the Mission Support Centre (MSC) right away.

Other relevant capabilities assessed during BASALT-2 but not implemented with respect to EV support tools were [122]:

1. Simple text-based displays providing EV with data on critical suit consumables, elapsed time, remaining time in phase of EVA, bingo times to key milestones namely the latest time in which particular information must be sent by the MSC to keep the EV crew on schedule [12], suit health and breathing atmosphere pressure/composition.

2. Head-mounted display (HMD) that overlays virtual traverse waypoints and annotations on the surroundings and allows EV, IV and MSC to overlay text and annotations on the surroundings.

3. VR/AR/telepresence system that displays EVA terrain allowing IV and MSC to virtually join EV in the field.

Several of the new capabilities mentioned as being useful in BASALT-2 were included and evaluated during BASALT-3 such as virtual training environments and HMD technologies for the EV crew and telepresence systems for MSC and IV crew to 'join' EV crew in the field. One example of these is Holo-SEXTANT [123], for more details see subsection G.2.2. The field test results from the BASALT-2 and BASALT-3 are being integrated into EVA concept of operations documents which are owned and maintained by the NASA JSC EVA office [122].

Critical technologies identified during analogue missions were also investigated and are reported hereafter. It was found that constant video feeds from crew-mounted cameras are useful, however they rarely provided contextual information above and beyond what could be deduced from photographs taken by the EV crew provided that, those images followed the predetermined and systematic imaging protocol. The ability to send close-up photographs at reasonable resolution was the primary mean to transmit scientific information from EV through IV and to MSC. Furthermore, directing photographs by the highly trained EV crew proved to be significantly better compared to remote operation, as EV crew is in the best position to assess the best subject, scale and contrast. It should be noted that also other analog programs have examined questions on decision-making and how science teams can operate in simulated extraterrestrial missions, nevertheless there is not much literature available on the details regarding the background of particular protocols chosen, or in detail examination of the results regarding the choices made [12].

During the BASALT program, an auxiliary role was the one of a stenographer who was in charge of transcribing verbal communications by EV crew [12]. It is envisioned that future science operations management systems could feature automatic transcription of all kinds of voice communications.

Finally, important support capabilities for EVA execution have been investigated and presented by Marquez et al. [98], these are reported hereafter:

- Timeline Management
- Life Support System Management
- Physiological Management
- Communication Management
- Science Operations Management

It can be concluded that innovative technologies are needed to ensure the success of future human lunar and planetary exploration missions as cuff-checklists will not be sufficient to support and guide crew during surface operations. Moreover, despite several modern human-machine interfaces are being tested during analogue campaigns already, an augmented reality hands-free application able to aid astronauts during geological site inspections, and to provide with critical suit consumable data is recommended but has neither been developed nor been tested yet.

3
Methodology

In this chapter an overview of the methodology used for this research project will be given, more specifically the development of the AR-IoT tool is described together with the evaluation strategy. In section 3.1 the hardware used and the software implementation of the system are elucidated. Then, the methodology used to design the user interface and to evaluate it is described in section 3.2.

3.1. Application Design

In this section, first the hardware used to develop the AR-IoT system is described and then the software implementation is explained.

3.1.1. Hardware

Several hardware components have been utilised for this research project and will be presented hereafter.

Microsoft HoloLens 1

For this research project the Microsoft (MS) HoloLens [163], more specifically the 1st generation MS HoloLens, represents the most crucial hardware component. The MS HoloLens is an augmented reality (AR), six degrees of freedom (DOF) Head-Mounted Display (HMD). It features an estimated field-of-view (FOV) of 30° x 16.5°, a stereoscopic display with 1268 x 720-pixel resolution for each eye, spatialized audio technology, Wi-Fi as well as Bluetooth wireless technology. These augmented reality glasses have an onboard computer which includes a general-purpose processor as well as a custom Holographic Processing Unit (HPU). The device is untethered and can thus be used easily and independently. An overview of the device including the hardware components is shown in Figure 3.1 [55].

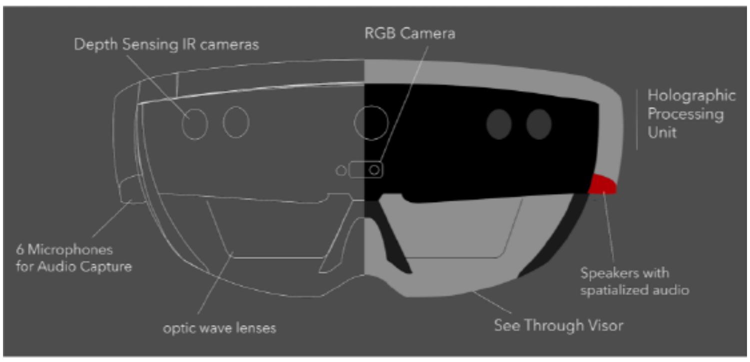


Figure 3.1: Overview of the Microsoft HoloLens (1st gen) hardware components (highlighted) [55].

To interact with the HoloLens the user can make use of hand gestures, gaze and voice inputs. HoloLens' AR and Mixed Reality (MR) capabilities permit images to be projected on the optic wave lenses in front of the user's eyes and become visible to the user in the form of holograms in the environment. The HoloLens is capable of mapping the user's surroundings by creating a 3D map through its sensors and can "place" holograms on physical components. The HoloLens is also capable of keeping track of these locations by caching them, this way allowing the user to return to previous locations and finding the earlier placed holograms. The MS HoloLens provides with state-of-the-art capabilities for augmented reality, furthermore it has an easily-accessible developer platform. It uses an incorporated Inertial Measurement Unit (IMU) and sensors to handle relative position and rotation in order to represent visual information accordingly, reducing the user's physical and mental taxing responsibility of orienting and visualizing. Additionally, the device allows the user to transform their view within an environment by rotating their head in the HMD [163]. This device was chosen for this project based on my previous experience with the HoloLens during my internship at the European Space Agency's Astronaut Centre (ESA-EAC) and as it represents state-of-the-art hardware for augmented reality technology. It should be noted that it is also the only augmented reality HMD type available at ESA-EAC.

Despite being a state-of-the-art device, the hardware shows limitations. The HoloLens is an indoor headset and the core technology of the device is designed for that. Moreover, it utilizes Wi-Fi signals to localize and time of flight IR sensors to map the user's surroundings. It should be noted that the display brightness is maintained suitable indoor under artificial lighting.

HoloLens clicker

The HoloLens clicker [162] specifically designed for the HoloLens 1st gen was used as well. It provides the user with another way to interact with holograms, in fact it can be used instead of hand gestures to select, scroll, move, and resize applications. For the application developed for this research project it was solely used to select. The "select" action can be performed by pressing on the indentation which can be seen in Figure 3.2 [132] .

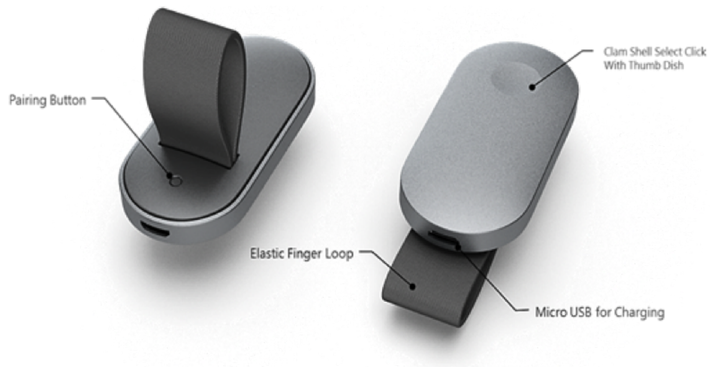


Figure 3.2: HoloLens (1st gen) clicker, the elastic finger loop is highlighted on the right-hand side where one can see an indentation as well on which one can press using their thumb [132].

The clicker is connected to the HoloLens via Bluetooth.

Raspberry Pi

A Raspberry Pi 3 (see Figure 3.3) [202] was used to create the link between the HoloLens interface and the Internet-of-Things components.

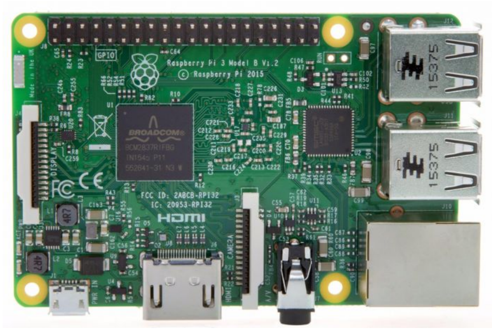


Figure 3.3: Raspberry Pi 3 [202].

Most notably, Raspberry Pi 3 has four USB 2.0 ports and built-in Ethernet, as well as a 1.2 GHz 64-bit quad-core ARMv8 CPU, 802.11n Wireless LAN and Bluetooth 4.1 Low Energy (BLE). The powerful CPU together with Wireless LAN and Bluetooth 4.1 radio renders this minicomputer an ideal candidate for IoT projects, because several sensors can be simultaneously connected to it. In addition, the Raspberry Pi has a 40-pin GPIO (General Purpose Input Output) connector to interface with external sensors [202].

In subsection 3.1.2 more details on the integration of the internet-of-things and augmented reality, hence between the Raspberry Pi and MS HoloLens will be explained.

3.1.2. Software Implementation

The AR-IoT tool runs the Unity 3D game engine [227] to visualize and render on top of the real world environment. This section presents the technical implementation methods and encountered challenges.

AR User Interface Development

The application has been developed using the Unity [227] framework (2018.4.13) and C# as the scripting language along with the open source 2017.4.1.0 version of the Mixed Reality Toolkit (MRTK) [161]. The 2017.4 version has been chosen as it represented the most stable option at the time of development. The MRTK is an open source collection of scripts, components, and tools for input to accelerate application development suited for the MS HoloLens. An overview of the AR-IoT application development in the Unity environment is shown in Figure 3.4. Please note that all application scripts are available upon request.

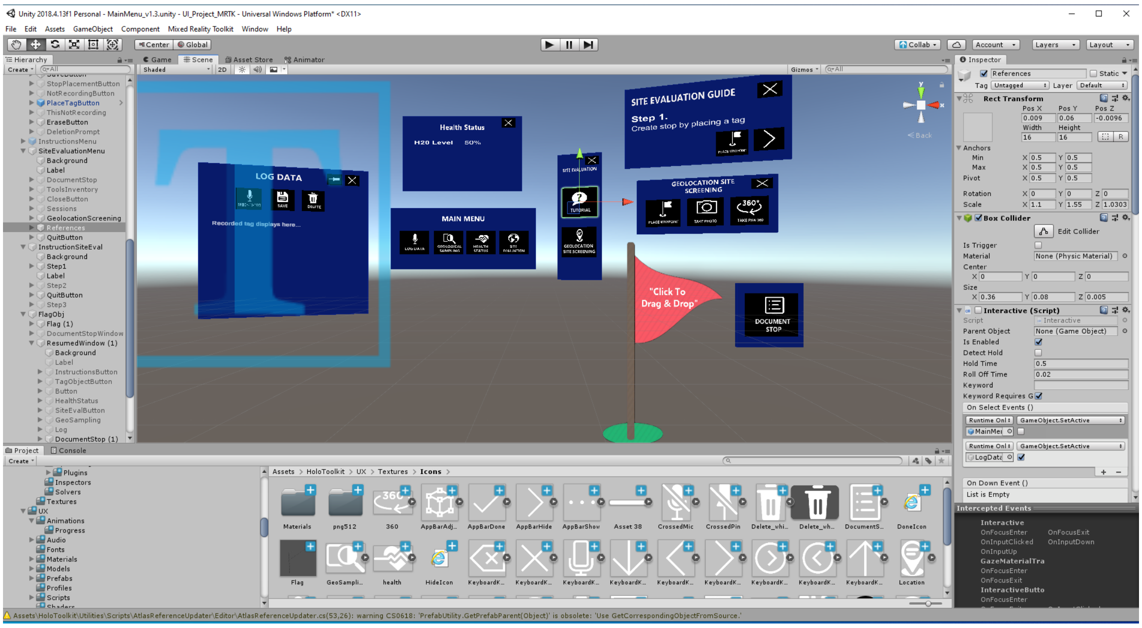


Figure 3.4: Overview of the AR-IoT application development in the Unity environment

AR IoT integration

To ensure communication between the HoloLens and the IoT devices, for instance sensors which provide with biomedical and consumable levels data, an IoT client-server architecture via Message Queuing Telemetry Transport (MQTT) protocol has been adopted and is illustrated in Figure 3.5.

The HoloLens and the server, also called broker, communicate via MQTT, a lightweight publish/subscribe protocol [203]. The communication between the Raspberry Pi [202] and the HoloLens is as follows, the MQTT broker receives information to send to the HoloLens from the Raspberry Pi which receives data from the sensors directly.

The open source message broker called Mosquitto [196] that implements the MQTT protocol and has a command line publisher/subscriber client was selected. For the HoloLens the open source M2Mqtt library [188] was implemented. This library includes a main class MqttClient that represents the MQTT client to connect to a broker. One can connect to the broker providing its IP address or host name and optionally certain parameters that are specifically related to the MQTT protocol.

Once the connection with the broker is established one can use the Publish() method to publish a message to a topic and the Subscribe() method to subscribe to a topic and receive messages published on the specific topic. It should be noted that the MqttClient class is events based so that one receives an event when a message is published to a topic one subscribed to, one also receives events when a message publishing process is completed, or when one has subscribed or unsubscribed to a topic.

The Raspberry Pi and the broker also communicate via MQTT protocol. The open source Python MQTT client Paho Eclipse [59] was embedded in the Raspberry Pi, additionally for the time being, an MQTT publishing script, that is able to send dummy sensor values, was written and implemented in Python. The MQTT Publisher script for the Raspberry pi is also available upon request. The idea behind this is that, when sensors able to publish meaningful data will be available, the actual data will be retrieved and formatted in the same way as the MQTT publisher script momentarily does. Then the sensor values will be directly linked to the specific topic, as an example *topic/heartrate* and random heartrate values will not have to be simulated anymore. The publisher script will therefore be adjusted and integrated such that the Raspberry Pi handling incoming sensor data can publish the incoming values to specific topics to the broker. It should be noted that for the sensors to be able to send the data to the Raspberry Pi, a Wi-Fi chip such as a ESP8266 chip [64], ZigBee [250] or LoRa [140] will have to be connected to the sensor.

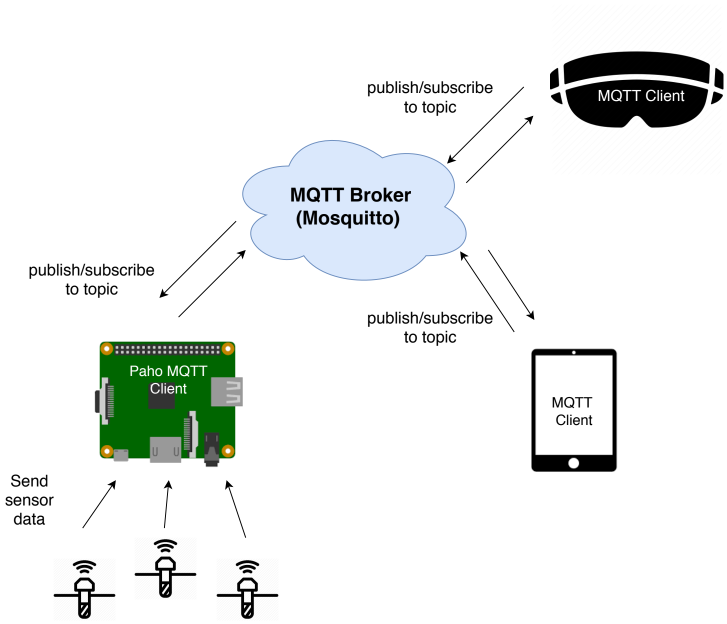


Figure 3.5: AR-IoT infrastructure: Mosquitto MQTT broker managing telemetry transfer between Microsoft HoloLens [88] with embedded MQTT client, arbitrary display [89] with embedded MQTT client and Raspberry Pi [193] with embedded MQTT client receiving and handling data from different sensors [88]

Several tests were performed to assess whether the connection and data transmission between the Raspberry Pi and the HoloLens was working. First, the connection and data transmission between the broker and the HoloLens was successfully tested using the embedded Mosquitto [196] command line publisher client. Then, the connection between the embedded MQTT client on the Raspberry Pi and the broker was tested successfully as well using the embedded Mosquitto [196] command line subscriber client.

Finally, the whole setup was successfully tested. For this, it was ensured that: all devices are connected to the same network and the IP address of the broker as well as the topic to subscribe/publish to (in the test case "topic/water") was correctly specified for the embedded MQTT client in the HoloLens and in the Raspberry Pi. The open-source MQTT Mosquitto broker [196] was used for these tests. Therefore, it can be concluded that the transmission of data between the Raspberry Pi, sending data from its potentially connected "dummy" sensors, and the HoloLens is working.

As can be seen in Figure 3.5, an arbitrary display with embedded MQTT client is shown as well, this is because the initial evaluation methodology, which could not be pursued due to Covid-19 and the closure of ESA-EAC, included controlled experiments in which two scenarios would be compared, one involving the innovative AR-IoT tool and the other using conventional media such as a cuff-checklist and a simple Graphical User Interface (GUI) on an arbitrary device. In fact, to make the comparison fair the same amount of information has to be provided. Therefore, to make sure that the IoT sensor data information is provided in both scenarios, a device with a simple GUI displaying the same dummy or real sensor data as on the HoloLens, was considered essential. Together with Sebastian Lorenz, intern at ESA-EAC, after providing the specifications of the IoT MQTT setup used for the HoloLens, a concept for the software architecture was discussed. Consequently, the Android application was developed by Sebastian Lorenz. Here, the main scope was to demonstrate data transmission of dummy sensor values via a Raspberry Pi and MQTT broker to the HoloLens as well as an Android device and this was achieved.

According to the specifications of Sebastian Lorenz, similarly to the HoloLens IoT integration, the open source MQTT client Paho Eclipse [59] library was used. This time the library written in Java to develop Android applications was used. An MQTT subscriber able to subscribe to the same topics as the HoloLens, capable of retrieving consumable levels and biomedical data, was then developed and integrated in the application. Additionally, a script which enables the user to specify the IP address of the MQTT broker, once the application is started, was developed. It should be noted that the application was written in Kotlin.

The layout of the application was kept simple and is shown in Figure 3.6 and Figure 3.7.

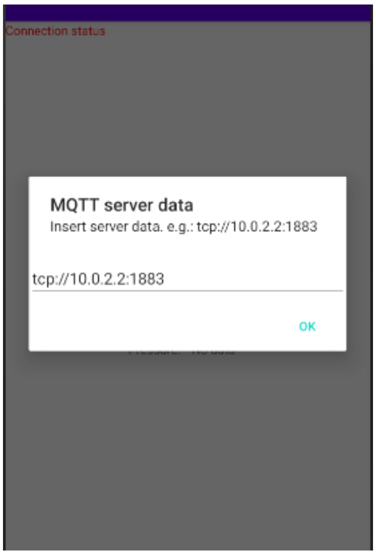


Figure 3.6: Custom-made graphical user interface of the Android application requesting the input of the MQTT broker/server IP address.

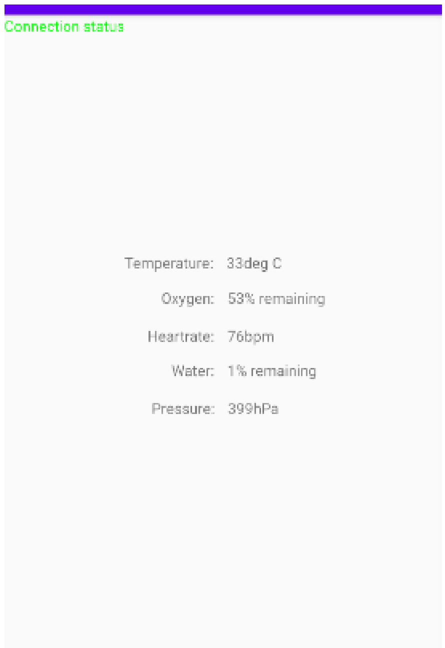


Figure 3.7: Custom-made graphical user interface of the Android application.

It should be noted that the application was tested successfully. More specifically, the dummy sensors values could be successfully displayed on the application on an Android device. The following procedure testing was applied: after deploying the application on an Android phone and inputting the server IP address, data transmission was tested by running the developed Python MQTT publishing script on the Raspberry Pi. Again, it was ensured that the server IP address specified was the same in the publisher script and in the subscriber script and that all devices were in the same network.

3.2. Research Design

As suggested by literature (see section G.5), when designing for people, especially when designing an interface, a user-centered design method is highly recommended. As stated by Swan and Gabbard [97] it is crucial to understand perceptual and cognitive characteristics of the users, hence develop AR based on performed user studies [97]. As mentioned in chapter 1, past AR studies did not often adopt user-centered design methodologies, in fact less than 10% of all the AR papers between 2005-2014 were user study papers. Moreover, the majority of the studies included little field testing and few heuristic evaluations. That survey however, included only one AR space-related application study, namely the one by Markov-Vetter and Staadt [160]. Most space-related AR studies did focus on the user and frequently gathered feedback from the users through direct observations, surveys and questionnaires (see Table D.1 in Appendix D) and in a few cases performed usability tests. Nonetheless, a clear outline of data collection and analysis is often missing.

Based on a review of design models, principles and standards (see section G.5 and section G.6) it was concluded that a user-centered design method is the most feasible for a research project of this kind and was therefore adopted. It should be noted that this model is in line with the human-centered design model used by NASA [173] and the principles proposed by Gould and Lewis [116] suggested by Lee et al. [96] (for further details, the reader is referred to section G.5). The user-centered design cycle that has been applied to the augmented reality interface design in this research is depicted in Figure 3.8.

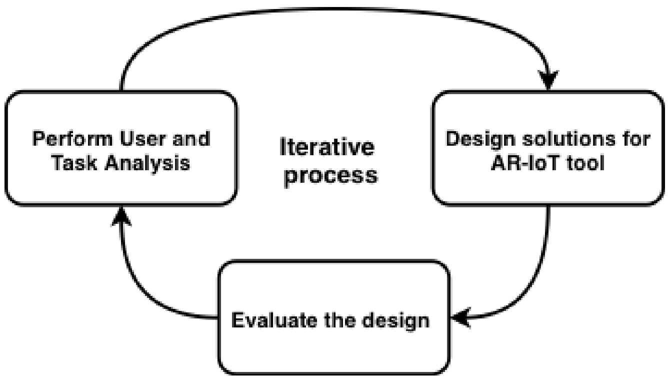


Figure 3.8: User-Centered Design cycle applied to the Augmented Reality Interface Design

From Figure 3.8 one can see that the first step is to determine the requirements of the users [149], in this case the target group of this tool is: current and future generation astronauts, as well as the context in which the application will be used. A user and task analysis has been carried out already during the literature study (see chapter 2 and for more details Appendix G) to define design considerations and a relevant surface EVA scenario. Extravehicular activities (EVAs) with a special focus on the Apollo surface operations and future EVA scenarios have been investigated (see chapter 2 and for more details section G.1). This included gathering important recommendations from Apollo astronauts based on summary reports [118] and interviews [154] on surface operations and medical operations [197] (for more details the reader is referred to subsection G.1.3 and subsection G.1.2). Nevertheless, as the design process is iterative, user and task analyses were not only performed at the start of the user interface design but throughout the various design phases, it should be noted that the outcome is reported in subsection 3.2.1.

After a task and user context analysis has been performed, a concept can be created which combines the knowledge of human characteristics, principles of human behaviour and interface guidelines [96]. Attentional, perceptual, memory and mental model are important design principles one needs to keep in mind (for more details see subsection G.6.1) .

In Figure 3.9 one can see how the selected design process is reflected in the AR-IoT tool design. Several concepts were created in the early design phase, the so called conceptual phase, before a first prototype was created. Details on the design of the different concepts and the first prototype for the AR-IoT tool are elucidated in subsection 3.2.2.

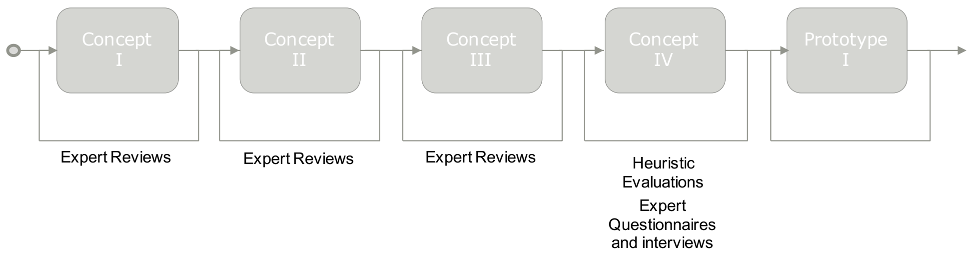


Figure 3.9: User-Centered Design Process applied to the design of the AR-IoT tool. Four concepts were created before the design of the first prototype. All these concepts were constantly evaluated and re-iterated.

Finally, from Figure 3.8 and Figure 3.9, it can be seen that the process is highly iterative, after the creation of concepts and prototypes across the different design phases, constant evaluations with experts and target group users are performed and iterations follow each of these evaluations. The methodology used for the tool evaluation phase is reported in subsection 3.2.3.

3.2.1. User and Task Analysis

As mentioned earlier, the first step to take when designing a user interface while following the user-centered design methodology is a user and task analysis. This analysis was started during the literature study phase and continued throughout the design phases of the AR-IoT tool.

Based on the common EVA tasks identified and astronaut training activities performed at the European Space Agency's Astronaut Centre (ESA-EAC) more specifically Caves & Pangaea analogue campaigns, the selected EVA scenario is an **astronaut geological field activity**.

Gathered recommendations from Apollo astronauts that carried out surface operations include [154], [197]:

- It is useful to have access to navigation and position data, consumables and biomedical data on demand;
- A Head-Up/Head-Mounted voice-activated display is useful considering eye and hand coordination due to pressure suit, however interference issues with other voice loops have to be considered;
- It is suggested to display relevant information on safety and risks only and if possible more details on a call-up basis;
- Support from habitat intravehicular (IV) crew can be useful to support operations rather than only ground support.

As mentioned in chapter 1, in view of the Artemis missions [179] to the Moon and future missions to Mars and beyond, predicted risks and issues will be (see subsection G.1.6 for more details) :

- Long communication latencies causing limited availability of ground support which will require more autonomy by the astronaut crew to manage both system performance and mission execution;
- More complicated tasks will have to be performed compared to the Apollo mission tasks.

Based on the information gathered, it becomes possible to answer the question identified by NASA [173] (see subsection G.6.1) concerning the kind of interaction and possibilities the user can have with the display. Hand coordination for an EVA as mentioned in subsection G.1.2 is altered by the limits of the pressure suit hence hand gestures are not considered to be feasible to interact with an AR interface. Voice commanding could be a feasible means of interaction, however it might cause interferences with other voice communication loops, that is why it is often recommended to avoid voice input unless strictly necessary. Eye gazing seems to be a feasible solution, however technology limitations and usability has to be investigated.

Before starting the development of the first concept a discussion on useful functionalities that an augmented reality geological inspection tool can provide with was performed. EVA experts and geological astronaut training experts from Caves & Pangaea at ESA-EAC were consulted. From the discussions it became clear that the main astronaut geological activities include task and route planning as well as operations.

These activities include:

- Surface Exploration Traverses
- Site Inspections
- Geological Sampling during Pangaea training
- Experiments during Caves training

For the above mentioned activities, the following useful functionalities were identified:

- Create waypoints in the augmented environment to mark points of interest to which gathered information can be attached. Apart from allowing systematic site inspections, this can ease navigation and planning of following EVAs;
- Log gathered environmental information through speech-to-text conversion via speech recognition;
- Display procedural steps to aid the flow of operations;

- Retrieve and display environmental or physiological data from IoT devices;
- Provide tool locating capabilities;
- Provide navigation capabilities;
- Provide temporal data (e.g. elapsed time and time remaining to complete assigned operations);
- Retrieve and display data from a rover supporting the activities.

It is out of the scope of this research to include all useful functionalities identified, therefore a selection of a coherent and feasible set of functionalities was performed. Finally, all above mentioned functionalities except: tool locating capabilities, navigation capabilities, retrieving and displaying of temporal data as well as rover status data were selected and integrated. More information on the features and rationale can be found hereafter and in chapter 4.

3.2.2. Concepts and Prototypes Design for the AR-IoT tool

After the initial user and task analysis, the design of concepts can be initiated. The concepts created during the early design stage, as well as the iterations performed which led to a final concept and after that to a first prototype will be elucidated hereafter.

During the initial design stage, regular expert reviews were held to brainstorm specifically on the application functionalities and feasible user interaction strategies as well as user inputs. Several digital concepts were created during this early design phase, discussed during these expert reviews reported in section 6.1 and consequently iterated upon. After this early but still conceptual design phase, the final developed concept, namely concept IV of the AR application was evaluated through three heuristic evaluations, reported in section 6.2. Further information on the type of tests to evaluate the digital AR application concepts will be provided when describing the evaluation methods later on in subsection 3.2.3. It is important that in the beginning of the design phase, the functions rather than the layout of the display are investigated. It is stressed again that the process is highly iterative as can be seen in Figure 3.8 and Figure 3.9 and is aimed at understanding and meeting the users needs.

It should be noted that for starters several digital concepts of the AR application have been created in Unity, excluding the IoT integration. After evaluating the AR concepts, the integration with IoT devices was carried out and tested. Further information on the integration and testing can be found in subsection 3.1.2.

Concept I

The first digital concept was kept simple, it provides with basic functionalities and allows for the addition of further planned capabilities during the development of later versions. Please note that concept I had the aim to solely provide with the required functionalities and types of user interaction and did not focus at all at making an attractive user interface (UI) design. To view a video demonstration of the AR-IoT application concept I, the reader is invited to click on: AR-IoT Tool Concept I Video Demonstration.

The UI consists of:

- A “Main Menu” shown in Figure 3.10 which has two options: “Tag Object” and “Instructions”. The user can select these two options using gaze input. The user is also provided with feedback upon gaze thanks to a light blue “loading bar” animation overlaid on the respective options.

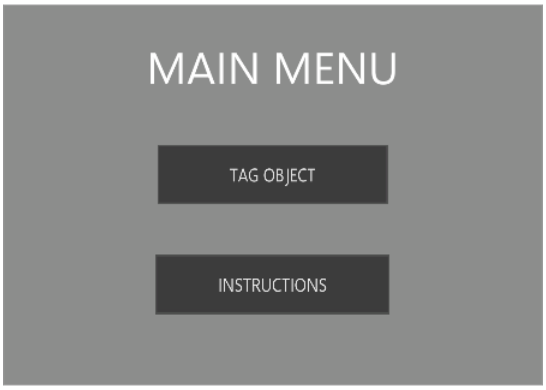


Figure 3.10: Concept I: Main Menu with two options: Tag Object and Instructions

- The “Tag Object” window shown in Figure 3.12 allows the user to: record the tag via speech-to-text conversion with speech recognition, place the tag via the “place tag” button which once selected, makes the tag follow the user’s gaze and can be placed wherever the user wants through a tap gesture (see Figure 3.11 [165]) and finally save the recorded tag. The "save" option was not yet implemented in concept I.

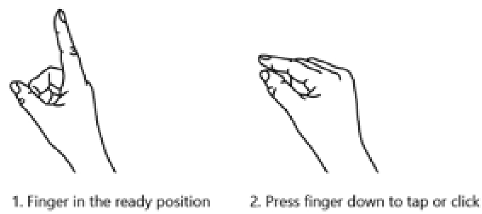


Figure 3.11: Air tap gesture recognised by the HoloLens 1st gen [165].

- The “Instructions” window shown in Figure 3.13 provides with two options: "Inspection Instructions" where the idea at the time was to implement a currently employed rock classification flowchart that allows to identify rocks for geological sampling activities, and "Inventory of Tools" which provides with information on the location of the tools. Both options were not implemented in concept I.
- Both “Instructions” and “Tag Object” windows can be closed by using gaze input on the close window button on the top right corner. Gaze feedback is provided to the user as well in this case.

Figure 3.12: Concept I: Tag Object Menu with three options: log information via voice input e.g. speech recognition, save the logged information and place the tag where desired.

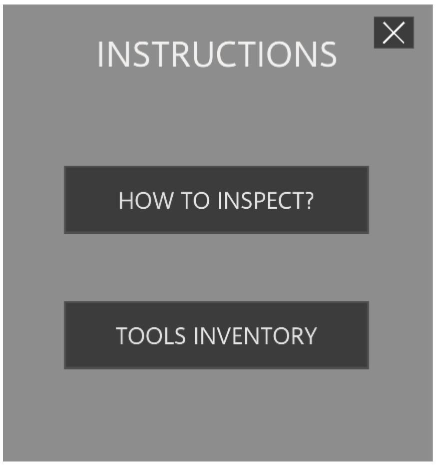


Figure 3.13: Concept I: Instructions Menu with two options: access to inspection instructions and access to tools inventory

Concept II

Based on the first three expert reviews on concept I, concept II was designed and developed. Concept II was restructured and features were added based on common site inspection activities during geological training mentioned in chapter 4. Common activities performed during Pangaea-eXtension in 2018, incorporated in the Electronic Field Book (EFB) [135] were taken as a reference for this application concept. Capabilities such as tool inventory and sample inspection guidelines were not included. In fact, geological sampling involves performing a mineral screening with additional analytical tools to retrieve the spectra [135]. Geological sampling is an activity that follows a site evaluation and for this proof of concept, the sampling capability was considered out of the scope of the project, it was however added in the menu for completeness purposes. The tool inventory, which involves the localisation of the tools used during geological field activities would require the HoloLens application to know where the tools are located, for this the HoloLens would also have to know what its location is. To enable this, a Bluetooth GPS receiver would be required to connect to the HoloLens application and send GPS data constantly to provide it with geolocation awareness, this adjustments have been made for Holo-SEXTANT [55] for instance. As geological sampling capability was for the time being considered out of the scope of this project and the tool inventory capability is required specifically for sampling activities it was also considered outside of the scope of this project. As can be seen in Figure 3.14 to Figure 3.17, the layout of the UI was enhanced for concept II. To view a video demonstration of the AR-IoT application concept II, the reader is invited to click on: AR-IoT Tool Concept II Video Demonstration.

The UI of this concept consists of:

- A "Main Menu" with the following activity options: "Stop Description", "Geological Sampling", "Health Status" and "Site Evaluation" (see Figure 3.14)

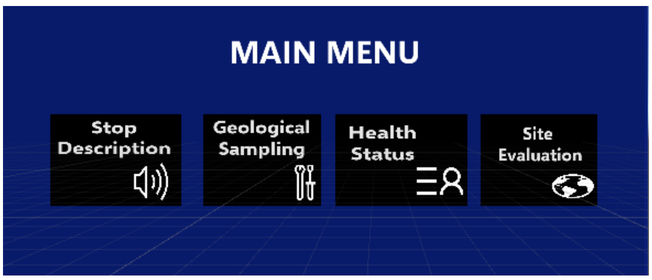


Figure 3.14: Concept II: Main menu with four options: Stop Description, Geological Sampling, Health Status and Site Evaluation.

- "Stop Description" shown in Figure 3.15 has similar logging functionalities as concept I, allowing the user to log data of the surrounding environment by recording a tag via speech recognition, saving the log and pinning the tag somewhere in the augmented environment. The save functionality was still not enabled for this concept.

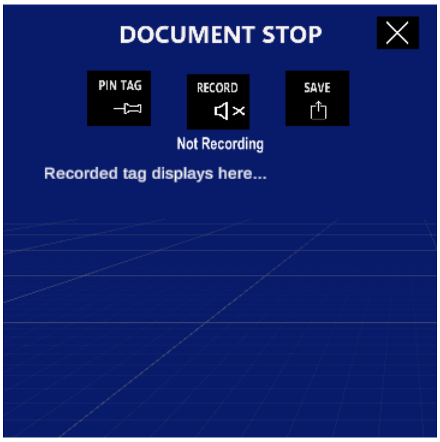


Figure 3.15: Concept II: Document Stop with three options: log information via voice input e.g. speech recognition, save the logged information and place the tag where desired.

- "Geological sampling" is an activity that was included in the set of Pangaea activities and in the EFB [135], nevertheless the focus of this concept is on site evaluation activities and the option was therefore not enabled.
- "Health Status" serves to inform the user on their physiological state and consumable levels, nevertheless this had not been implemented at that stage as the IoT network had not been integrated yet. However, it was used to showcase what functionalities could be implemented to experts reviewing the application.
- "Site Evaluation" is an activity which involves several steps according to the Pangaea training activities, namely "Geolocation Site Screening" and "Documenting a Stop". The user in this case is also offered the possibility to consult "References" which are essentially instructions. The site evaluation menu is shown in Figure 3.16.
- "References" involve instructions on how to perform a "Geolocation Site Screening", namely creating a stop, taking photos and panoramic 360 photos (see Figure 3.17).
- "Create a Waypoint" gives the user the option to place/anchor multiple flags in the augmented environment to virtually mark the sites that have been inspected by the user.
- "Take Photo" allows the user to take a photo using the embedded camera of the HoloLens and saving it in a local folder which can be accessed externally by other users through the Windows Device Portal.
- "Take Pan 360" allows the user to take a panoramic photo, more specifically a video and save it similarly to the photo to a local folder which can be accessed externally by other users through the Windows Device Portal.

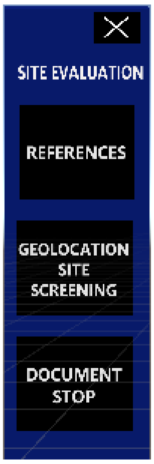


Figure 3.16: Concept II: Site Evaluation Menu with three options: access to references, access to geolocation site screening activities as well as stop documentation

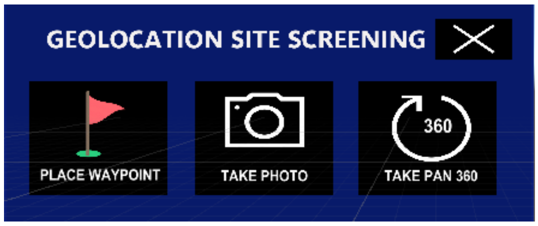


Figure 3.17: Concept II: Geolocation Site Screening Menu with three options: placing a waypoint, taking a photo and taking a panoramic 360 photo.

For concept II the HoloLens clicker (see Figure 3.2) was integrated as a means of interaction. It is used to confirm a gaze input action. The user is required to press on the clicker after the loading bar on the button is completed (a so called "action text" is displayed above the button to indicate that an action is required, namely "click to confirm"). This is to “double check” with the user whether the action is desired before the user performs an action. The double confirmation procedure was suggested and described as strictly necessary by one of the Pangaea & Caves experts who reviewed concept I.

In terms of layout the background colour of the UI was changed from grey to blue to enhance the visibility of the interface as suggested by Anandapadmanaban et al. [55]. Furthermore, the shape of the buttons was changed from rectangular to squared as suggested by one of the experts. In fact, the squared shape of a button is considered a standard. Some icons were added to the labels to make the interface more intuitive and user-friendly. Icons have been custom designed inspired by the standard icons provided by the MRTK [161] and conventional existing icons frequently used for smartphones as well as other common user interfaces nowadays.

Concept III

Concept III is an iteration of concept II, more specifically the core functionalities are the same as concept II, nevertheless some additional features have been added. Based on the feedback from two Human Machine Interface (HMI) design experts, the following changes have been applied to the UI: all menu windows have been arranged in front of the user in a circular concave manner to allow the user to maintain a good overview. All buttons have an icon and a label, furthermore to achieve consistency all icons and labels have been adjusted to be of the same type, colour, size and arrangement. Custom icons created are displayed below (see Figure 3.18 - Figure 3.29). The black on white version is shown here, however the white on black is required for the HoloLens and is the one implemented in the application. It should be noted that the icons integrated in the application that are not displayed below such as the cross for the quit button, the delete icon, the move forward icon, the check-mark icon, the photo icon and the Earth icon for the site evaluation button have been retrieved from the Mixed Reality HoloToolkit [161].

Figure 3.18: Microphone/Recording icon

Figure 3.19: Crossed Microphone/Stop Recording icon

Figure 3.20: Geological Sampling icon

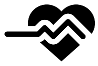


Figure 3.21: Health Status icon

Figure 3.22: Geological Site Screening icon

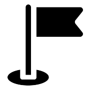


Figure 3.23: Waypoint/Flag icon

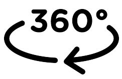


Figure 3.24: Panoramic 360 Photo icon.

Figure 3.25: Pin icon.

Figure 3.26: References icon.

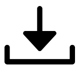


Figure 3.27: Save icon.

Figure 3.28: Document Stop icon.

Figure 3.29: Minimize icon.

Additionally, feedback on the "Take Photo" action was added once the photo would be taken.

The loading animation providing gaze feedback is faster now as suggested by MS HoloLens developers and by experts who reviewed the application. Moreover, the "action text" that appears requiring the user to click to confirm the action, was adapted such that it appears once the loading bar animation is complete instead of at the beginning of the loading, to avoid an overload of information for the user.

Additionally, the location of the pin button, which enables the users to place the recorded tag anywhere they want, was adapted and placed next to the "quit button" to signal that the action involves the "Document Stop" window. In fact, similarly to the "quit button" which closes the "Document Stop" window, the "pin button" allows the user to relocate the entire "Document Stop" window.

Regarding the creation of a waypoint, the location in which the flag would appear was adapted such that it would appear in the field-of-view (FOV) of the user.

An additional feature that was added is the possibility to input specific information related to an identified site such as: longitude, latitude and a brief site description. This information can be entered in specific input fields for the user to have a structured log document. Users are able to dictate the coordinates and these are converted into the degrees, minutes, second format when longitude and latitude are recorded. Whereas for the description input field speech-to-text conversion occurs in a standard manner. It should be noted that despite the fact coordinates would probably be filled in automatically, based on another external device, the feature was added to showcase the potential and accuracy of speech recognition and the possibility to programmatically convert these to the right format. Please note that all application scripts are available upon request.

Concept IV
Concept IV is an iteration of concept III, more specifically the core functionalities are the same as in concept III, nevertheless some additional features have been added. This time the concept has been reviewed by the same experts that reviewed concept I (see section 6.1). To view the demo of the AR-IoT application Concept IV the reader is invited to click on: AR-IoT Tool Concept IV Video Demonstration.

One of the experts suggested adding functions directly to the references, this means that the references do not only show which steps the user has to follow but they also allow the user to perform the action while consulting the references. Furthermore, as the pin functionality was driven by gaze, making it rather difficult for the user to choose the exact location where to place a window and interact with it, the same principle that applies for the flag placement is applied to the "document stop" placement, namely gaze and then click to select, drag with gaze and click to drop. Moreover, customised icons were added for all buttons for consistency (see Figure 3.30 to Figure 3.36).

Figure 3.30: Concept IV: Main Menu with four options: data logging, geological sampling, health status and site evaluation.

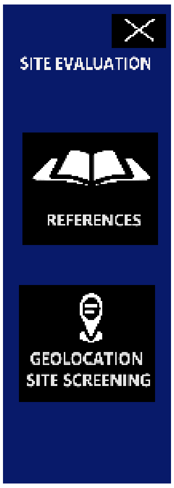


Figure 3.31: Concept IV: Site Evaluation Menu with two options: access to references as well as to geolocation site screening activities.

Figure 3.32: Concept IV: References for site evaluation activities in form
of guided procedures

Figure 3.33: Concept IV: Geolocation Site Screening Menu with three options: placing a waypoint, taking a photo and taking a panoramic 360 photo.

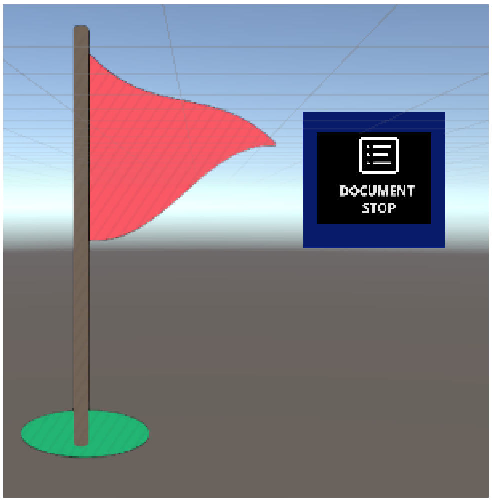


Figure 3.34: Concept IV: Flag and minimized Document Stop Menu

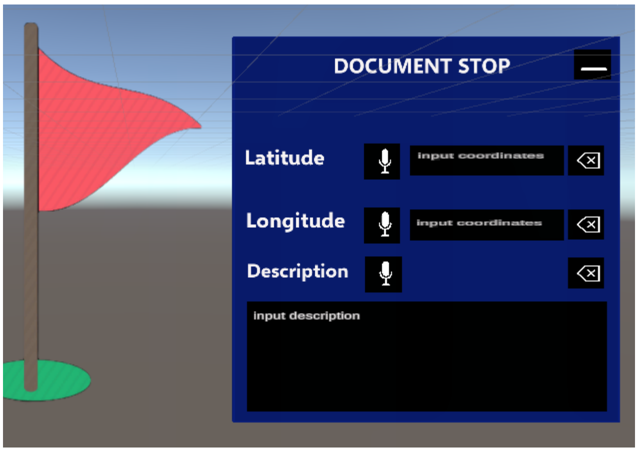


Figure 3.35: Concept IV: Flag and expanded Document Stop Menu with three options: logging latitude, longitude and a general description, all with the possibility to delete incorrect entries.

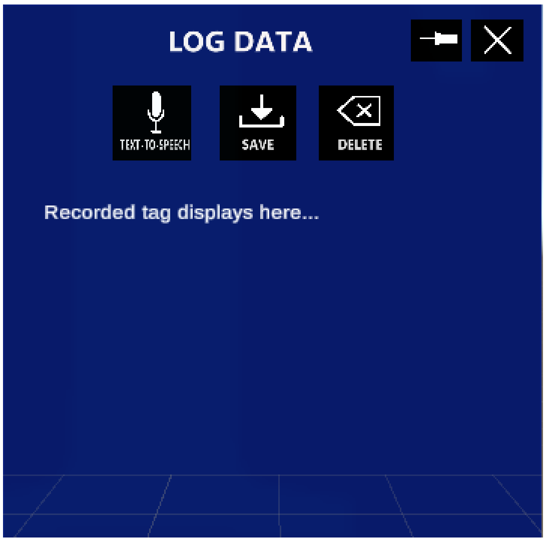


Figure 3.36: Concept IV: Log Data Menu with three options: log information via voice input e.g. speech recognition, save the logged information, delete the information in case of mistakes, and place the tag where desired.

As suggested by one of the experts, to increase the visibility of the flag, the size of it would have to be increased and the flag would have to always appear in the same location in proximity of the geolocation site screening menu such that it is in the FOV of the user and easily accessible. This has been ensured. Also, the document stop menu is now directly linked to the flag. Note that it is possible to gaze and click on the document stop button and expand the window (see Figure 3.34 and Figure 3.35).

Moreover, an erase button was added to the log data menu (see Figure 3.36) and to the document stop window (see Figure 3.35) to allow the user to be more free to control his/her actions.

As mentioned during one of the expert reviews, a crucial capability that was missing was the possibility to take a photo wherever required or desired, as the geolocation site screening menu was initially static and would not tag-along with the user. That window is now enabled to move with the user such that he or she can take a picture at any desired location.

Prototype I

Prototype I is an iteration of concept IV. Figure 3.40 - Figure 3.49 provide with an overview of Prototype I. To view the video demonstration of the AR-IoT Prototype I the reader is invited to click on: AR-IoT Tool Prototype I Video Demonstration. For a more detailed description on what can be seen in the two video demonstrations the reader is referred to section 4.2.

This prototype includes the proposed design solutions to the heuristic violations discussed during the debriefing session performed after the three heuristic evaluations (see section 6.2). The main modification regarding the user interaction is that the loading animation feedback upon gaze has been removed and replaced by a direct change of color upon gaze. With respect to the interface, modifications include:

- the addition of a "drag & drop" action text on the flag to clarify the way the user should interact with the object (see Figure 3.46 and Figure 3.47).
- the change of the label "references" in the site screening menu to "tutorial" (see Figure 3.16).
- the addition of a "go back" button to the "site evaluation instructions" to increase the consistency and the user's freedom to control (see Figure 3.42 - Figure 3.44).
- the increase in font size of the feedback prompt "photo taken" to enhance visibility.
- the enabling of the "save" functionality such that the text recorded via speech recognition can be saved to a .txt file in a local folder accessible via the Windows Device Portal (see Figure 3.49 and Figure 3.47).

- the addition of an extra deletion confirmation prompt when the user gazes and clicks on the delete button to increase the user's freedom of control.

At this stage, a consistent and user-friendly plan for the displaying of the button states and feedback has to still be created and implemented.

After the implementation of the solutions discussed, an MQTT client has been implemented in the AR interface. This client can subscribe to a certain topic and receive strings via MQTT protocol. The transmission of data, namely strings, has been tested with an open source message broker called Mosquitto [196] that implements the MQTT protocol and an MQTT publishing client, which sends dummy sensor values, on a Raspberry Pi. It is possible to successfully send strings to be displayed on the HoloLens. For more information on the AR IoT integration, please refer to subsection 3.1.2.

The proposed icons for the save, delete and tutorial functionalities have been designed and are illustrated in Figure 3.37, Figure 3.38 and Figure 3.39. The black on white version is shown here, however it should be noted that the white on black is required for the HoloLens and represents the version implemented in the application.

Figure 3.37: New design for the delete icon for the HoloLens application proposed after the heuristic evaluation.

Figure 3.38: New design for the save icon for the HoloLens application proposed after the heuristic evaluation.

Figure 3.39: New design for the tutorial, former reference, icon for the HoloLens application proposed after the heuristic evaluation.

It should also be noted that all standard icons used such as the save, camera, microphone, minimize, delete, help/tutorial, log/document data and check-mark icon are in line with the standard icons known by the target group, namely astronaut crew and the astronaut training division, mentioned in the SSP 50313 "Display and Graphics Commonality Standard" Document [181].

Figure 3.40: Prototype I: Main Menu with four options: Stop Description, Geological Sampling, Health Status and Site Evaluation.

Figure 3.41: Prototype I: Site Evaluation Menu with two options: access to a geolocation site screening tutorial, namely guided procedures and the geolocation site screening set of activities.

Figure 3.42: Prototype I: Instructions in form of guided procedure (step 1: place a tag/waypoint).

Figure 3.43: Prototype I: Instructions in form of guided procedure (step 2: take a photo)

Figure 3.44: Prototype I: Instructions in form of guided procedure (step 3: take a panoramic 360 photo).

Figure 3.45: Prototype I: Geolocation Site Screening Menu with three options: placing a waypoint, taking a photo and taking a panoramic 360 photo.

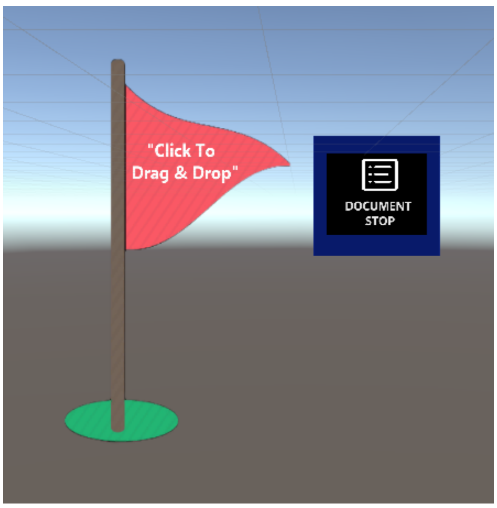


Figure 3.46: Prototype I: Flag and minimized Document Stop Menu

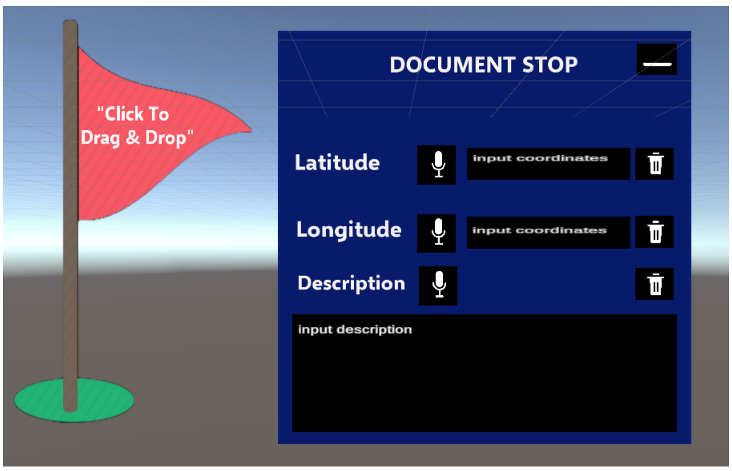


Figure 3.47: Prototype I: Flag and expanded Document Stop Menu with three options: logging latitude, longitude and a general description with the possibility to delete incorrect entries.

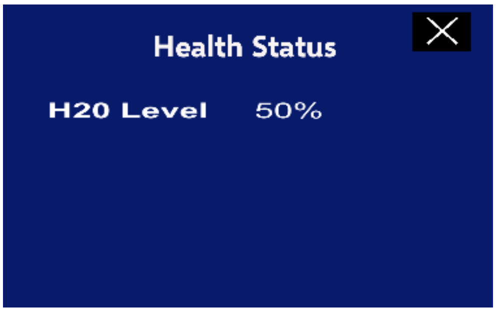


Figure 3.48: Prototype I: Health Status

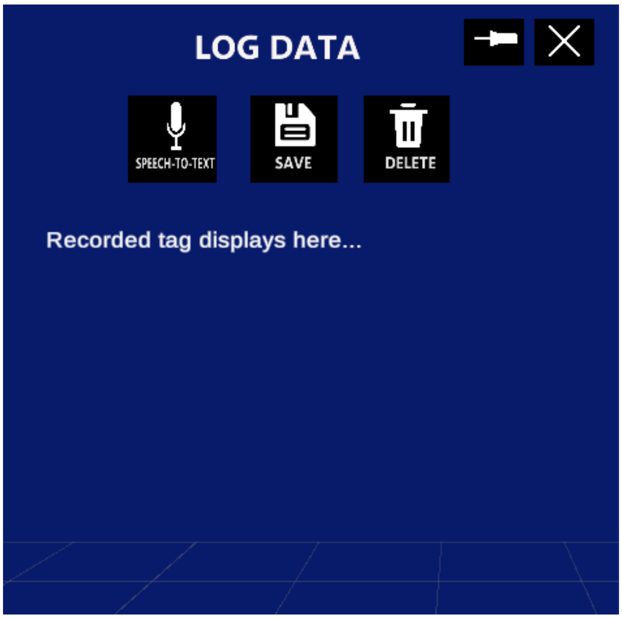


Figure 3.49: Prototype I: Log Data Menu with three options: log information via voice input e.g. speech recognition, save the logged information, delete the information in case of mistakes, and place the tag where desired.

3.2.3. Evaluation

As mentioned earlier, the user-centered design approach is a highly iterative process involving constant evaluations and iterations after the creation of each concept and prototype. In this subsection, the methodology used for the evaluation of the different concepts and prototypes is explained. The initial plan included controlled experiments with human subjects, nonetheless this plan had to be disregarded for the time being due to exceptional circumstances caused by Covid-19. Nevertheless, the planned concept of operations to be adopted for the experiments is described in chapter 5, while more details on the comparison scenario and the planned controlled experiments are described in Appendix A. As mentioned by Lee et al. [96] during the initial design phase, concepts are created and evaluated through heuristic evaluations, after that prototypes are created and evaluated through usability tests, then the pre-production model is created and evaluated through user studies and finally the product can be released and is evaluated through in-service monitoring. For this research, an orientation phase that included expert reviews before heuristic evaluations has been regarded as necessary and crucial and was therefore included.

The expert reviews held in the conceptual phase after the creation of concept I, II and III were especially helpful to answer certain research subquestions posed at the start of the project, described in chapter 1, more specifically with respect to aspects concerning: which activities are relevant during astronaut geological field work and which are feasible to be designed for in the user interface as well as the type of information that is required by the user to accomplish the given geological site evaluation task. The reviews with experts also served to answer the subquestion regarding what kind of interaction with the AR-IoT tool is considered feasible for astronaut crew during lunar and planetary EVAs.

The heuristic evaluations were required to understand whether important user interface design principles were fulfilled, this was necessary to answer one of the research subquestions posed at the start of the project, described in chapter 1, more specifically whether the identified heuristic principles are fulfilled. The process on the heuristic evaluations started with the identification of relevant design principles which helped in answering another research subquestion posed at the start of the project, described in chapter 1, on which design principles among others which heuristics are considered relevant. It should be noted that the design process and evaluation for this research project reach the prototype creation phase.

Besides the expert reviews performed in the conceptual phase and the heuristic evaluations performed before the development of prototype I, the evaluation included expert requirements compliance questionnaires and in-depth interviews. The questionnaires were sent out exclusively to astronaut geological field training experts including ESA astronauts and support engineers. Similarly, semi-structured interviews were performed with ESA-EAC experts to get additional insights on important usability, user interaction and implementation aspects for new technologies such as augmented reality. The expert requirements compliance questionnaires together with the interviews helped to answer all research subquestions posed at the start of the project, described in chapter 1, in a more targeted manner. Moreover, these evaluations helped specifically to answer the subquestions on whether the AR-IoT tool is perceived as useful, helpful and operationally feasible for future geological site inspection activities by astronauts and astronaut geological field activities experts when compared to the current tools and media. In addition, they helped to answer the subquestions on relevant and feasible concept of operations (ConOps), as well as on whether it is believed that the mental workload is reduced when the AR-IoT tool is used compared to the current tools and media. It should be noted that the adopted evaluation strategy is, as mentioned earlier already, in line with the strategy usually adopted by space-related AR studies, namely to gather feedback from the users through direct observations, surveys and questionnaires (see Table D.1 in Appendix D).

Expert Reviews
In section 3.2 the methodology used for the developed user interface has been explained. It is mentioned that several digital concepts have been created before the first prototype, which was supposed to be tested with human test subjects during controlled experiments, could be created. Every concept was subject to two to three expert reviews, including Human Machine Interaction (HMI) and geological astronaut training experts.

The expert reviews were held in the form of open informal discussions between the interface developer and the expert. During the expert reviews, the experts were instructed on the actions to perform with the interface and afterwards they were asked to mention positive and negative aspects of the interface as well as suggestions for further implementations (see section 6.1). Informal open discussions were chosen for these reviews as it allowed for a fruitful and productive exchange of ideas, which is crucial in the initial conceptual design phase. These expert reviews were extremely important in the conceptual phase of the interface development, as only with the input from the Caves & Pangaea experts it was possible to develop a useful application and, only thanks to the input of the HMI as well as human and robotic exploration experts, it was possible to develop a feasible interface that is consistent with standard display design principles. The results of these expert evaluations are shown in section 6.1.

Heuristic evaluations
After several expert reviews which led to a more defined concept, namely concept IV, three heuristic evaluations were performed. Since evaluators tend to miss certain aspects and usually different evaluators identify different aspects, Nielsen [104] suggests using a minimum of three, preferable five evaluators to perform heuristic evaluations. Due to time constraints and a limited number of design experts available, three evaluators were included in the heuristic evaluations.

First of all, a checklist with the design principles, also called "heuristics", that were considered relevant for the interface at stake was created. The checklist was created based on the 10 Usability Heuristics for User Interface Design identified by Nielsen. These general principles are listed below [105]. Furthermore, the 15 display design principles mentioned in subsection G.6.1 as well as the requirements mentioned in NASA's SPACE FLIGHT HUMAN-SYSTEM STANDARD VOLUME 2: HUMAN FACTORS, HABITABILITY, AND ENVIRONMENTAL HEALTH document [174] were considered.

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use

8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

The complete checklist includes the following "heuristics":

1. The **system status feedback** is visible for every user action (1)
2. The **response times** are appropriate to the user (1)
3. The **icons** are concrete and familiar (2), [V2 10044, NASA STD-3001 vol.2]
4. The user is **free to control** their actions (3)
5. The user interface is **consistent** (4) [V2 10038, NASA STD-3001 vol.2]
6. The user interface is **legible** (4) [V2 10039, NASA STD-3001 vol.2]
7. **Font size and type** ensures acquisition, readability, and interpretability of the display allowing for timely and accurate processing of information. [V2 10050, NASA STD-3001 vol.2]
8. The **prompts** are brief, unambiguous and imply the user is in control (6, 7)
9. The **color and brightness** contrast is good (6)
10. **Instructions** for use of the system are visible or easily retrievable (6, 8)
11. Dialogues contain **relevant information** to the current task (8) [V2 10041, NASA STD-3001 vol.2]
12. Displays and controls are **visible and easily accessible** [V2 10029, NASA STD-3001 vol.2]
13. The system provides the display area to present all critical task information **within the user's Field-of-View (FoV)** [V2 10037, NASA STD-3001 vol.2]
14. The displayed information is **relevant, sufficient, but not excessive**, to allow the crew to make decisions and perform the intended actions [V2 10040, NASA STD-3001 vol.2]

It should be noted that the reference to the principles mentioned by Nielsen [105] is attached in round brackets to every selected principle in the checklist as well as a reference to the requirement specified in NASA's SPACE FLIGHT HUMAN-SYSTEM STANDARD VOLUME 2: HUMAN FACTORS, HABITABILITY, AND ENVIRONMENTAL HEALTH document [174] in square brackets.

Once the checklist was reviewed by an HMI expert at ESA-EAC, heuristic evaluations on the final concept, namely concept IV, could be performed.

The heuristic evaluations were performed by having each individual evaluator inspect the interface alone. To ensure that the evaluations were independent and unbiased only after all evaluations had been completed, the evaluators were allowed to communicate and have their findings aggregated. The results were recorded by the evaluators and the observers separately. The evaluator was requested to fill out the checklist with the 14 "heuristics" mentioned above and rate each principle on a 5-point Likert scale [232] (strongly disagree, disagree, neutral, agree and strongly agree) as well as add comments and explanations to each rated principle. The observer recorded the comments and discussions, as the evaluator was going through the interface, to ensure that each rating was clearly argued and reasoned. In fact, it was not sufficient for evaluators to simply say that they did not like something [104]; they were required to explain why they did not like it with reference to the heuristics .

During the evaluation sessions, the evaluators were free to choose the approach they liked the most. Some decided to first go through the interface thoroughly and then rate and provide feedback on the interface via the checklist all at once while going through the interface again. Others decided to go through the interface while evaluating each principle separately, inspecting the various dialogue elements and comparing them with the list of recognized usability principles, the so called "heuristics".

As mentioned by Nielsen [107], since the evaluators are not using the system as such, namely to perform a real task, it is possible to perform heuristic evaluations of user interfaces that are in a very early stage of the usability engineering lifecycle. As a real task did not exist yet and could not be given to the evaluator, the observer, which in this case was the interface developer, was allowed to assist the evaluators in operating the interface in case of issues as well as aid the evaluators and answer their questions as they had limited domain expertise and needed to have certain aspects of the interface explained. In fact, heuristic evaluations do not have to be performed with users that are part of the target group (e.g. astronauts, trained field geologists) but with experts in user interface design. This is in line with the guidelines for heuristic evaluations provided by Nielsen [104].

As mentioned earlier, to ensure independent and unbiased results, the evaluations were performed by each evaluator separately and only afterwards a de-briefing session with every evaluator including the observer, which in this case was the interface developer, was held. The ratings of the checklist of each evaluator were merged to get a better overview. Then the comments were merged to get of rid of redundant feedback and allow for a time-efficient but complete discussion session. The de-briefing session was conducted primarily in a brainstorming mode and focused on discussions of possible redesigns to address major usability problems and general problematic aspects of the design. Each point on the list of usability heuristics and related issues were discussed separately and solutions were recorded. The results as well as the outcome of the brainstorm de-briefing session are reported in section 6.2.

Questionnaire for astronauts and astronaut geological field training experts
A customized questionnaire for astronauts and astronaut geological field training experts, to assess whether the developed AR-IoT application is in line with the specified usability requirements and whether it has potential to be used by them in the future, was created. It should be noted that this assessment approach is often used in the space industry, in fact it is frequently the case that astronauts are approached by hardware and software developers to give feedback on a product, which is already developed to a certain level, by checking its requirements compliance.

The questionnaire is based on different guidelines, standards, design principles and requirements reported hereafter:

- Usability and learnability requirements that originate from guidelines and standards specified in the SSP 50313 - Display and Graphics Commonality Standard Revision F Document [181];
- Usability principles proposed by Nielsen [103];
- The System Usability Scale (SUS) [93], [2] (see Appendix E);
- A cross-check with the heuristic design principles (see Table B.1 in Appendix B) identified for the heuristic evaluations has been made to ensure that all principles that are considered applicable and feasible for this evaluation are addressed in the usability questionnaire for the experts.

It should be noted that only application relevant and evaluable aspects have been included in this questionnaire. Therefore, only the considered feasible and relevant usability and learnability aspects from the SSP 50313 document [181] have been included, in fact no requirements on response times, input control, multiple displays and error recovery were included. Moreover, certain usability aspects of the SUS could not be included for this test as ease of use, confidence level of the user while using the application and cumbersomeness of usage cannot directly be assessed without having the evaluator wear the HoloLens with the application running on the device. In fact, experts were asked to rate the requirements compliance on the basis of a video demonstration of the application concept IV (the latest recorded version available at that time due to Covid-19 circumstances). It should be noted that not using the HoloLens for these evaluations was an external constraint (Covid-19) and not a choice. Similarly, aspects from the heuristic evaluation related to adequate colors and brightness, field-of-view and response times are difficult to evaluate under the given circumstances, hence they have not been addressed in this questionnaire. Finally, the main goal of this questionnaire is to understand the potential in terms of operational feasibility, helpfulness and usefulness of this tool for future geological field activities, hence these aspects have been added in form of customised requirements.

Before sending the questionnaire to the experts, two dry-runs were performed. The questionnaire and video demonstration of the application concept IV were sent to two subjects with no experience with the use of augmented reality, this was done to see whether all items in the questionnaire are rateable. The dry-runs were successful as all questionnaire items were rated.

The questionnaire for the experts includes the following items, each item is supposed to be rated according to the 5-point Likert scale [232] (strongly disagree, disagree, neutral, agree and strongly agree). Moreover, experts were free to leave any comments, feedback or suggestions they might have next to each item. There were no restrictions on the length of answers, as respondents were given as much space as they required. Moreover, no particular efforts were made to encourage participants to add comments, feedback or suggestions. It should be noted that references to each requirement specifically selected from the SSP 50313 document [181] have been added.

1. The application is **useful and helpful** for future astronaut geology field training and geology field exploration activities on the Moon.
2. The application is **operationally feasible** for future astronaut geology field training and geology field exploration activities on the Moon.
3. Only data that is **important to mission success and significant** in terms of crew interface is provided (SSP 50313).
4. The overall display design is based on the geological site inspection tasks that will be performed with the display (SSP 50313).
   (i) Specific data shown, the **display layout and groupings**, and the choice of display **elements** is **driven by operational requirements.**
   (ii) Information is **logically grouped** according to purpose, function, or sequence of use (e.g., either a left-to-right or top-to-bottom orientation).
   (iii) The display follows operational flows and allows **task completion with the minimum number of actions.**
5. The display is **consistent** when grouping/ordering display elements (SSP 50313).
6. The interface elements, colors and provided feedback are **consistent**.
7. The interface elements (e.g. text, icons, labels, objects) are **clear and relevant**.
8. **Information density** is held to a **minimum** in displays used for critical geological site inspection tasks (SSP 50313).
9. Primary information required for performing a geological site inspection task is on a **summary display** (SSP 50313).
10. **Information layering**, via secondary displays or dialog boxes, is implemented to provide supplemental information in support of the primary display (e.g., specify options available to the crewmember or to provide details) (SSP 50313).
11. The interface is designed for **efficient use of crew time** and to **minimize** crew and flight controller **training time** (SSP 50313).
12. The number of **user inputs** e.g. gestures/voice/gaze needed to perform simple or routine functions is **minimized** (SSP 50313).
13. A **help function** is accessible to the crewmembers (SSP 50313).
14. The display design facilitates **error-free operations** (SSP 50313).
15. **Data** is **protected** from inadvertent errors and hardware failures e.g. frequent saves (SSP 50313).
16. When a process is initiated or completed, crewmember **feedback** is provided (SSP 50313).

17. When an input is required, an indication is provided to the crewmember, e.g. a cursor change (SSP 50313).

18. If the completed command implies the need for further crewmember action, the **need for action is indicated** (SSP 50313).

19. The application responds to crewmember interaction with **appropriate feedback** (SSP 50313).

20. **Navigation** through the interface is clear and intuitive.

21. The application **minimizes the user's mental workload**.

22. The interface is **satisfying** for crewmembers and training members.

23. The required information is **easily found and accessed**.

24. It is possible to understand what the features of the application represent and to realize the **integration** of the different media.

25. The application is **"astronaut crew-friendly"**.

26. The application has the **potential to become "flight proven"** assuming that augmented reality technology will be an integral part of the astronauts' EVA helmet.

ESA astronauts which have either been both trained for EVAs through analogue missions and performed EVAs in the past on the ISS or have only been trained for EVAs through analogue missions as well as participants and/or support engineers of Caves & Pangaea analogue missions were contacted via email. Following a brief introduction to the research project, they were asked whether they would like to participate in this evaluation by filling in the questionnaire based on a video demonstration of concept IV (the latest recorded version available at that time due to Covid-19 circumstances). To view the video demonstration of the AR-IoT Concept IV the reader is invited to click on: AR-IoT Tool Concept IV Video Demonstration. Finally, additional information which consists of a brief explanation of the assessment was provided as follows:

> *"A short introduction of the application that you will be assessing will be given first, then the aim of the tool will be explained, and a brief description of the features is provided to clarify what you will be seeing in the demo. After that you are requested to watch the demo as many times as you wish, finally I kindly ask you to fill out the questionnaire."*

The introduction is as follows:

> *"The video (2 minutes) presented to you is a demo of the Augmented Reality (AR) Application proof of concept I am developing."*

The aim is explained as follows:

> *"The aim of the tool is to ease future lunar surface exploration activities more specifically geological site evaluation activities. The site evaluation activity involves different actions, namely a stop description, a geolocation site screening, an analytical screening and documenting the stop and the area. The tool focuses on all aspects except the analytical screening which involves performing a mineral screening with some kind of analytical tool and communicating to the support team the spectra found. In fact, this action could not be included in the tool as it was considered out of the scope of this proof of concept. "*

The description includes the list of features of the AR-IoT tool as described in chapter 4.

The data from the questionnaire was analysed with Python and validated with the statistical analysis software, SPSS [87].

The Likert scale is considered to be an ordinal scale, hence responses can be ranked or rated, however the distance between responses cannot be measured [73], [213]. Therefore, the difference between responses cannot be assumed equidistant despite the numbers assigned to those responses are [73]. Due to this observation, experts over the years have claimed that for ordinal data, the median should be employed as the measure of central tendency [73], [213] as the required arithmetical manipulations to calculate the mean for ordinal data are inappropriate [213]. Additionally, ordinal data can be explained using frequencies/percentages of response in each category [213]. Therefore, the median is calculated to describe the data obtained from the questionnaire as well as the frequencies/percentages of response for each point in the questionnaire. To measure the dispersion in the responses the Interquartile range (IQR) is calculated.

The comments that respondents left were first explored to understand what contribution these could make to the study overall. It was then decided that an analysis of the feedback responses was desirable for several reasons: the comments provided with valuable additional insights related to the items in the questionnaire; a large number of participants took the time to write comments; and because of the strength of personal opinions expressed in many of the comments. All results of the questionnaires are reported in section 6.3.

Interviews with astronauts and astronaut geological field training experts
Semi-structured interviews based on a pre-defined protocol that consisted of open-ended questions to complement the questionnaire responses from the experts and to gain further interesting insights on the usability of the AR-IoT tool, based on the experience and opinion of the interviewees, were performed. The invitation email informed the participants that the interview would be via Skype or Zoom and that it would consist of a set of open-ended questions based on their experience and opinions.

The invitation email was sent only to experts in astronaut extravehicular activity (EVA) training, more specifically it was sent to 3 groups of people:

- ESA astronauts who participated in analogue missions including Caves & Pangaea
- Caves & Pangaea support engineers

Consent to audio-record the interview was gathered via verbal consent or written consent depending on the participant's preference prior to the interview.

As mentioned earlier, the interviews were semi-structured based on a pre-defined protocol which consisted of seven open questions. First, introductory questions were asked, these were aimed specifically at understanding what kind of experiences the participant had with augmented reality technology, astronaut geological field training and/or extravehicular activities in general during analogue missions as well as testing of and exposure to new technology during these missions. Hence, the following questions were considered fundamental:

1. What is your experience with augmented reality?
2. In how many analogue missions have you taken part in? Which ones?
3. What new technology was tested during these analogues?

Other questions focused on required and useful new technology that could enhance the preparatory training for future human lunar and planetary surface exploration as well as the actual planned missions. This question issued a set of sub-questions, these are all reported hereafter:

4. What do you think would be useful/helpful tools for astronaut geological field activities for both training and for lunar and planetary exploration?
   a. What were you missing in terms of supporting tools during analogue expeditions e.g. Pangaea-X or NEEMO to perform EVAs?
   b. What activities/tasks were very demanding in terms of mental workload?
   c. What activities were most sensible to mistakes?
   d. Do you think such a tool (AR-IoT tool) can aid astronauts in future lunar and planetary EVAs? How? Why not? When?

The following question regarding usability was included:

5. What do you think an application needs to be astronaut-friendly, specifically for EVAs?

The following questions regarding the expert opinion on the future outlook on innovative technology development and integration were addressed:

6. Do you think AR technology will become an integral part of the astronaut EVA suit?
    a. What do you think will be the obstacles for AR technology to be part of the astronaut EVA suit? When? Why?
7. What is your general opinion on this application?
    a. What is missing?
    b. Do you foresee it being implemented in the future?
    c. What do you think of the means of interaction? Gaze and click, speech recognition?
    d. Can you think of any other potential applications of the developed AR-IoT application?

These questions were formulated such that they are as open-ended and generic as possible, while keeping them "answerable" by the interviewee. Moreover, these questions were considered a guideline, the order in which they were addressed during the interview was kept flexible and opportunities were taken as soon as they would arise to go more into depth when the addressed topic was considered insightful by the interviewer. In fact, the aim of these interviews was to gather additional insights on whether and when according to the experts the application developed is useful, operationally feasible as well as usable during future analogue missions and future human extraterrestrial exploration missions. Additionally, it was investigated whether the application has the potential to become flight-proven in the future and what the obstacles are or could be.

The audio-recorded interviews were first transcribed verbatim according to the guidelines mentioned by McLellan et al. [58]. No transcription software was used for this process to avoid data privacy issues.

The software Atlas.ti 8.4 [24] was used as a support tool for the qualitative interview analysis, more specifically to enable a structured text analysis.

After a thorough review of qualitative data analysis methodologies described by Braun and Clarke [229], Mayring [186], Jamieson [214] and Vaismoradi et al. [157] the qualitative content analysis (QCA) approach was adopted. This approach was first described by Mayring [186] and is now a well established approach for text analysis. It is a suitable method to analyze interviews, to code documents and observations [148]. Mayring [187] suggests two main procedures for QCA, inductive and deductive category development. Inductive category development is data-driven while deductive category development is theory-driven. The principles for inductive category development were adopted for the coding process as the scope of this analysis is explorative. The strategy implies that the researcher's goal is the formulation of a criterion of definition, derived from the theoretical background and/or the research question(s) as a selection criterion to detect relevant material from the text [187]. A step-by-step procedure, consisting of: selection, reduction, generalization, construction, combination and integration, is systematically applied by the researcher [148], [187]. This procedure was adopted by the author.

First, the transcripts of the interviews were examined line-by-line applying usual steps of text analysis, as for instance highlighting of text, writing notes and seeking for keywords in the text [187] using Atlas.ti 8.4 [24]. The transcripts were then reread to define initial categories. As mentioned earlier, the coding process occurred using the inductive category development process as described by Mayring [187]. An overview of the steps and associated description is given in Figure 3.50.

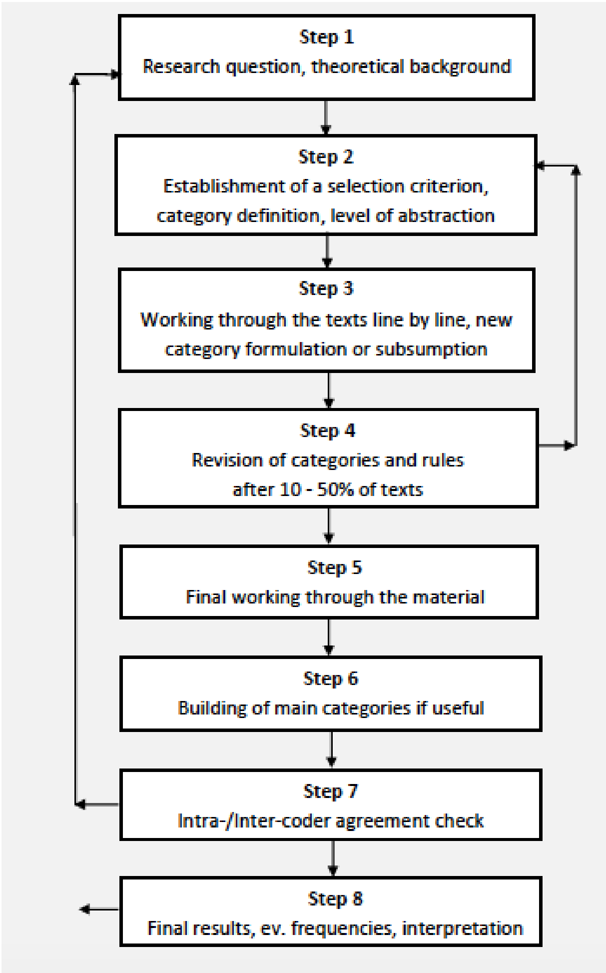


Figure 3.50: Inductive Category Development step-by-step procedure [187].

The **category definition** was formulated as follows:
Experiences and opinions in and around analogue missions related to testing and application of new technologies as well as new concepts of operation.

The **level of abstraction** is:
Experiences and opinions with new technologies focusing on aspects related to usability, user interaction and implementation. Experiences with new concepts of operation.

Secondly, the different identified categories were grouped into categories and subcategories. This step was performed deductively by introducing theoretical considerations while formulating the main categories and assigning the subcategories, as suggested by Mayring [187].

Thirdly, the number of quotations was counted as well as the number of classified quotations. It should be noted that when the quote was referring to more than one subcategory, the quote was assigned to every subcategory it belonged to. After that, subcategory frequencies and respective main category frequencies within the material were evaluated. Then, the number of respondents per subcategory was counted. It should be noted that subcategories with one respondent only were omitted.

Finally, illustrative quotes of each subcategory were selected to corroborate the validity of the analysis and to conciliate the reader of the connection between the interpretation and the evidence [56], [168], [225]. The principle of “prototypical and outlier illustrations” for each subcategory [56] was used for this purpose. The assumption was made that the importance of a subcategory is proportional to the number of times it gets mentioned by the respondents. Consequently, it was decided to report the number of quotes (see section 6.4) such that these are proportional to the number of respondents mentioning the corresponding subcategory. One quote per respondent was accepted for each of the subcategories as the maximum to avoid that a subcategory is dominantly described by a single respondent.

Illustrative quotes were therefore selected as such:

- Subcategories mentioned by 2 to 3 respondents are represented by a minimum of 1 and a maximum of 2 quotes
- Subcategories mentioned by all 4 respondents are represented by a minimum of 2 and a maximum of 3 quotes

4
AR-IoT Tool

In this chapter, the developed proof of concept of the innovative Augmented Reality Internet-of-Things (AR-IoT) surface exploration tool created to aid future extravehicular (EV) crew during geological field activities, specifically for geological site inspections, is presented. First the features of the AR-IoT tool are explained. After that, video demonstrations of prototype I, described in subsection 3.2.2, created as a result of the iterative conceptual phase (see section 3.2), are elucidated and shown.

4.1. Features

Based on research on geological field activities previously outlined in chapter 2, analogue campaigns such as BASALT [51] and Caves & Pangaea [134] as well as discussions with experts from Caves & Pangaea at the European Space Agency's Astronaut Centre (ESA-EAC), it could be concluded that an augmented reality tool for site evaluation activities had not yet been developed but is in line with future needs and requirements. It should be noted that this research project is essentially a proof of concept, therefore it was important to focus on one specific task with a set of defined sub-tasks that could easily be tested.

As stated in chapter 2, geological field work involves different activities namely: site evaluation, site preparation and documentation and sampling [130], [135]. These activities are reflected in the Electronic Field Book (EFB) as well. Thus, the hierarchical structure of these activities integrated in the EFB [135] was used as reference during the development of the AR-IoT tool, this was done keeping in mind a potential future integration of the tool in the Caves & Pangaea test campaigns. Moreover, suggested enhancements, necessary and/or desired capabilities and recommendations gathered during analogues such as BASALT-2 [51] and from the testing of the Holo-SEXTANT concept [55] have been considered.

The developed AR-IoT tool focuses specifically on the site evaluation aspect of geological field activities. The site evaluation activity involves different actions, namely creating a stop description, making a geolocation site screening, carrying out an analytical screening and documenting the stop and the area [135]. The developed AR-IoT tool focuses on all aspects except the analytical screening, which involves performing a mineral screening with some kind of analytical tool and providing the support team with the spectra found. The analytical screening support feature was not included in the tool as it was considered out of the scope of this proof of concept.

The features of the innovative AR-IoT tool will be elaborated upon hereafter. Please note that more details on the various functions, iteration and rationale behind certain interface layout decisions and means of interaction can be found in section 3.2, where the user-centered design methodology applied to create the tool's proof of concept is elucidated.

Constant access to a main menu

The main menu is shown once the application has been loaded. It gives the user the possibility to select one of the following options: log data, geological sampling, health status, site evaluation. The main menu is meant to give users the possibility to: log data at any point in time during their planned traverse, recall their health status which includes biomedical data and consumable levels at any point in time and perform site evaluations when required. It should be noted that the geological sampling option has not been enabled for this proof of concept, it was simply added for completeness purposes.

Hands-free interaction
The tool offers hands-free interaction to the user. In fact, the users avail themselves of gaze input to pick the action that needs to be performed. The chosen action is then highlighted upon selection and requires double confirmation through one simple click performed by pressing on the given clicker (see subsection 3.1.1) once, while gazing on the item of choice. It could be argued that the act of pressing a button does not represent complete hands-free interaction nevertheless, the assumption is made that the button will be incorporated in the astronaut spacesuit glove. The reason why the alternative means of interaction by voice input was not employed in the first place is because voice loops interference, as initially recommended by Apollo astronauts [197], should be avoided when feasible.

Hands-free logging of data
The log data capability allows the user to record data hands-free (e.g. via speech recognition). The data log can be geospatially pinned to the desired location at which the user has recorded the data. This way the user can create linked electronic field notes with the aim to increase efficiency and reduce workload. In fact, the users do not have to take their hands off the tools they are momentarily using. The interaction is based on gaze input and a simple button click to start the recording. While the pin function is on a drag & drop basis, more specifically gaze on the pin button, click enable dragging, gaze to drag and drop by clicking again.

Saving of logged data
The logged data, namely the text recorded and transcribed via speech recognition can be saved to a .txt file in a local folder in the HoloLens and is then accessible via the Windows Device Portal by other ground control crewmembers or intravehicular (IV) crew.

Deletion of recorded/logged information
The interface allows the user to make corrections on recorded information. The user is free to delete recorded text when necessary. Moreover, before deleting the text definitely the user will be prompted to give an extra confirmation as a window requesting a confirmation for the action will pop up.

Access on demand to consumables levels and biomedical data
The health status button provides with a simple text-based display providing the user with data on critical suit consumables and biomedical data as suggested by Stevens et al. [12] and Apollo astronauts [197].

Access to tutorials in form of references/procedures
In case the user does not recall how to perform a certain activity, in this case a site evaluation, the users can avail themselves of guided procedures. This is aimed at reducing mental workload and increasing efficiency. The user can readily perform the action required at each step without having to switch to different windows again reducing the workload and time required to perform an action.

Hands-free taking photographs and panoramic views/videos
The user can take photos of desired locations and record videos in form of panoramic views hands-free. The photographs and videos/panoramic 360 photos are automatically saved to a local folder on the HoloLens and can be accessed via the Windows Device Portal by any other ground control crewmember or IV crew.

Hands-free creation of waypoints
The user can easily and promptly mark field locations of interest without the need to take their hands off the tools they are using in the field at that specific instant in time. Furthermore, the user can readily link data to the waypoint represented by a flag, namely location information and descriptions of the area. This is meant to help extravehicular (EV) crew in mapping an area and highlighting features during a traverse while making it accessible/visible to both subsequent EV crew but also to monitoring crew inside the habitat or on ground.

Real-time live stream of HoloLens
Other crewmembers in this case IV crewmembers can potentially see anything the HoloLens user sees in-

cluding the surrounding environment and the holograms. This would allow IV crew to be able to constantly monitor EV crew and increase their situational awareness.

4.2. Video Demonstration

For the reader to be able to visualise the final product of this thesis, two video demonstrations are made available. Both video demonstrations are representing prototype I created as a result of the iterative conceptual phase described in subsection 3.2.2. A more detailed description of what can be seen in the videos is described hereafter. Finally, the links to the videos are provided.

In the first demonstration named ***Demo 1***, the AR-IoT application as a whole is shown. The reader is referred to section 2.4 for background information on the activities carried out during geological site evaluations to better understand what is shown and the rationale behind the selected set and sequence of activities. Once the application is started, the main menu with the different options to be selected namely: log data, recall the health status which includes biomedical data and consumable levels, perform site evaluations and geological sampling, is shown. Then, the site evaluation option is selected via gaze and click input. One can see that the white cursor is directed on the option through gazing, which makes the icon highlight while the click confirms the selection. Note that a clicker (see Figure 3.2) is being held in the hands at that time. The site evaluation option gives users the possibility to follow a tutorial or if users are acquainted with the site evaluation process, they can directly select the geological site screening option. The latter case is shown first in the video. After that, in the same manner, the geological site screening option is selected. The user now selects the waypoint creation option, a flag is created and by gazing at the flag, clicking, selecting the location where the flag should be placed through gaze, and by clicking again, the flag can be dropped at the desired location. The user can then record GPS data and a general description of what is seen in the area via speech-to-text conversion through voice recognition. The user is supposedly seeing basaltic rocks and records the information, moreover the user dictates the coordinates for the latitude. After that, the user decides to delete the recorded data to showcase the functionality. As many waypoints as desired to mark an area can be created and placed. The use of these waypoints is described more thoroughly in section 4.1. Finally, as is common practice for geologists on the field photo-documentation is done by taking photos and panoramic 360 photos in form of videos.

Carrying out geological site inspections aided by the tutorial option under the site evaluation menu is then shown, the importance of this feature is described in section 4.1. The same steps as described earlier are shown. Upon completion the user decides to check their health status which for the moment shows the water consumable level of the simulated portable life support system (PLSS) and shows that 50% is remaining. Then the user decides to add some extra notes based on observations of the area that is being surveyed, e.g. colour of the landscape and number of rocks. The data logging is done by recording data via speech recognition. The data log can be pinned to any landmark or placed next to a waypoint by dragging and dropping it, similar to the flag placement. Moreover, it is shown that the data can be saved and deleted if considered necessary. As described in section 4.1 the saved log is accessible via the Windows Device Portal of the HoloLens. All interactions are hands-free, note that the importance of this feature is described in section 4.1.

In the second demonstration named ***Demo 2***, the integration of the IoT in the AR application is explicitly shown. As described earlier, and specified in section 4.1, the user can access consumables levels and biomedical data on demand. This feature was achieved specifically with the integration of IoT (for more details, refer to subsection 3.1.2) into the AR application. Once the application has loaded, one can see that the user selects the health status option to view the remaining water available in the simulated PLSS, which at the start is 85%. The user then carries out the usual set of geological site inspection activities as described earlier for *Demo 1* while keeping an eye on the dropping level of the water consumable to get an idea of how much time is still left to carry out operations until an eventual return to base. In fact, finally 5% of water are remaining and operations are ended.

To view the two video demonstrations of the AR-IoT prototype I proof of concept (***Demo 1*** and ***Demo 2***), the reader is invited to click on: AR-IoT Tool Video Demonstrations.

5
Operational Scenario

In this chapter, the proposed operational scenario to be adopted during the initially planned controlled experiments detailed in Appendix A is presented. The development of this concept of operations (ConOps) was not only aimed at creating a feasible operational scenario for human subject testing but it was also essential throughout the design of the AR-IoT tool. In fact, it provides with an overall picture of the operation from the perspective of the users who will operate the tool [173]. First, the design considerations related to the operational scenario are elucidated, then the assumptions that have been made are presented, after that the operational setup including background information, task design, roles and responsibilities, test equipment, data flow, task instructions and guidelines are described.

The selection and design of this test scenario takes into consideration the Concepts of Operations (ConOps) investigated during different analog field experiences such as NASA's Biologic Analog Science Associated with Lava Terrains (BASALT) [121], the Desert Research And Technology Studies (DRATS), Pavilion Lake Research Project (PLRP) and NASA Extreme Environment Operations (NEEMO) program as well as the ConOps of the Apollo missions.

5.1. Research Design Considerations

Important considerations to be accounted for while planning a simulation and evaluating results have been made. In 2008, NASA's Directorate Integration Office hired the Exploration Analog and Mission Development (EAMD) team to manage and conduct several spaceflight analog mission evaluations (including at DRATS, PLRP, NEEMO, and BASALT) employing a set of operational methods and metrics that would allow for an iterative development, testing, analysis, and validation of evolving exploration architectures. This method ensures that the data collected stays pertinent to NASA's strategic architecture and technology development goals and allows for data-driven, actionable recommendations. For this study, key points considered for the methodology adopted by EAMD have been extracted and considered [121]. These include:

- the development of a study that integrates all required tasks to address objectives and questions as well as a plan to gather quantitative and qualitative data;
- the documentation of the assumptions made;
- the selection of test subjects that are representative of the target population;
- the execution of the research design with sufficient fidelity of the operational environment and integration of relevant technologies;
- the use of test subject consensus results to create a single set of data.

5.2. Assumptions

Important assumptions have been made to provide with a framework for the selected concept of operations or proposed test scenario.

1. It is assumed that during the execution of the EVA by crew, additional levels of information will be retrieved via contextual sampling and presampling of sites of interest which can modify the science tasks, traverse plans and science priorities [121]. In fact, the goal while conducting exploration EVAs by human crews is to retrieve as much scientific information as possible to fulfill mission science objectives [122].

2. It is assumed that a crucial element to this new vision of lunar and planetary field geology is the notion of coordinated human and robotic exploration. Currently as well as in the predictable future, it is assumed that no autonomous robots exist with the capacities of human explorers for: expert analysis of geologic associations established upon basic training and real-world experience; intelligent and quick decision-making with respect to the most critical observations to make or samples to take; on-the-fly data synthesis and the utilisation of that synthesis for traverse replanning; as well as adjudicate when satisfactory observation and sampling has been performed [130].

   At the same time, human exploration is risky and necessitates the use of limited resources in an optimal manner (e.g., oxygen and water) to keep even primary life support for the astronaut crew. Resource requirements are more easily managed for robots. When remotely controlled by humans or with short-term autonomy, robots can perform notably well when it comes to: systematic collection of observational data, such as different types of imagery; systematic sampling, in-situ analysis, and the archiving of samples which necessitates straightforward functional capabilities; deployment of scientific experiments; and repetitive, rule-based surveys [130].

3. It is therefore assumed that robotic precursor missions, both orbital and on the surface will have collected imagery and data which determined the plan for the science operation extravehicular activity (EVA) traverses [91], [121] [130].

4. It is assumed that crewmembers are provided with geological site inspection assistance tools such as an Augmented Reality Internet-of-Things interface.

5. It is assumed that communication issues including latencies of 2.5 seconds in the case of a lunar exploration scenario, interferences and interruptions will be occurring [121] between extravehicular (EV) crew and Mission Support Centre (MSC).

6. It is assumed that there are two EV crew on site [121]. This does not imply that the simulated operational scenario has to be carried out with two test subjects. One can assume that the other EV crew (EV2) has other tasks to perform that are not related to the activity of EV1. In certain currently predicted future lunar scenarios, EV1 would be in charge of operations and EV2 in charge of science [121]. However, thanks to new innovative technologies that can relief the mental and physical workload of the EV crew, it is considered realistic to assume that EV1 and EV2 are in charge of both operations and science just in different locations to maximise the scientific outcome. Hence, there is no need to be specialised in one activity only and require the direct support from the other crewmember but instead rely on the support of innovative aiding technology.

7. It is assumed that during real operations on the Moon, two intravehicular (IV) crew are located on the Human Landing System (HLS) and assist the two EV crew during surface operations. In that case, EV1 would be in charge of operations and EV2 in charge of science, two IV crew would be needed as IV1 would be in charge of timelines, operational tasks, procedures and constraints, whereas IV2 would be the science lead, and primarily interact with the EV crew and MSC (via SCICOM, Science Communicator) on science tasks, priorities, and recommendations. In fact, the HLS is planned to be designed for four crewmembers [91], moreover, the Lunar Gateway is also planned to host a total of four crewmembers and so is the Orion spacecraft [172]. Nevertheless, it is possible that the Gateway will be expanded and crew will be present on the Gateway while four crewmembers are on the surface. In case EV1 and EV2 are both in charge of operations and science, specialised IV1 and IV2 crew in charge of both operations and science respectively, could still be useful. Another possibility could be to have personalized assistance in both science and operations from IV1 to EV1 and from IV2 to EV2 provided that IV1 and IV2 avail of innovative aiding and guiding technology as well. It is however also feasible and realistic to assume that as during the Apollo missions crew is assisted by the Mission Control Centre (MCC).

8. It is assumed that MCC assists crew on surface rather than IV crew. This can be reasoned by the scope of this research project namely assessing whether the designed AR-IoT tool can aid crew operations under communication issues including latencies and interferences. Additionally this is in line with the concept of operations during the Apollo missions [57]. It should be noted that during the BASALT analogue it was assumed that MCC is limitedly available to assist crew on the surface. In fact, IV crew was assumed to communicate with an Earth-based Mission Support Centre (MSC) that supplies with scientific expertise and operational guidance across communication latency and bandwidth limitations [51]. It should be noted that Mission Control Centre was then Mission Support Centre to reflect the advisory role of the team in case of limited availability when working in presence of communication delays [121]. Nevertheless, the focus of the BASALT analogues was specifically human exploration on Mars.

9. It is assumed that operations take place during Artemis III, hence in 2024, where the first crew lands on the Moon with a lunar lander also known as HLS [180].

10. It is assumed that there is no habitat or base in place on the lunar surface yet by the year 2024 [180].

11. It is assumed that the spacesuit of the crew is equipped with sensors able to measure consumable levels, medical health data and environmental data. These sensors belong to an IoT network which ensures communication with an augmented reality interface and/or monitoring displays integrated in the suit.

12. It is assumed that there is only one voice loop between EV crew and MCC for the specific concept of operation chosen. In fact, as IV crew will not be playing a crucial role in this case a voice loop for IV and EV crew as well as for IV and MCC is not necessary as EV communicates directly with MCC. It should be noted however that during BASALT analogues for example, it was assumed that there are two voice communication loops one between IV and EV crew which happens real-time and is open for listening by MSC, and one between IV and MSC which is affected by communication latencies and interruptions. From the acceptability assessments of BASALT analogue missions [122] it appeared that voice communication between IV and EV crew was essential/enabling while voice communication from MSC to IV was said to be providing with little or no enhancement by EV crew as the messages were received across latency. EV crew discouraged voice communication from MSC as these kind of interactions were perceived as distracting or even detrimental to the science and operations flow. Communication between IV crew and MSC was essential. Nevertheless, for this research it is considered acceptable to assume that EV crew receives direct instructions from MCC. In fact, the focus is on the assessment of the capabilities of new technology by EV crew rather than on the whole operational scenario architecture.

13. It is assumed that there is no impact on the results caused by the fact that the subjects are not wearing a spacesuit as the planned comparison scenario with conventional media used during Apollo missions to assess the AR-IoT would just like the scenario involving the AR-IoT tool not involve wearing a suit. In fact, it is assumed that the disadvantages due to the use of a suit affect the two scenarios equally.

14. It is assumed that a wireless network connection is available.

5.3. Operational Setup

In this section, first background information on concepts of operations drawn from previous analogue missions is given, after that the task design or the set of EVA tasks is shown, then the roles and responsibilities together with the test equipment, data flow, task instructions and guidelines are elucidated.

5.3.1. Background Information

This operational scenario draws its framework from several simulated scenarios during the BASALT analogues [121], the geological site inspection activities described by Hodges and Schmitt [130] and the set of activities performed during a site inspection in the Pangaea-eXtension, also known as Pangaea-X, campaign [135]. During Pangaea-X in 2018, several different approaches for the scenarios were simulated, tested and analysed [134]. Different sites, with specific geologic and environmental characters were made available within the Lanzarote Geopark. A complex logistic framework was set up with the aim of maximising efficiency and allowing an interconnection between the different experiments. The specific goal of the campaign was to evaluate potential applications and developments for future mission scenarios and to increase operational relevance of training [134]. One of the technologies tested during the Pangaea-X campaigns is the Electronic Field Book

(EFB) for tablets. The European Space Agency's Astronaut Centre (ESA-EAC) initiated the development of the EFB, with the aim of integrating geological reference, sampling protocols, traverse preparation, navigation and documentation, analytical tools as well as providing real-time awareness to support mission control personnel and scientists during operations. Please note that the functionalities provided by the EFB have been previously outlined in section 2.4. For the reader to better understand the set of actions involved during site evaluation activities, it is highly advised to refer to section 2.4 where these are described.

Finally, it should be noted that several verbal discussions to analyze and validate the relevance and feasibility of the chosen operational scenario, for this research project specifically, were successfully held with Caves & Pangaea geological field training experts at ESA-EAC.

5.3.2. Task Design

An overview of the different tasks to be performed by EV crew and MCC as well as the media needed during the operational test scenario, in chronological order, is shown in Figure 5.1. It should be noted that the scenario outlined includes the use of the HoloLens. The reader is referred to Appendix A for more information on the planned comparison scenario with conventional media. At the start of the EVA, the EV crew provides with a brief report of system checks before egressing the airlock and then prepares for airlock egress. After that, EV crew egresses the airlock. With the assumption that a rover has already mapped the area and a target area for the crew has been determined, the crew starts the traverse by moving to the target area. During the target area approach phase, the EV crew evaluates the surrounding terrain and determines whether there are important features that characterise the area and/or sites to be noted and reported. The EV crew also determines whether there might be relevant areas for other science objectives or sites that might be worth visiting later or returning to during another separate EVA. On arrival at the predetermined target area, EVA crew creates a site and conducts a verbal contextual survey from the distance and then from different perspectives. Consecutively, EV crew records relevant imagery and video footage. Finally, after checking the completeness and correctness of the delivered documentation crew returns to the airlock.

During analogues such as BASALT, the visitation order of the different stations was prioritized in such a way that the first two stations were the ones with the highest priority whereas the third represented a secondary objective [14]. To allow for some flexibility in responding and adapting to new discoveries during an EVA [130] targets of opportunity (TOP) were introduced. TOP are features or sites of interest identified by EV crew during a target area approach or within the target area which had not been identified as a specific EVA objective before [14]. These TOP have the potential to become a top scientific objective.

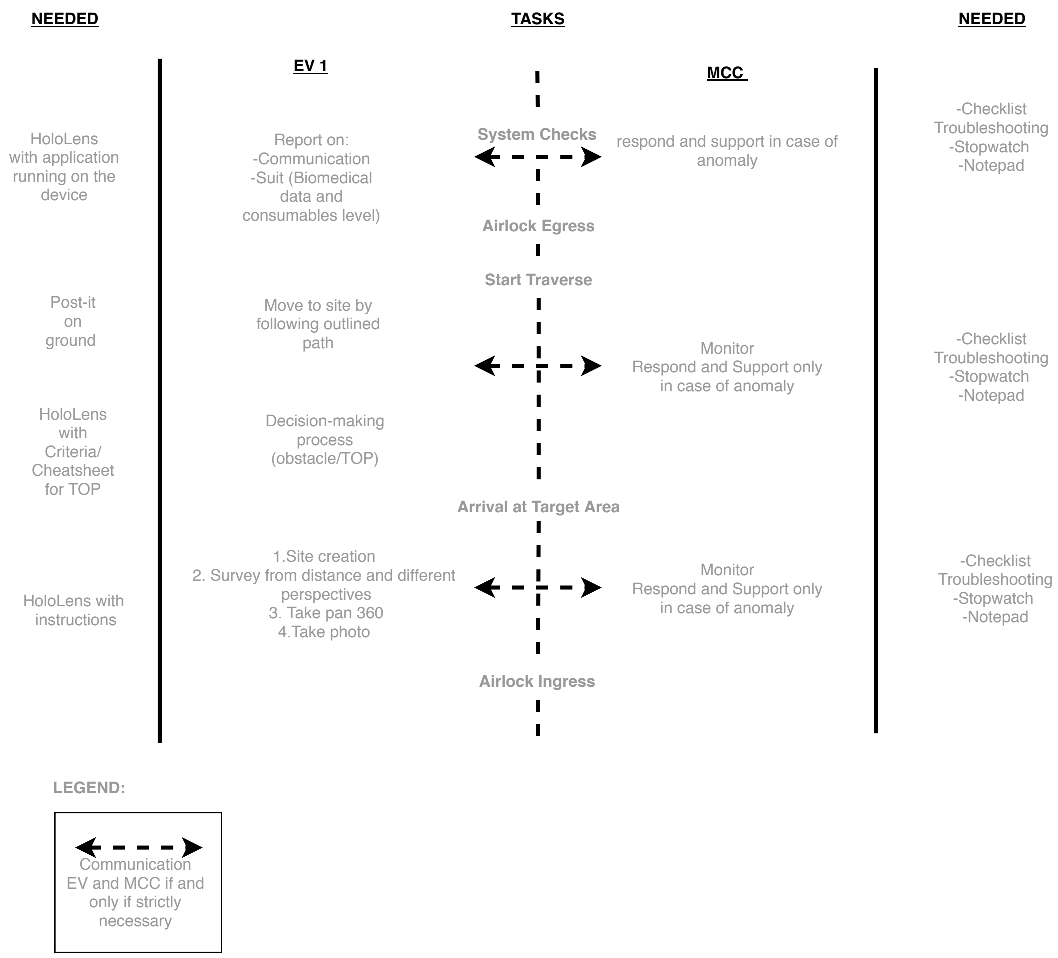


Figure 5.1: Extravehicular crew (EV1) and Mission Control Centre (MCC) tasks in chronological order throughout the different EVA phases and the required instructional media. It should be noted that this scenario includes the HoloLens.

5.3.3. Roles and Responsibilities

For analogue missions such as NASA's BASALT it has been common practice to have two EV crewmembers cooperating on the field to complete science tasks. More specifically EV1 would be operations lead managing the timeline, traverse navigation and operational tasks, whereas EV2 would be science lead. two IV crew were present inside an IV workstation supporting the EVAs. IV1 was assigned to the communication with EV crew and MSC (via CAPCOM) on operations, procedures, timelines and constraints, while IV2 would interact with EV crew and MSC (via SCICOM) regarding science tasks, recommendations and priorities. The MSC segment would not only involve a CAPCOM and SCICOM but also a Flight Director having the authority over all operational recommendations from MSC, an EVA planner, a Science Team Lead, a Biology Lead, a Geology Lead, an Instrument Lead, an Imagery Lead, a Leaderboard Lead, a Tactical Awareness Lead and a Science Team member [121].

Considering the roles and responsibilities assigned during analogue missions such as BASALT and considering the scope of this study, which is to evaluate the usability of an innovative technology able to support EV crew during surface operations, it is considered sufficient to perform the simulations with one EV crew, as test subject, performing geological field operations and communicating with MSC and one simulated MSC crew playing a key role especially for the comparison scenario in supporting EV crew and collecting reported data. It follows that there will be only one voice loop.

In fact, the scope of this study is neither to develop, test, analyse and validate new exploration architectures as it was for BASALT nor to develop future mission scenarios and increase the relevance of training as it was partially for Pangaea-X, the goal of this study is to evaluate new crew support technology.

Furthermore, it should be noted that the new technology being tested is a proof of concept and to be able to assess the efficiency and effectiveness of it, it is important to keep the complexity of the scenario low, else the more variables are introduced the more bias is introduced and the more difficult the assessment becomes. The decision to keep the scenario as simple as possible to reduce the noise factors during the test led to the choice of only one test subject as EV crew. It should be noted again that the aim of this research is to assess the effectiveness and the efficiency of technology on its own and not the ability of the technology to create and enhance the collaborative environment between the two EV crew. Additionally, despite the fact a collaborating rover able to support crew is a safe assumption to make when considering future lunar exploration mission, it has been decided not to include this component in the tests to reduce the complexity of the simulation and the mental effort for the test subject. It is believed in fact that introducing a rover represents an undesired noise factor.

The considerations mentioned earlier in section 5.1 still apply, namely that the research design has sufficient fidelity with the operational environment, the test subjects are representative of the target population, consent to analysing the results is provided by the latter and that the study integrates all required tasks to address the objectives and a plan to gather quantitative and qualitative data.

The EV crew (test subjects) will be geological field activity experts, hence either instructors or astronauts. While the MCC will be simulated by the same trained person in each test to reduce individual bias.

5.3.4. Test Equipment

For the scenario in which the AR-IoT tool is evaluated, the test equipment consists of the Microsoft (MS) HoloLens [163], a Raspberry Pi [202] and voice communication devices. An available Wi-Fi network is also necessary. The IoT network consists of a Raspberry Pi which receives and manages either simulated data or data from several different types of sensors equipped with a Wi-Fi chip, ESP8266 [64], Zigbee [250] or LoRA [140]. It should be noted that this IoT network is modular such that different types of sensors can send data to the Raspberry Pi. The AR cues are generated using in-house software that was programmed in Unity [227] and are displayed on the MS HoloLens. For more information on the hardware used as well as the hardware and software integration the reader is referred to subsection 3.1.1 and subsection 3.1.2

5.3.5. Data Flow

In Table 5.1 an overview of the data type generated by the simulated Mission Control Centre (MCC) and the simulated extravehicular (EV) crew, together with the way this data is transmitted through the different interfaces at hand is given. Furthermore, it should be noted that latencies and/or interruptions are simulated in voice communications between EV and MCC. Latencies should be introduced in the data transmission as well. Additionally, the latencies and/or interruptions in this scenario are most crucial on voice communications between MCC and EV crew, latencies in other data transmission are not affecting the operations thus the results in this case.

Table 5.1: Data Flow between Mission Control Centre and EV crew through various interfaces. Information on the data type, who generates it, transmission modality and whether latencies/interruptions occur or not is specified.

Data Type	Generated by	Accessible by	Transmission Modality	Communication Issues
Planned traverse	MCC	EV, MCC	Post-its on the ground describe the path	No
Planned activity	MCC	EV, MCC	Briefly outlined beforehand and during the activity through cuff-checklist or HoloLens	No
Notes/Tags (text)	EV	EV, MCC	Spatially locatable, accessible through the Windows Device Portal	No
Waypoints	EV	EV, MCC	Spatially locatable, accessible through the Windows Device Portal	No
Photos	EV	MCC	Spatially locatable, accessible through the Windows Device Portal	No
Video	EV	MCC	Spatially locatable, accessible through the Windows Device Portal	No
Voice Communication	EV, MCC	EV, MCC	Transmitted with delay and simulated interruptions	Yes
IoT sensor data of suit consumable levels and biomedical data	Sensors and/or Script on RaspberryPi simulating sensor data	EV, MCC	Accessible via command line client by the MCC and via HoloLens or device with the Graphical User Interface	No

5.3.6. Task Instructions

In subsection 5.3.2 an overview on the task design was given. More detailed information on the actual tasks are given hereafter. Task instructions given to the test subject are limited. At the start of each test, participants are told that they have to perform tasks related to geological site inspection activities for which they will be given sufficient instructions to be followed. More specifically, the test subjects are requested to:

"perform the tasks as accurate and as fast as possible while maintaining adequate situational awareness and following the instructions given by the tool at hand".

When using the MS HoloLens, the user is first familiarized with the tool, especially with the user hands-free interaction possibilities. A quick interaction trial by gaze and click is performed, moreover the user is made aware of the fact that the device provides with speech recognition.

After that, each participant is requested to locate themselves at the airlock and perform system checks in communication with ground control who will provide with all the required instructions to accomplish this task.

In fact, from Figure 5.1 it can be seen that the first task that crew has to perform is to carry out system checks. In this case, crew has to report to simulated ground control also referred to as Mission Control Centre (MCC) whether voice communication is working nominally and in case the HoloLens is used, whether the HoloLens application is functioning nominally. These system checks do not require any knowledge by the participant as all instructions will come from the simulated ground control. Moreover, in both cases crew is requested to report simulated suit data consisting of biomedical data, namely suit pressure, heart rate, temperature and consumable levels more specifically water, oxygen and carbon dioxide. System checks performed with MCC represent a pre-task. It means that the time taken to accomplish this task and the accuracy in carrying out this pre-task are not considered in the results analysis.

Upon system checks completion, the test subject is instructed to egress the airlock and start the outlined traverse by means of post-its on the ground until the marked target area is reached where the test subject is requested to perform all the requested site evaluation activities. All tasks should be performed with the sole use of the tools provided without requiring communication with MCC (in the case the AR-IoT tool is evaluated). Special attention should be taken during the traverse and interesting sites, so called targets of opportunity (TOP) should be "logged" or reported using the HoloLens.

It should be noted that the path outlined by the post-its will not have been visible to the test subject prior to that moment.

Once the airlock is egressed the stopwatch is activated. The participant will soon reach the first TOP, quite evidently marked by a certain number of rocks placed in the middle of the path where he or she should be logging data with the HoloLens as accurately as possible following the instructions on the HoloLens regarding the features to document. The test subject will know which option to click on the HoloLens to log the data of a target of opportunity (TOP) as it is named accordingly. The participant will have to report on:

- Number of rocks

- Colour of the rocks
- Other notable characteristics of the surrounding area

After that, the test subject will continue the traverse through the outlined path to the target area where he or she has earlier been requested to perform a site evaluation activity. The participant will know which tasks to perform as they are specified under the option "site evaluation" on the HoloLens. Upon arrival at the marked target area the participant will have to create a waypoint by placing a virtual flag in the augmented world and document the area based on the specified criteria, which are the same as for the TOP evaluation. After that, the subject is requested to take a regular photo and a panoramic 360 photo with the HoloLens.

The stopwatch will be stopped once the test subject will signal that he or she is done with the site evaluation task.

The idea behind the given task was to keep them as simple as possible to reduce the workload and task complexity for users that are completely new to augmented reality technologies.

It should be noted that the role of the IoT in this test scenario is among others marked by the presence of a predetermined traverse map, in fact in this scenario it is assumed that a lunar or planetary rover, which is part of the IoT network, has previously mapped the area and the data gathered was then used to create a map of the area to determine interesting sites or target areas and the associated traverse. Moreover, the sensors, supposedly embedded in the astronaut suit and the life support system which provide with simulated data, are also considered part of the internet-of-things network. The data obtained and displayed to the participant during the test is not meant to influence the results, it is however meant to enhance realism and showcase a useful and expandable feature for later geological field activities to the experts. In fact, during Caves & Pangaea campaigns carry-on "sensor boxes" containing different environmental sensors, required for instance to explore caves, were employed. Integrating sensor data from these sensors into the AR-IoT tool is considered useful and straightforward to implement.

5.3.7. Guidelines

For crewmembers to know which kind of information is required to be collected during the different phases of the EVA, guidelines are usually established by the science team prior to the expedition. During BASALT analogues verbal descriptive observations were required from crewmembers over the course of the EVA [14] such as: weather conditions, general observations regarding the outcrop scale, environmental discomfort, potential TOP, other unforeseen conditions and/or obstacles as well as detailed observations regarding the outcrop and sample scale [12].

The procedures created for the operational scenario of this research are inspired by the cuff-checklists created for the BASALT analogues [12] used by the EV crew as prompts for what descriptions they should be reporting at different stages in the EVA (see Figure 5.2).

MISSION PHASE	TASKS	TASK DETAILS
Station Approach (15 min, once into comm coverage)	• Translate to station perimeter base • Give local report • Record contextual video + still imagery • Look for TOPs • Arrive at station perimeter base	• Local Report: o Local time o Wind speed & direction o % cloud cover & sun ∠ o temp & precipitation
Station Contextual Survey (5 min, 5 m out)	• Position SA Camera • Give contextual report, including still imagery and video (as bandwidth allows)	• Contextual Report: o Orientation & shape o Condition & color o Presence of water, other fluids, & biomass o Specifics for today's EVA objectives
Sample Location Search (25 min, 1 m out)	• Get candidate sample location markers from FST • Search for, mark, describe, & record video + still imagery (as bandwidth allows) for each candidate location • Translate back to station perimeter base	• Sample Location Report: o Orientation & shape o Condition & color o Presence of water, other fluids, & biomass o Specifics for today's EVA objectives
Pre-Sampling Survey (60 min, < 1 m)	• Get candidate sample markers from FST • Translate to Sample Location • Search for, mark, describe, & record video + still imagery (as bandwidth allows) for each candidate replicate • Take ASD scans • Take still imagery of ASD screens • Translate back to station perimeter base	• Replicate Report: o Orientation & shape o Condition & color o Presence of water, other fluids, & biomass • Report: o ASD minerals + general spectrum
Sampling (30 min, < 1 m)	• Perform bio sterilization • Translate to replicate location • Break rock • Hold samples + bags up to video camera (if bandwidth allows) • Bag samples • Arrange bags on original outcrop and take still imagery of bags	• Sample Report: o Bag number o Sample type (bio, mcmaster, geo, archive) o Color o Surface slope & azimuth o Friability o Texture, vesiculation o Volatiles o Primary, Secondary Minerals o Presence of biomass

Figure 5.2: Cuff-checklist providing EV crew information on what is required for them to so at each stage of the EVA. TOPs stands for Targets of Opportunity, SA for Situational Awareness, FST for Field Support Team, ASD for Handheld VNIR spectrometer [12].

Interestingly, scientists and engineers of different disciplines frequently use unique language and distinct terminology, hence it became crucial to implement formal project definitions for terms that would be used to plan and execute the BASALT EVAs. The set of definitions comprised references to spatial locations, employed in the formulation of EVA tasks, geological/environmental terms used to describe the scientific objectives of each individual EVA and the alteration types being selected [14]. Stevens et al. [12] reported a brief glossary cheat sheet (see Figure 5.3).

FIELD GLOSSARY CHEAT SHEET	
FEATURE	**DETAILS**
Color (includes variations in color)	• black, grey, tan, reddish-brown, orange • banded or zoned
Grain Size	Approximate size ranges as follows: • very coarse 5-10 mm • coarse 2-5 mm • medium 1-2 mm • fine <1mm • aphanitic -- not visible, too fine to see unaided.
Alteration	Degree of alteration or weathering is based on two factors: 1. **Thickness** of alteration rind. a. none = fresh, cannot measure b. low = <5 mm c. medium = 1-2 cm (can find fresh rock inside) d. high = 2-10 cm (have to work hard to get fresh rock) e. intense = > 10 cm, or even pervasive (generally not applicable to young rocks in dry climates; however, some rocks as young as a few 10 ka might have infillings of caliche to depths exceeding "medium" or "high") 2. **Intensity** of mineralogical changes from original rock texture, regardless of the rind thickness. a. none = cannot depict any change from fresh rock b. low = original visible minerals are present, but they may be discolored; groundmass is discolored. c. medium = both visible minerals and groundmass are discolored, some degradation of original fabric (groundmass texture). d. high = cannot determine original mineralogy or groundmass; original textures completely lost; discoloration complete. e. intense = cannot distinguish whether altered part is actually original rock or has been replaced; totally discolored and disfigured.

Figure 5.3: Field Glossary Cheat Sheet [12].

For the selected concept of operation the following dummy observational data has to be collected.

- Number of rocks
- Colour of the rocks
- Other notable characteristics of the surrounding area

The following guidelines in form of cuff-checklist, shown in Figure 5.4, are given during the test scenario employing traditional media. The HoloLens application avails itself of the same instructions under the site evaluation option. The same instructions of data logging of the target of opportunity are incorporated under the "log data" option in the augmented reality interface. It should be noted that despite the fact the naming "log data" and "data logging of the target of opportunity" are not consistent yet, this will have to be ensured before the test. While the possibility of integrating a cheat sheet, as shown earlier in Figure 5.3, should be kept in mind, in this case it was not considered necessary as the tasks requested to be performed are consciously simplified to avoid the need of prior training or briefings of inexperienced users.

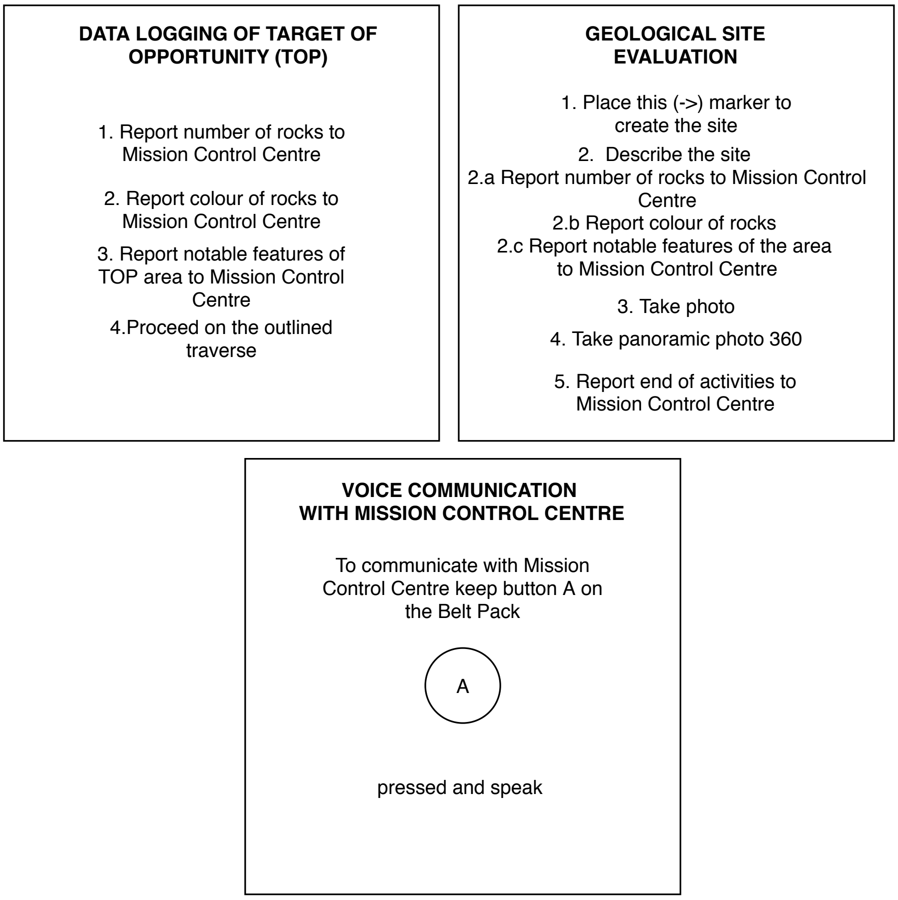


Figure 5.4: Cuff-checklists provided as instructions for the TOP evaluation and the target area evaluation during the controlled experiment with conventional media.

6 Results

In this chapter, the results from the different evaluations performed through the design of the AR-IoT tool are discussed. First, the results from the informal expert reviews of the initial design concepts are reported. These were performed during the conceptual design phase. After that, the results from the heuristic evaluation are discussed, this was performed once the initial concept of the application had been reviewed several times and an actual final concept was conceived with a more concrete structure and layout, more specifically concept IV. Two further evaluations were performed through a requirements compliance questionnaire (see Appendix C) and complementary semi-structured interviews directed to ESA astronauts and astronaut geological field training experts. The results from both the questionnaires and the in-depth interviews are reported in this chapter as well.

6.1. Expert Reviews Results

During the conceptual phase, the primary design phase of the AR application, expert reviews were performed. These were informal open discussions on the capabilities and means of user interaction the application should have to become operationally feasible and to be helpful and useful for the target group e.g. astronauts. Experts had the opportunity to mention positive and negative aspects as well as to suggest further features to be implemented.

Concept I was reviewed by three experts at the European Space Agency's Astronaut Centre (ESA-EAC), these experts included a former participant of the Pangaea-X 2018 test campaign, a software developer of the Caves & Pangaea team and an expert in human and robotic exploration. The results are reported in Table 6.1. Generally, the initial concept, including the capabilities of the interface, was defined as good and clear as well as the gaze interaction with the gaze feedback through the horizontal loading bar on the button.

With respect to the layout it was suggested to replace the rectangular shaped buttons with large square buttons. Regarding the structure of the menus, it was suggested to group features and arrange these in different levels depending on whether the actions have to be performed simultaneously or sequentially.

With respect to the interface capabilities, it was suggested to include the display of system date and time as well as the request for registering the user's ID upon login. It was also suggested to enable the personalisation of the interface depending on the user's needs. Moreover, regarding the tool localisation features it was suggested to direct the user with augmented visual feedback to the tools.

With respect to the user interaction, the gaze input with the gaze feedback, through the horizontal loading bar on the button, was perceived as effective and clear. Nevertheless, it was mentioned that a double confirmation should be requested to the user before proceeding with an action to give the user more control over their actions. More specifically, instead of executing an action after loading completion of the button, which could happen accidentally as the user is inactively gazing on a button, it is important to request for an additional action through a different or similar means of interaction to confirm the request. The response time of the interface more specifically the loading speed of the loading bar was perceived as too slow. Interestingly, as some features were fully interactive through gaze while others through the use of hand gestures, as for instance the "place tag" button, it could be observed, from the interface interaction by one of the experts, that hand gesture interaction is very troublesome for inexperienced users, in fact the subject was not able to successfully complete the action.

Concept II was reviewed by two experts in human machine interaction (HMI) interface design. The results are reported in Table 6.1 as well.

Positive and negative aspects related to the interface as well as suggestions for further implementation are reported hereafter.

Generally, the interface was defined as well structured and the capabilities as clear, the colour scheme was defined as good and clear as well as the icons and labels. It was noted however, that the menu layout was occasionally lacking consistency. Additionally, it was noted that the flag created once the "create stop" button was pressed was not appearing in the field-of-view (FOV) of the user.

Suggestions regarding the interface included: placing the "pin" button next to the "quit" button for consistency as these actions are both involving the whole menu, moreover it was suggested to provide additional feedback when the pin window capability is enabled by highlighting the borders of the menu. Other suggestions included displaying all menus in such a way that they are all in the FOV of the user, angled such that the setup resembles a "semi-circle", highlighting buttons when active and displaying the label of an icon only upon hovering with the gaze to reduce the mental overload of the user.

Positive and negative aspects related to the means of interaction as well as suggestions for further implementation are reported hereafter. The double confirmation through a combination of user inputs, namely gaze and click was perceived as intuitive, user-friendly and effective. It was noted however that the "click to confirm" prompt to remind the user to use the clicker to confirm an action upon gaze was displayed too soon. Furthermore, it was mentioned that the response time of the loading bar was yet too slow and loading was not resumed upon gazing away from the button. Finally, it was pointed out that unfortunately no feedback was provided to the user once a photo had been taken successfully.

Concept III was reviewed by the same experts that reviewed concept I. The results are reported in Table 6.1.

With respect to the interface the experts mentioned that it was clear and comprehensive, especially the icons and labels of the buttons. It was noted that all the required features for geological site evaluation activities were included. Nevertheless, experts mentioned that prompts were not always legible and the flags created upon selection of the "create stop" button were perceived as too small. Finally, a negative aspect mentioned was that the information of the stop documentation was not linked to the flag.

Suggestions for further implementation included the definition of an architecture for the different menus as well as the direct inclusion of the capabilities related to each step in the references to ease the user and reduce mental workload. It was also suggested to include a logbook in the interface that allows the user to track concluded actions. Additionally, it was suggested to include a "delete" option to increase the control of the user, namely allowing the user to delete previously recorded data, if desired. Finally, it was suggested to perform colour testing to evaluate whether good visibility of the interface is always ensured under different lighting conditions e.g., indoors and outdoors.

Regarding the user interaction, it was mentioned again that loading should resume when the user does not gaze on a button, moreover it was noted that it was not possible to take photos at a desired location as the "geolocation site screening" menu was set as static and would not follow the user. Finally, it was mentioned that the type of interaction, namely gazing on the object, click to select, drag and drop to place an object at a desired location should apply to not only the placement of the flag but also to the "pinning" or dragging and dropping of a menu.

Table 6.1: Informative expert reviews on initial three concepts of the application. Positive and negative aspects of the interface are reported as well as suggested features for further implementation. The reviewers are labeled as REV followed by a number.

Concept	Positive Aspects	Negative Aspects	Suggested Features
I	[REV1] idea is good and solid [REV2] Gaze feedback through loading bar is good and effective	[REV3] Use of gestures is troublesome [REV2] Missing request for confirmation to proceed with the action once the loading bar has completely loaded [REV2] Response time of the loading bar should be faster	[REV1] get system date and time upon log-in and EV ID [REV1] personalisation of GUI for each user [REV3] Use big square buttons instead of bars [REV3] Tool localisation capability with highlighted edges to direct the user [REV2] Add double confirmation through final additional loading bar or other means of interaction [REV2] Create an application architecture to define the different levels depending on the actions that have to be performed simultaneously or sequentially
II	[REV4] Good to have icons and labels [REV4] Well structured [REV4] Capabilities are clear [REV4] Good use of colours [REV4] Gaze + click is effective	[REV5] Waypoint flag is not in the field-of-view [REV5] 'click to confirm' prompt appears too soon [REV5, REV4] Resume loading is not enabled when user does not gaze [REV4] Missing consistency in layout [REV4] Response time for the loading bar should be faster [REV4] Missing feedback upon picture taken	[REV5] pin tag button should be next to the quit button [REV5] illuminate borders of the window as feedback as pinning function is enabled [REV5] show label complementing the icon on hover [REV4] when a button is active highlight it [REV4] arrange the windows cinema-screen like, all around, angled. And when one is active bring it in the foreground
III	[REV3] Clear interface, clear icons [REV3] Speech-to-text via speech recognition [REV3] Includes required features [LT] Good interface	[REV3, REV2, REV1] Resume loading is not enabled when user does not gaze [REV1] Unable to take the photo at desired location [REV1] Prompt text not always readable [REV2] Flag is too small [REV2] Information is not attached to the flag	[REV3] Menu arrangement tag along, highlight, bring forward [REV3] Link stop documentation to the waypoint/flag [REV3] Link capabilities directly to the steps in the references [REV3] Add logbook to the steps accomplished by the user [REV1] Use same interaction as for the flag placement for the pin tag [REV1] Erase field with transcribed text option [REV1] Perform colour testing

6.2. Heuristic Evaluation Results

After the conceptual phase which included several informal expert reviews, once a more complete, defined and structured concept was available, namely concept IV, three heuristic evaluations were performed. The evaluations were performed by three different Human Machine Interaction (HMI) design experts separately. Each evaluator was given a checklist with the identified usability principles, the so called "heuristics" and was requested to examine the interface and judge its compliance to these heuristics. Each heuristic evaluation took approximately one hour. For more information on the methodology behind this evaluation the reader is referred to subsection 3.2.3.

The rating of and the comments provided by the different evaluators are reported in Appendix B. It can be seen that for some of the listed and rated principles there is quite a spread in the assignment of the points on the 5-point Likert scale [232]. For principle 1, 4, 5, 8, 10 and 13, the ratings differed majorly between the subjects as can be seen when comparing the filled in questionnaires of the different subjects in Appendix B. A key feature of the heuristic evaluation questionnaire ratings process is the fact that comments and recommendations were documented for each numerical rating provided. While the numerical ratings are important and often the subject of considerable discussion and debate, their most important function in this case was to provide structure and consistency to the process through which the different heuristic principles identified were critiqued.

The comments provided during the heuristic evaluation were gathered and discussed during a de-briefing session together with all three evaluators. The feedback for each point is discussed hereafter. Please consult subsection 3.2.3 or Appendix B to see the details on each heuristic principle selected.

1. **System status feedback**
   is unclear with respect to:

   - recording status
   - saving status
   - take pan 360
   - dragging & dropping of the flag
   - button states
   - take photo

   is redundant with respect to the:

   - confirmation request "click to confirm"

2. **Response times** are inadequate when using the loading animation for the gaze feedback
3. Unclear/unfamiliar **icons** include:
   - Save icon which resembles a download icon instead
   - Reference icon which shows an inappropriate label
4. The user's **freedom to control** is not ensured during the "take photo" action and when going through the references in fact the user could not go through them backwards
5. The interface is rated as not fully **consistent** due to the loading animation which does not resume when the user gazes away and due to the references which do not portray all the steps in detail
6. The interface is rated as **legible**
7. **Font size** should be increased for certain prompts such as the "photo taken" prompt
8. The **prompt** for photo taken is unclear, while prompts for: recording, pan 360 and for the confirmation to delete the text are missing
9. **Color and brightness** are good

10. **Instruction** for required "drag & drop" action is missing
11. **Relevant information** is present
12. Displays and control are **visible and easily accessible**
13. The created flag and the references menu are not **within the field-of-view**
14. Information such as the written prompt "click to confirm" to confirm the action is excessive while **insufficient** for the flag "drag & drop" task

The discussion with all the evaluators aimed at finding solutions to the identified issues which led to a violation of the selected heuristic principles. For each principle certain questions had to be answered.

1. How to implement more intuitive feedback for the record, take pan 360, take photo, move flag and save button?
Evaluator 1: suggests the use of sounds and colours or other means of providing feedback.
Evaluator 2: is not in favour of mixing means of feedback, stressing the importance of staying consistent, while being in favour of using text and keeping in mind to not overwhelm the user.
Evaluator 3: suggests setting up a strategy to have consistent icons and means of interaction based on the various states of the buttons.

From the discussion on principle **1.** it became clear that the evaluators have contrasting opinions on how to best provide feedback to the user. An agreement on a strategy on how to make the recording button states more intuitive could not be found. Nevertheless, everyone agrees on displaying a "drag & drop" label on the flag.

2. How to solve the inadequate response time issue when using the loading animation for the gaze feedback?
Everyone agrees to get rid of the loading animation and to change the colour of the button, by highlighting it on gaze/hover instead.

3. How to make the save, reference and delete icons more intuitive?
Everyone agrees to use a trashcan icon for the "delete" button and folder with a horizontal arrow pointing into the folder from the left-hand side for the "save" button. Additionally, everyone agrees that "tutorial" instead of "references" should be used as label and for consistency it is suggested to have a tutorial function to perform every demanding action that includes several steps.

4. How to provide the user with more freedom to control?
Everyone agrees to add a "go backwards" button to the references/tutorial button and to allow the "geological site screening" menu to follow the user such that those actions can be performed at any location.

5. How to ensure consistency?
Evaluator 1: suggests to add more steps to the references/tutorial.
Evaluator 2 and 3: suggest to not add more steps, by stating that the steps are clear and detailed enough, in fact it is safe to assume that the target user (astronaut) does not require additional steps.

From the discussion on principle **5.** no conclusions can be drawn for now. Usability tests and controlled user experiments are foreseen to provide more insights regarding this matter.

8. How to make the prompts brief and less ambiguous?
Everyone agrees that there is a need for a consistent plan for prompts. However, as for principle **1.** no agreement can be found on how to submit feedback to the user. More specifically, it remains unclear whether it is actually best to use text to ensure consistency and clearness or not to use text as it may cause an increased cognitive workload.

Everyone agrees that a prompt requesting a deletion confirmation should be implemented.

13. How to ensure that everything is within the FOV of the user?
The FOV is limited by the technology.

14. How to keep information relevant, sufficient but not excessive?
Evaluator 1: suggests to use different colours for the flag to indicate the state of completion of the "geolocation site screening" action e.g. red if incomplete and green upon completion. The evaluator in this case was not entirely familiar with the human spaceflight colour coding scheme in fact it is very important to pay attention to the colours used for interfaces that are meant for astronauts as there is a clear set of colours and a clear meaning associated to them (yellow: caution, red: warning or emergency, green: normal state) [173], [181].

From the discussion it became clear that an agreement on a solution could not always be found due to contrasting opinions on design principles and/or methodologies as well as personal preferences. Nevertheless, the discussion provided with certain useful and implementable solutions. It should be noted that these solutions were implemented right away as can be seen in prototype I (see section 3.2). Additionally, some good insights from the different designers' opinions were obtained. It became evident as well that a clear, well-defined and structured operational scenario/concept of operations is needed to be able to provide the application with a more structured interface.

6.3. Astronaut and Astronaut Geological Field Training Experts Questionnaire Results

In this section, the results from the questionnaires are reported. In total, seven of the experts that have been contacted provided with the filled in questionnaire, more specifically:

- 3 ESA astronauts
- 1 participant of Pangaea analogue missions as well as support engineer
- 3 Caves & Pangaea support engineers and tool developers

Hereafter, the frequencies and percentages for each response are described together with the central tendency (Median, Mdn) and the variation (Interquartile range, IQR). The results were produced using a custom-made Python script and successfully validated using the statistical analysis software, SPSS [87]. It should be noted that all results except the IQR results could be validated. This is due to the fact that the methods used to calculate the percentiles differ, linear interpolation was used in Python while SPSS uses a weighted average method which is not available through the Python NumPy library used to describe the other results. Therefore, the results of the IQR computed by the Python script have been validated with hand calculations.

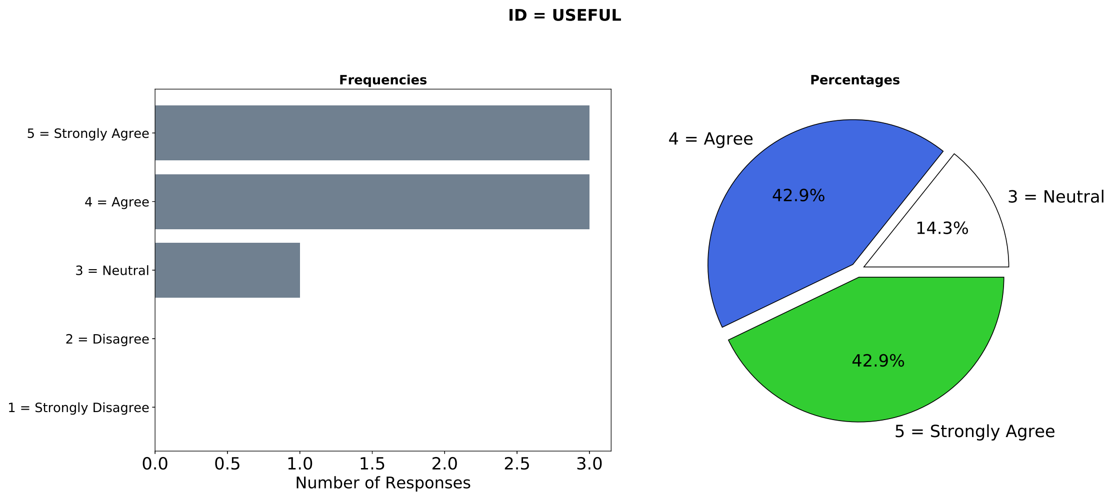
The comments added to questionnaire items were also explored, counted and elucidated. Here it should be noted that comments that said: *"I cannot evaluate"* or *"I can't judge this"* were neither counted nor reported as they are reflected in the respondents' neutral response or missing response on the Likert scale of the questionnaire, furthermore this choice was made to avoid repetitiveness.

In Figure 6.1 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: USEFUL, namely:

> *"The application is useful and helpful for future astronaut geology field training and geology field exploration activities on the Moon."*

Most respondents indicated agreement with the fact that the application is useful and helpful for future astronaut geological field activities (Mdn=4, IQR=1).

Three out of seven added a comment to this questionnaire item, of those three, two mentioned that the application is *"potentially helpful"* (R01) as at the moment there is only *"a short demo of a small subset of all the actions required, which is promising, but would need further assessment once the complete worksite operations are covered"* (R05), one of the two explained that: *"the helpfulness will also depends on the interface. Action confirmation by voice recognition shall be considered. To say sometimes "Confirmed", to validate an action, is not more interfering with the voice loop, than voicing up the site description during the recording"* (R05). The remaining respondent instead makes a comparison with the existing Electronic Field Book (EFB) on a tablet and says: *"it would be much easier and faster to use your application"* (R04).



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	5.0	1.0	0.0

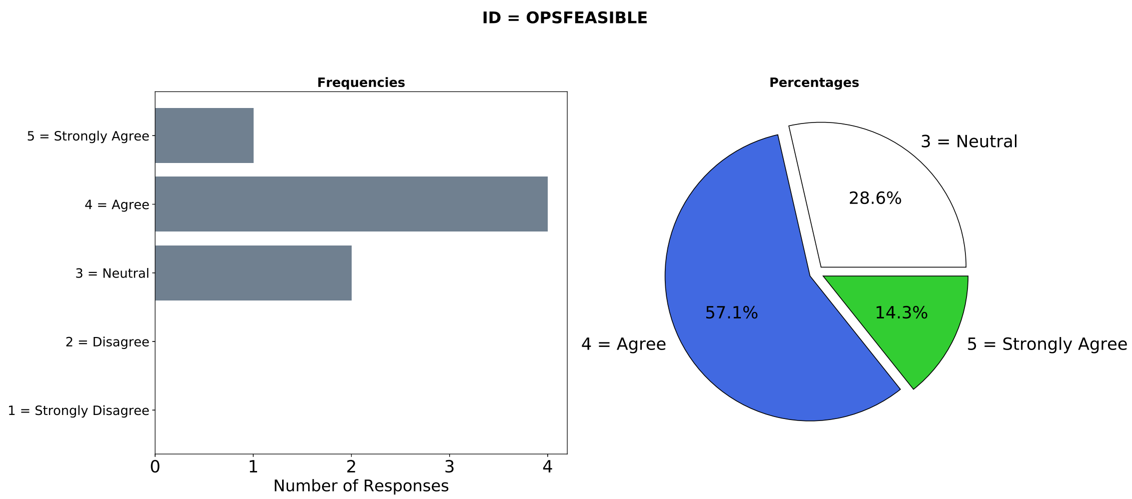
Figure 6.1: Frequencies and percentages for the response to ID: USEFUL, N = 7, missing response(s) = 0.

In Figure 6.2 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: OPSFEASIBLE, namely:

> *"The application is operationally feasible for future astronaut geology field training and geology field exploration activities on the Moon."*

Most respondents indicated agreement with the fact that the application is operationally feasible for future astronaut geological field activities (Mdn=4, IQR=0.5).

Four out of seven respondents provided with a comment, of those four, three agreed that the operational feasibility depends on *"the integration of the Hololens AR into the EVA suit helmet"* (R05). One of the three indicated that: *"it also depends on how reliable it is and how easy it is to operate from inside the suit"* (R05). The remaining respondent argues that it depends on whether the application will incorporate the capability of mapping the surrounding environment, saying that it: *"depends on the open field mapping"* (R03).

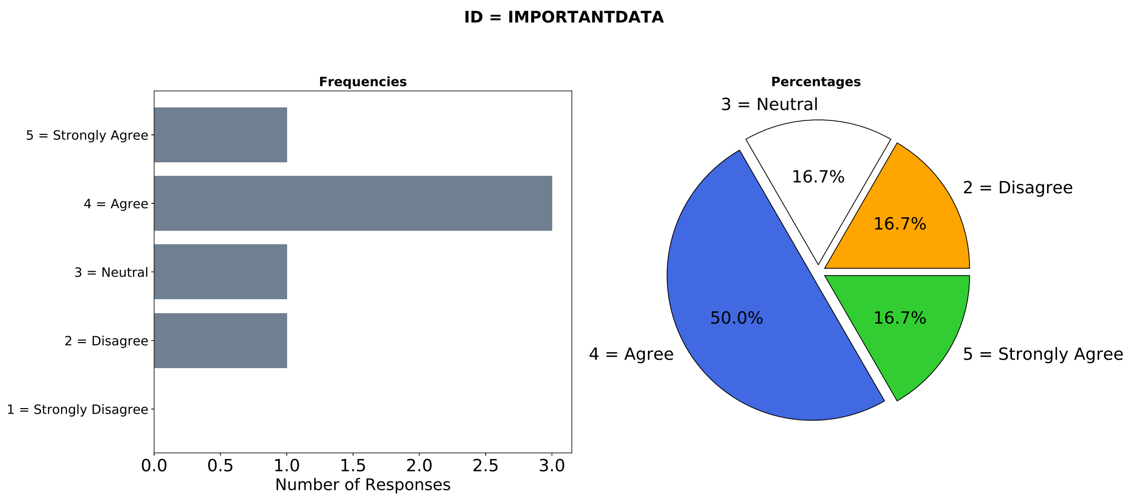


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	3.5	4.0	0.5	0.0

Figure 6.2: Frequencies and percentages for the response to ID: OPSFEASIBLE, N = 7, missing response(s) = 0.

In Figure 6.3 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: IMPORTANTDATA, namely:

> *"Only data that is important to mission success and significant in terms of crew interface is provided."*



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	4.0	5.0	3.25	4.0	0.75	1.0

Figure 6.3: Frequencies and percentages for the response to ID: IMPORTANTDATA, N = 7, missing response(s) = 1.

Most respondents indicated agreement with the fact that the application includes only data that is crucial for future astronaut geological field activities (Mdn=4, IQR=0.75). However, one response is missing.

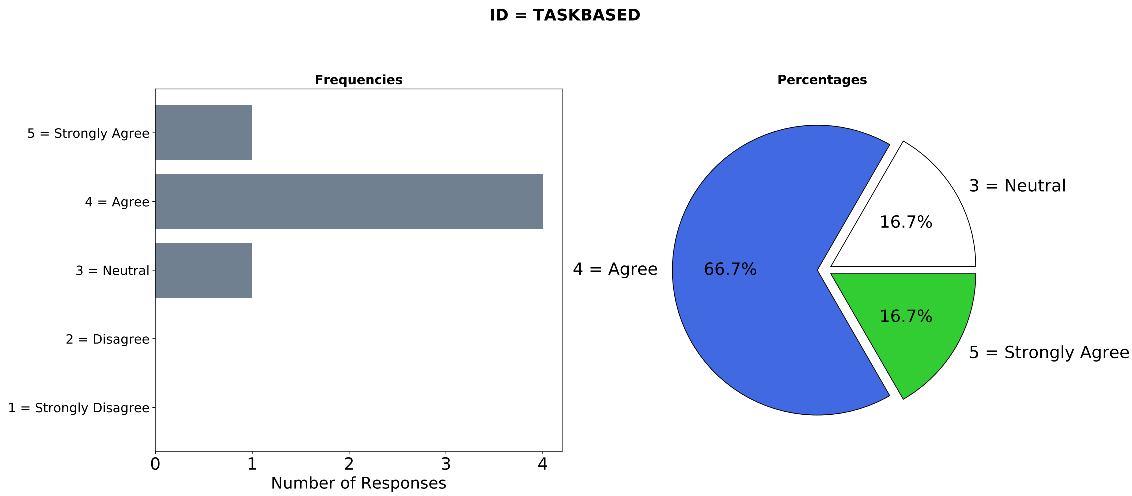
Four out of seven added comments, three of them agree that only *"a subset of"* (R03) the important data is provided as *"access to a map of the area, Augmented Reality identifications on horizon objects"* (R05) as well as the *"capability to recall photos or visualise them before recording them, distances appearing in AR between different locations, real calculated time allowance to reach a safe haven (based on suit consumable left, consumables consumption, distance, calculated consumables consumption required to reach the safe haven, AR flag of the position of the EVA buddy in real time (if hidden by an obstacle or far away distance), sample locations in AR, virtual AR sample markers"* (R05) are still missing. Moreover, two of these seven also agree that certain *"info displayed is still too much"* (R02) one of them explains that the longitude and latitude data should be coming from another device and says that these data should be *"directly transmitted and recorded automatically"* (R05), and that this should be done *"instead of reading the data, and entering it in the AR field"* (R05).

In Figure 6.4 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: TASKBASED, namely:

> *"The overall display design is based on the geological site inspection tasks that will be performed with the display."*

Most respondents indicated agreement with the fact that the application is based on geological site inspection tasks (Mdn=4, IQR=0). However, one response is missing.

Three out of seven left a comment regarding this, two of them agree that the application is a *"good start"* (R05) and that it is *"on a good track"* (R02), all three agree that however *"a lot of key steps are not covered yet"* (R05) and suggest to add places to record *"site lithology"* (R06) and *"structural relationships"* (R06) as well as *"validation of rejection feedback from the Control Centre to the astronaut"* (R05) such that the astronaut can keep track of which samples have been accepted, rejected and which ones have been newly selected. Moreover *"the confirmation by the astronaut of the taken sample locations, IDs and descriptions"* (R05) is also a task that is still missing.

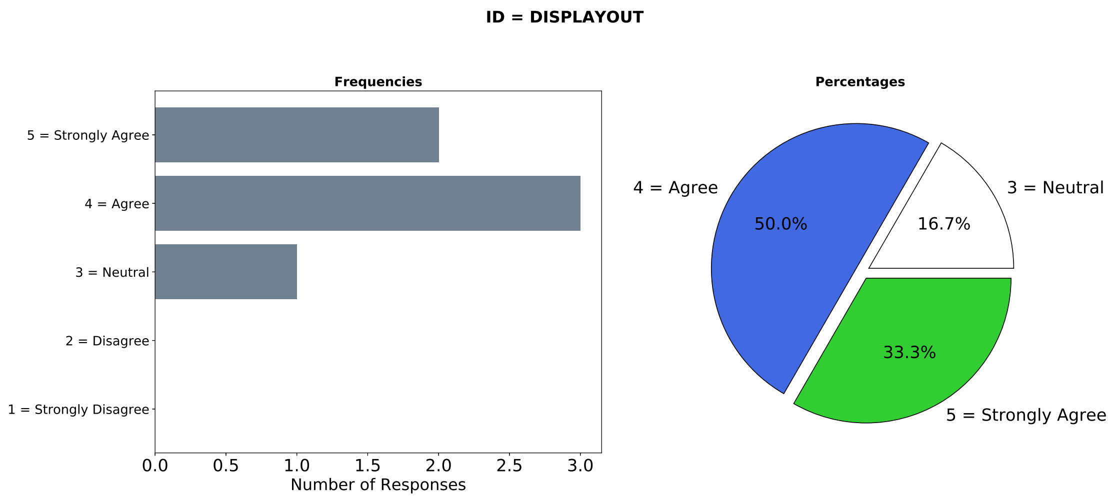


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	4.0	0.0	1.0

Figure 6.4: Frequencies and percentages for the response to ID: TASKBASED, N = 7, missing response(s) = 1.

In Figure 6.5 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: DISPLAYOUT, namely:

> *"Specific data shown, the display layout and groupings, and the choice of display elements is driven by operational requirements."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	4.75	0.75	1.0

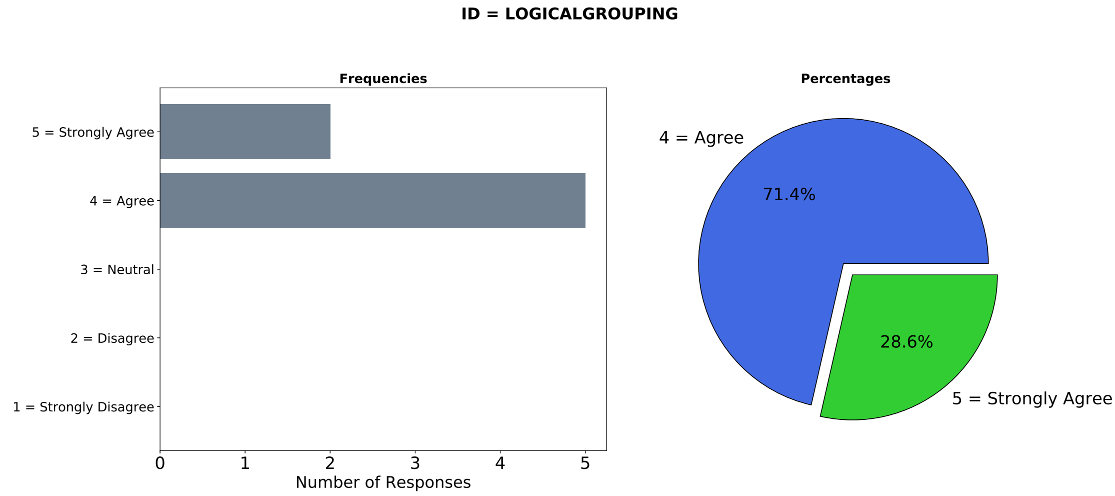
Figure 6.5: Frequencies and percentages for the response to ID: DISPLAYOUT, N = 7, missing response(s) = 1.

Most respondents indicated agreement with the fact that the application is driven by operational requirements when it comes to the display layout and groupings, and the choice of display elements (Mdn=4, IQR=0.75). One response is however missing.

Three out of seven provided with comments and all three addressed different aspects: one perceived that the *"groupings and file are sensible"* (R06) another one mentioned that *"too many windows open at the same time"* (R01) while another respondent suggested that *"operational requirements need first to be clearly identified and cross checked with the layout"* (R05).

In Figure 6.6 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: LOGICALGROUPING, namely:

"Information is logically grouped according to purpose, function, or sequence of use (e.g., either a left-to-right or top-to-bottom orientation)."



count	min	median	max	25%	75%	IQR	missing values
7.0	4.0	4.0	5.0	4.0	4.5	0.5	0.0

Figure 6.6: Frequencies and percentages for the response to ID: LOGICALGROUPING, N = 7, missing response(s) = 0.

Most respondents indicated agreement with the fact that the information within the application is logically grouped (Mdn=4, IQR=0.5).

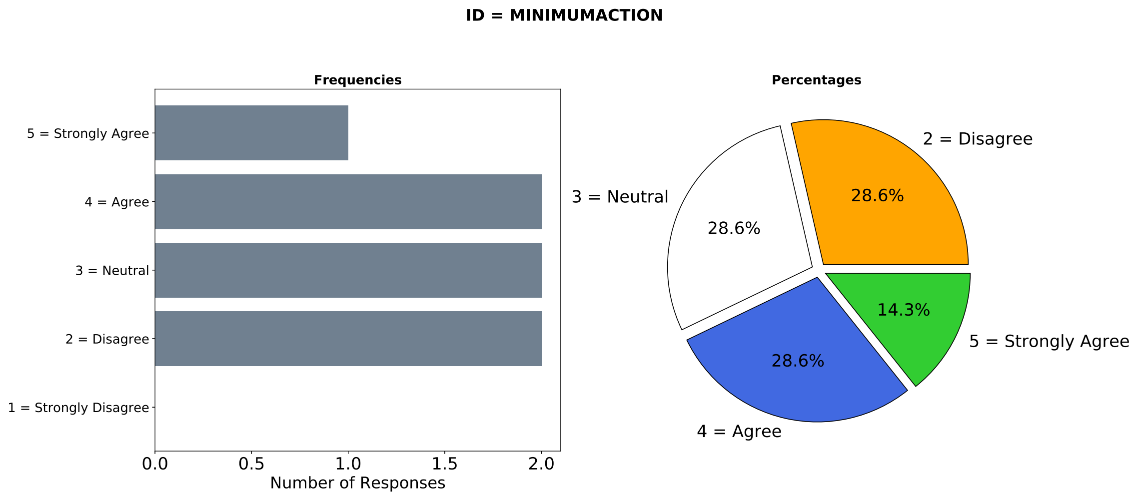
Two respondents out of seven left a comment, both agree that the grouping is *"logical for these first functions"* (R05) one of them adds that it is *"intuitive that information opens from top to bottom and left to right, like a computer GUI system"* (R06).

In Figure 6.7 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: MINIMUMACTION, namely:

> *"The display follows operational flows and allows task completion with the minimum number of actions."*

Opinions seem to be divided with regard to whether the application allows task completion with the minimum number of actions. While 28.6% (N=2) of the respondents remained neutral, 28.6% (N=2) expressed disagreement, but a roughly equal number (42.9%, N=3) indicated that they agreed or strongly agreed (Mdn=3, IQR=1.5).

Three out of seven left a comment, while one respondent believes that the number of actions can be *"minimised"* (R05) suggesting to use voice instead of gazing and/or clicking a button, explaining: *"saying "confirmed" to confirm an action is [a] less demanding action than pressing a button or glazing [gazing] carefully at a precise location on an AR display for 2 to 3 seconds"* (R05) another one believes that *"specific tasks could be simplified (e.g. planting the flag could be also a single button, allowing to have the flag in the ground in front of you automatically in one click, and with inside automatically geolocation coordinates)"* (R03). The third respondent expressed concerns about the fact that the *"windows are fixed in space and one has to move the gaze away from what one is doing to go back to watching the menu windows"* (R01).

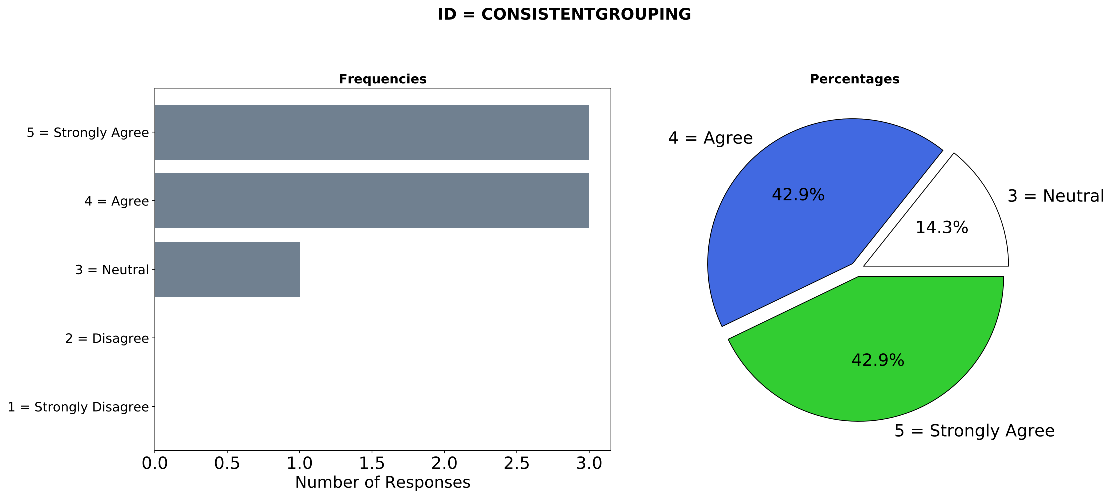


count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	3.0	5.0	2.5	4.0	1.5	0.0

Figure 6.7: Frequencies and percentages for the response to ID: MINIMUMACTION, N = 7, missing response(s) = 0.

In Figure 6.8 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: CONSISTENTGROUPING, namely:

> *"The display is consistent when grouping/ordering display elements."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	5.0	1.0	0.0

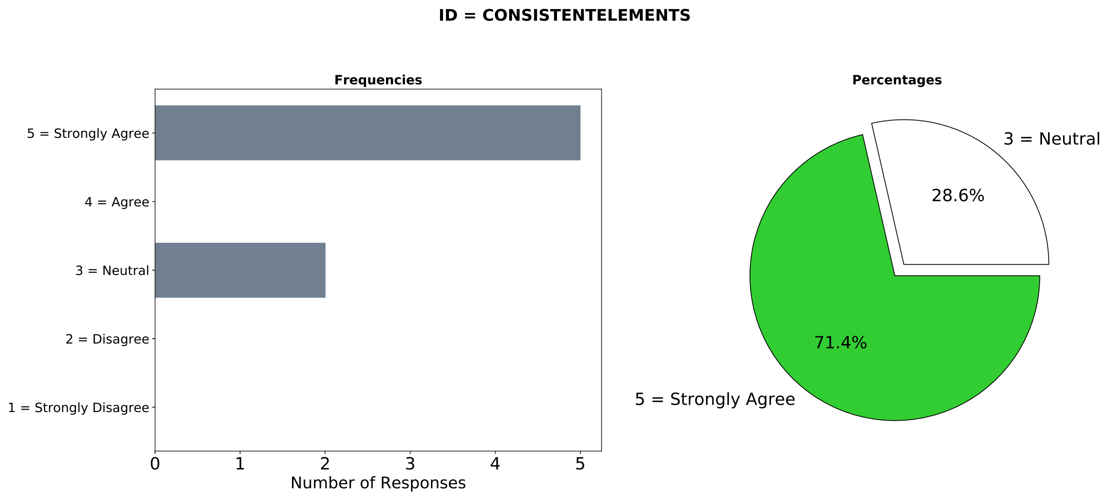
Figure 6.8: Frequencies and percentages for the response to ID: CONSISTENTGROUPING, N = 7, missing response(s) = 0.

Most respondents indicated agreement with the fact that the application shows consistency when it comes to grouping/ordering display elements (Mdn=4, IQR=1).

Only one out of seven left a comment agreeing with the statement and adding that *"the hierarchy tree is clear and consistent"* (R06).

In Figure 6.9 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: CONSISTENTELEMENTS, namely:

> *"The interface elements, colours and provided feedback are consistent."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	5.0	5.0	4.0	5.0	1.0	0.0

Figure 6.9: Frequencies and percentages for the response to ID: CONSISTENTELEMENTS, N = 7, missing response(s) = 0.

Most respondents indicated strong agreement with the fact that the application interface elements, colours and feedback given are consistent (Mdn=5, IQR=1).

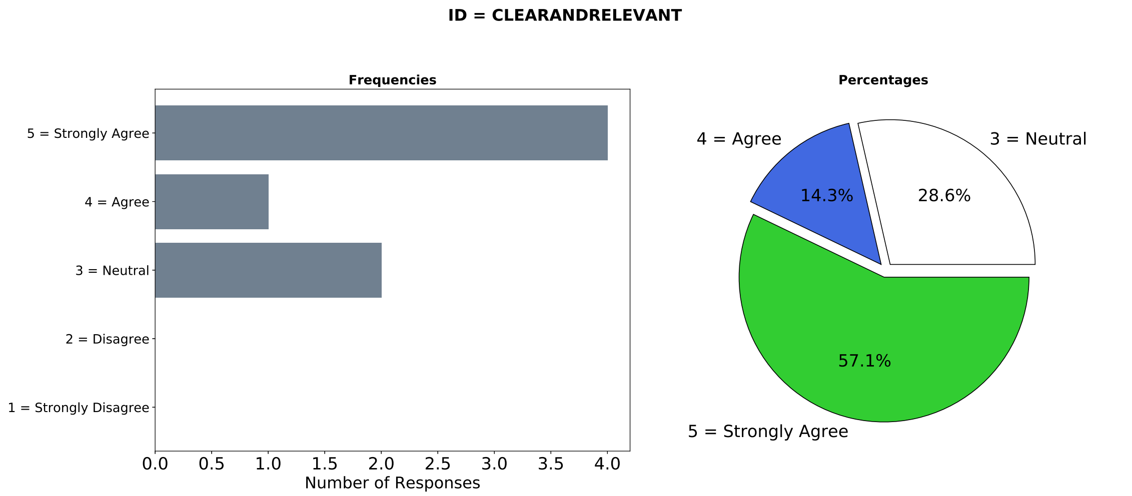
Two out of seven left a comment. Both expressed concerns about the colours used for the interface, while one is concerned about the saliency of information and suggests to improve the interface *"with additional colours: green for validated or done and red for delete"* (R05) the other one expressed concerns about the extreme colour contrasts on the Moon, explaining that *"the black sky is very black while the surface can seem very bright."* (R04) and suggests to investigate *"which colour would work best with those contidions [conditions] to make sure astronauts don't tire their eyes too much"* (R04).

In Figure 6.10 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: CLEARANDRELEVANT, namely:

> *"The interface elements (e.g., text, icons, labels, objects) are clear and relevant."*

Most respondents indicated either strong agreement or agreement (71.4%, N=5) with the fact that the application interface elements are clear and relevant (Mdn=5, IQR=1.5).

Two out of seven left a comment. While one agrees with the statement and adds that the icons are *"simple"* (R06) and *"clean"* (R06), the other one argues that it is *"impossible to judge without a test with realistic background and illumination conditions"* (R01).

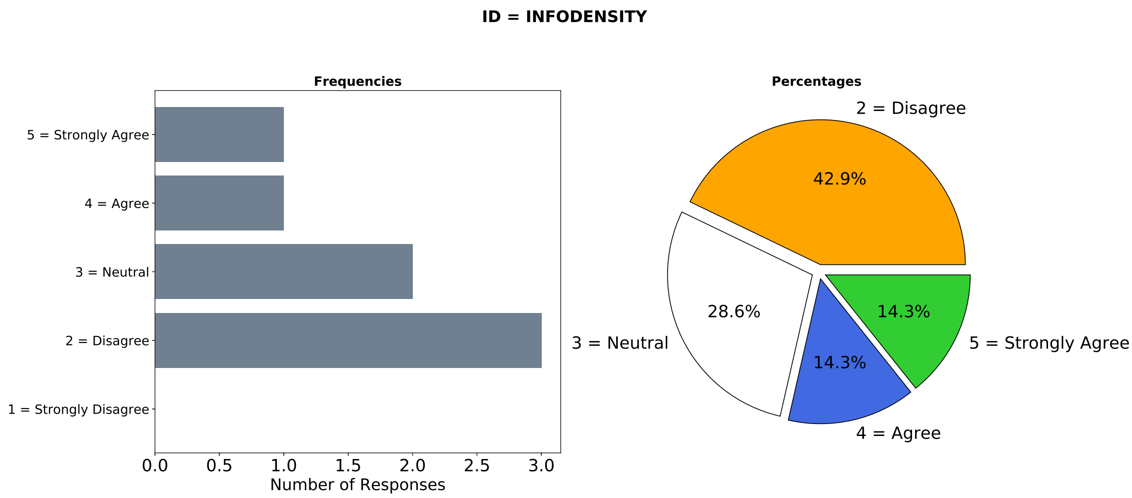


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	5.0	5.0	3.5	5.0	1.5	0.0

Figure 6.10: Frequencies and percentages for the response to ID: CLEARANDRELEVANT, N = 7, missing response(s) = 0.

In Figure 6.11 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INFODENSITY, namely:

> *"Information density is held to a minimum in displays used for critical geological site inspection tasks."*



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	3.0	5.0	2.0	3.5	1.5	0.0

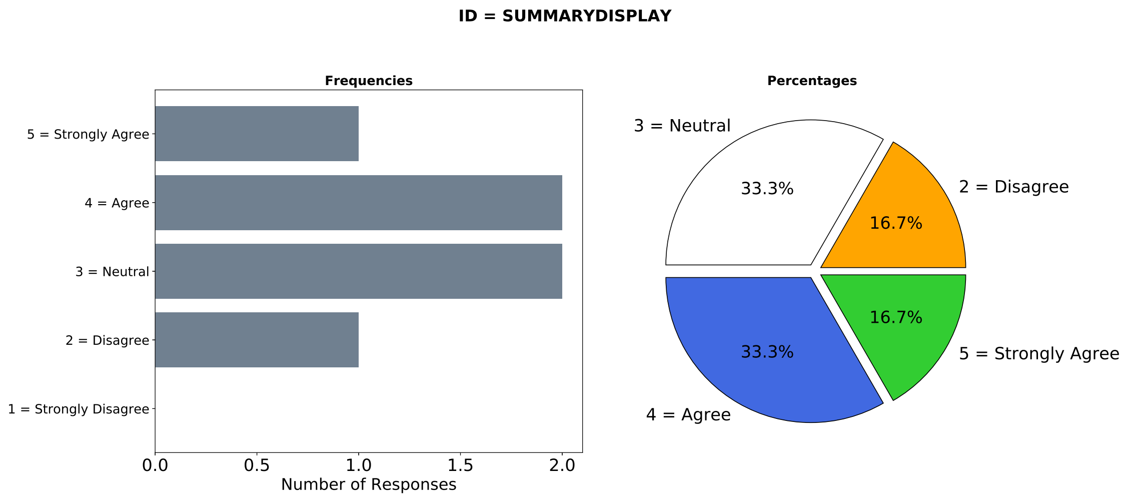
Figure 6.11: Frequencies and percentages for the response to ID: INFODENSITY, N = 7, missing response(s) = 0.

Opinions seem to be divided with regard to the fact that the density of information is held to the minimum for critical tasks within the interface. Many respondents (N=3, 42.9%) expressed disagreement, while 28.6% (N=2) remained neutral and another 28.6% (N=2) indicated that they agreed or strongly agreed (Mdn=3, IQR=1.5).

Four out of seven provided comments. While three of the four perceive that there is *"still too much"* (R02) information, one explains that it is because *"too many windows remain open at the same time"* (R01), while two of them propose to improve this by adjusting it *"depending on the situation (ex. A big waypoint flag seen from far away and almost transparent of smaller when on site)"* (R05) or *"grey out not needed info"* (R02). Another suggestion includes the size reduction of *"some of the blue panels"* (R05) as well as the *"waypoint flag"* (R05).

In Figure 6.12 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: SUMMARYDISPLAY, namely:

> *"Primary information required for performing a geological site inspection task is on a summary display."*



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	3.5	5.0	3.0	4.0	1.0	1.0

Figure 6.12: Frequencies and percentages for the response to ID: SUMMARYDISPLAY, N = 7, missing response(s) = 1.

Overall opinions are neutral leaning towards agreement with regard to the presence of a summary display for primary information (Mdn=3.5, IQR=1), in fact a third of the respondents (33.3%, N=2) remained neutral, another third (33.3%, N=2) indicated that they agree, while among the remaining respondents one disagrees and the other one strongly agrees. Note that one response is missing.

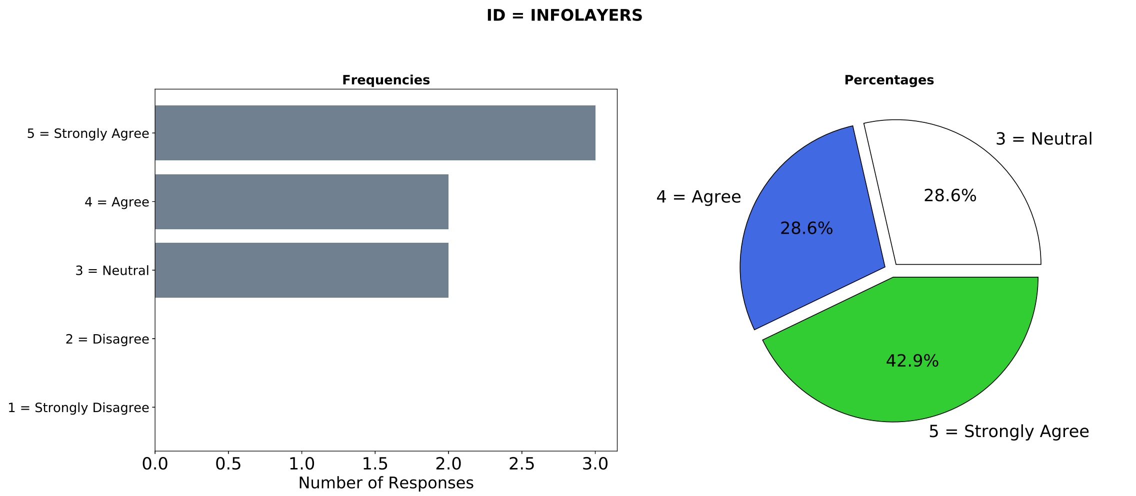
Three out of seven provided a comment, they all agree that *"information will have to be extended"* (R03), two of the three suggest adding the *"site description (verbally or written)"* (R06) task and *"AR name & distance overlays of horizon features (craters, monts, hills) and of vital assets (lander position, rover position, EVA buddy position) as well as access to maps and description cheat sheets or procedure guidelines"* (R05).

In Figure 6.13 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INFOLAYERS, namely:

> *"Information layering, via secondary displays or dialog boxes, is implemented to provide supplemental information in support of the primary display (e.g., specify options available to the crewmember or to provide details)."*

Most respondents (71.5%, N=5) indicated either strong agreement or agreement with the fact that the interface implements information layering for additional information (Mdn=4, IQR=1.5).

Three out of seven added a comment. While one of them believes that *"the options and information displays were well implemented"* (R06) the other two disagree and are concerned that it leads to having *"too many virtual displays around the astronauts, which would hide the geological details in his/her field-of-view and be detrimental to the situational awareness and the safety. With all the displays opened in the demo video, we can barely observe what is around us in the room"* (R05) one of the two suggests to *"minimise the maximum number of open displays"* (R02).

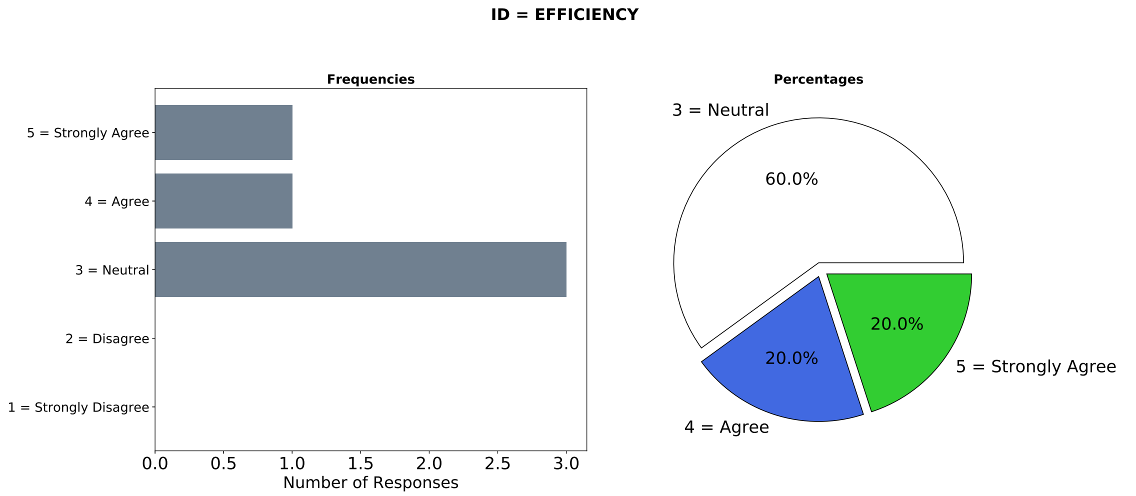


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	3.5	5.0	1.5	0.0

Figure 6.13: Frequencies and percentages for the response to ID: INFOLAYERS, N = 7, missing response(s) = 0.

In Figure 6.14 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: EFFICIENCY, namely:

> *"The interface is designed for efficient use of crew time and to minimize crew and flight controller training time."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	3.0	5.0	3.0	4.0	1.0	2.0

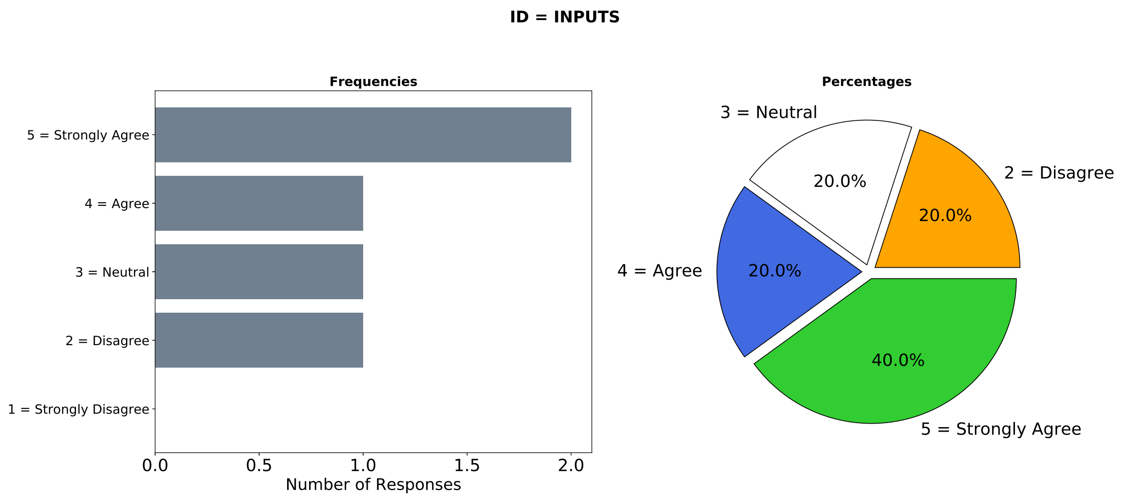
Figure 6.14: Frequencies and percentages for the response to ID: EFFICIENCY, N = 7, missing response(s) = 2.

Most respondents (60%, N=3) remained neutral about the fact that the interface is designed for efficient use of crew time and required minimum training (Mdn=3, IQR=1). Note that two respondents did not rate this item.

Three out of seven respondents added a comment. Two of them agree that there is room for improvement in terms of optimising the interface. While one expressed concerns about *"the tempo with the filling bar for the confirmation of what the eyes look at"* (R05) being *"too long"* (R05) another one fears that users will be *"spending a little too much time fussing with the boxes/where to pin icons"* and that *"some dialogue boxes, like “log data” box are bigger than they need to be by default"* will be getting in the way and for these suggests to *"expand as more data is added to them"* (R06). The other one of the three argues that *"this can only be judged in a comparative way"* (R01).

In Figure 6.15 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INPUTS, namely:

> *"The number of user inputs e.g. gestures/voice/gaze needed to perform simple or routine functions is minimized."*



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	4.0	5.0	3.0	5.0	2.0	2.0

Figure 6.15: Frequencies and percentages for the response to ID: INPUTS, N = 7, missing response(s) = 2.

Opinions seem to differ regarding the fact that the number of user inputs required are minimized, even though 60% (N=3) agrees or strongly agrees, 20% (N=1) disagrees and the other 20% (N=1) remains neutral, moreover two responses are missing (Mdn=4, IQR=2).

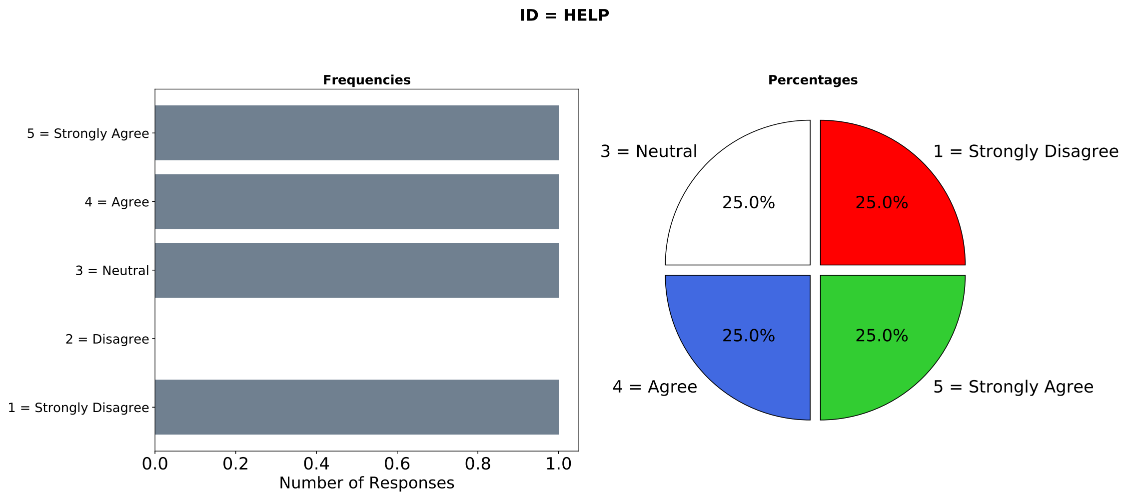
Three out of seven left a comment regarding this questionnaire item. While one of the respondents left positive feedback regarding the required user inputs noting that only a *"few ‘gaze-clicks’ are needed to perform important tasks"* (R06), the other two respondents raised concerns about the fact that the *"tempo with the filling bar for the confirmation of what the eyes look at is too long"* (R05) and the *"button as a user input"* (R07). It is argued that the button is *"rather impractical, and will delay (or even prevent) the implementation of such a system in a spacesuit"* (R07), the suggestion was made to *"alternate input signals, such as repeated eye-lid blinking (e.g. a triple blink) or short sounds that do not trigger the voice VOX, like clicking with the tongue"* (R07).

In Figure 6.16 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: HELP, namely:

> *"A help function is accessible to the crewmembers."*

Opinions seem to be divided with regard to the fact that a help function is provided (Mdn=3.5, IQR=1.75). Additionally, 42.8% of the respondents did not rate this item hence no actual conclusions can be made [95].

Two out of seven provided with a comment, of those three two are sceptical, one mentions to *"not see a help function"* (R01). Only one of the three seemed to have "identified" the help function and mentions that: *"there are help boxes site marking procedures"* (R06).

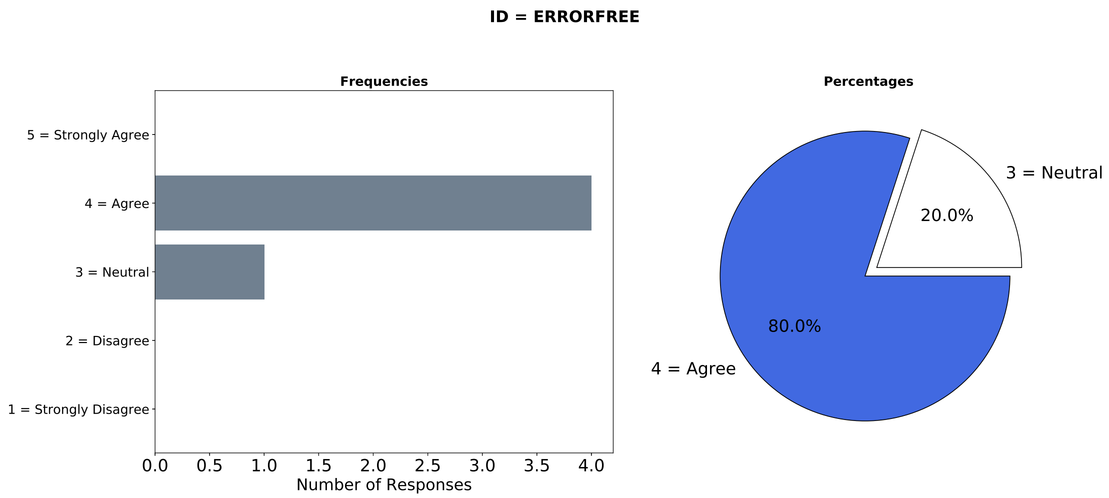


count	min	median	max	25%	75%	IQR	missing values
7.0	1.0	3.5	5.0	2.5	4.25	1.75	3.0

Figure 6.16: Frequencies and percentages for the response to ID: HELP, N = 7, missing response(s) = 3.

In Figure 6.17 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: ERRORFREE, namely:

> *"Display design facilitates error-free operations."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	4.0	4.0	4.0	0.0	2.0

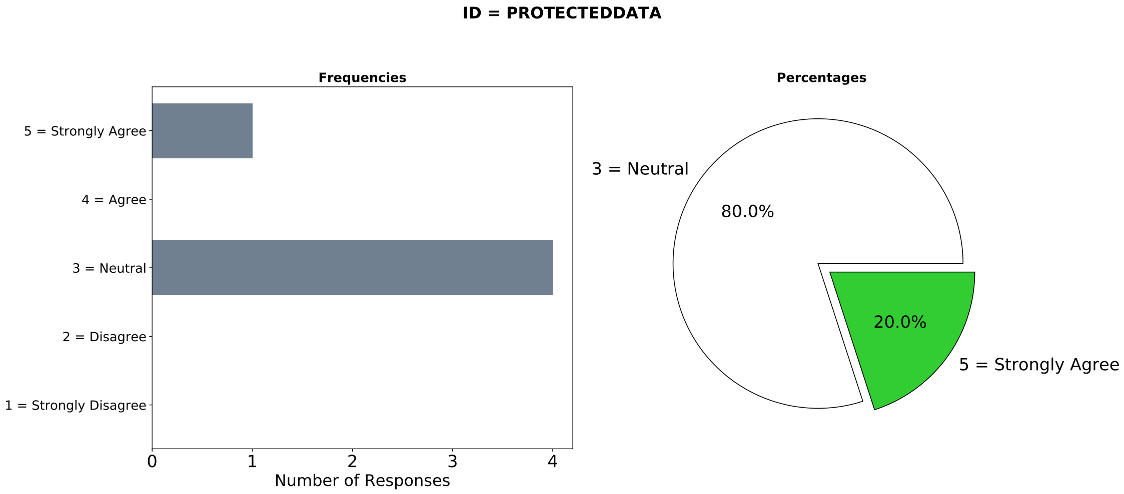
Figure 6.17: Frequencies and percentages for the response to ID: ERRORFREE, N = 7, missing response(s) = 2.

Most respondents indicated agreement with the fact that the application facilitates error-free operations (Mdn=4, IQR=0). Nevertheless, it should be noted that two responses are missing.

Two out of seven provided with a comment, both agree that there is room for improvements, one suggest adding a *" “back” or “undo” button"* (R04), the other mentions possible improvements depending on technology advancements in *"eye-tracking software"* (R06).

In Figure 6.18 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: PROTECTEDDATA, namely:

"Data is protected from inadvertent errors and hardware failures e.g. frequent saves."



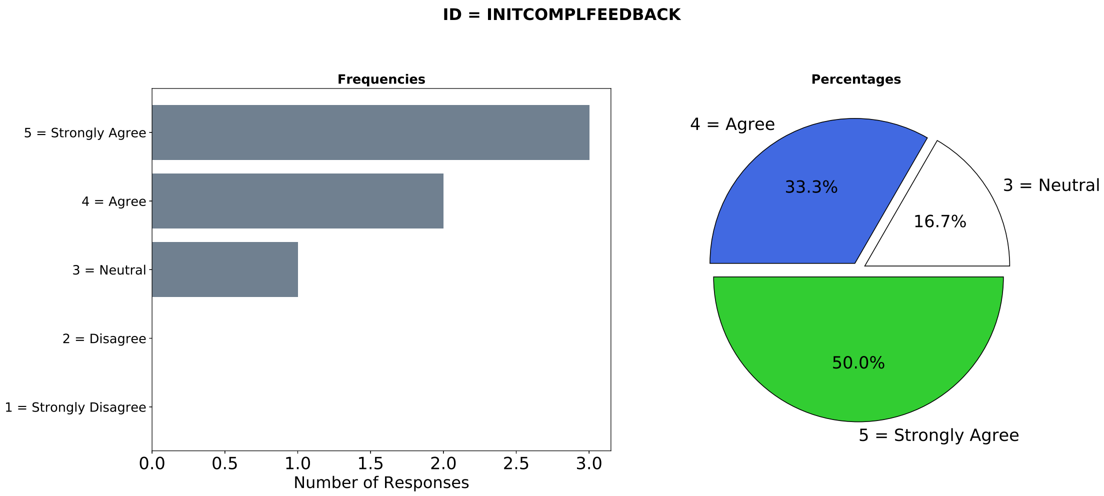
count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	3.0	5.0	3.0	3.0	0.0	2.0

Figure 6.18: Frequencies and percentages for the response to ID: PROTECTEDDATA, N = 7, missing response(s) = 2.

Most respondents remain neutral regarding the fact that the application ensures that the data is protected (Mdn=3, IQR=0). Also, two responses are missing.

In Figure 6.19 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INITCOMPLFEEDBACK, namely:

"When a process is initiated or completed, crew member feedback is provided."



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.5	5.0	4.0	5.0	1.0	1.0

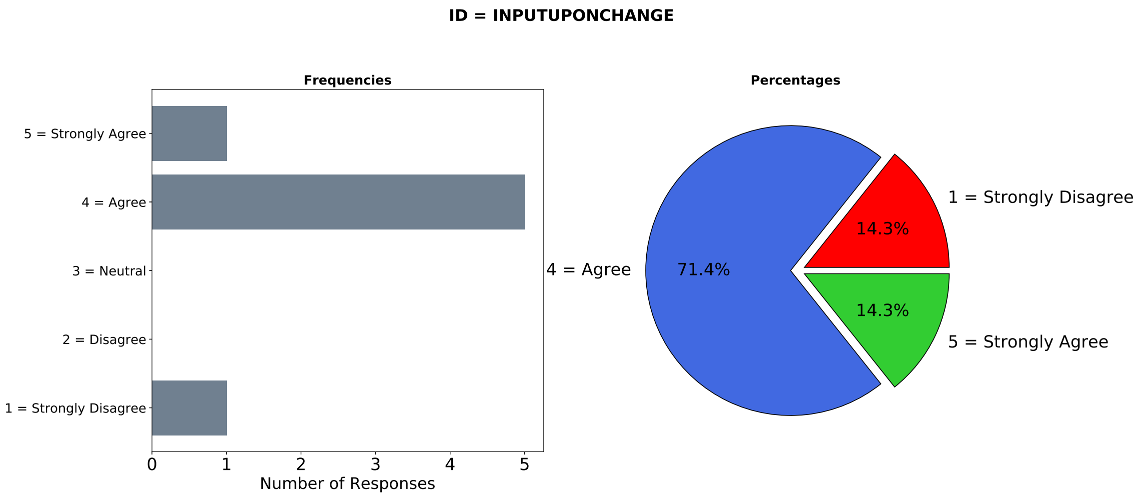
Figure 6.19: Frequencies and percentages for the response to ID: INITCOMPLFEEDBACK, N = 7, missing response(s) = 1.

Most respondents indicated agreement with the fact that the application provides with feedback upon process initiation and completion (Mdn=4.5, IQR=1). One response is missing though.

Only one out of seven left a comment agreeing with the statement acknowledging the fact that *"there was the little load bar on each icon"* (R06).

In Figure 6.20 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INPUTUPONCHANGE, namely:

> *"When an input is required, an indication is provided to the crewmember, e.g., a cursor change. "*



count	min	median	max	25%	75%	IQR	missing values
7.0	1.0	4.0	5.0	4.0	4.0	0.0	0.0

Figure 6.20: Frequencies and percentages for the response to ID: INPUTUPONCHANGE, N = 7, missing response(s) = 0.

Most respondents indicated agreement with the fact that the application informs the user when an input is required (Mdn=4, IQR=0).

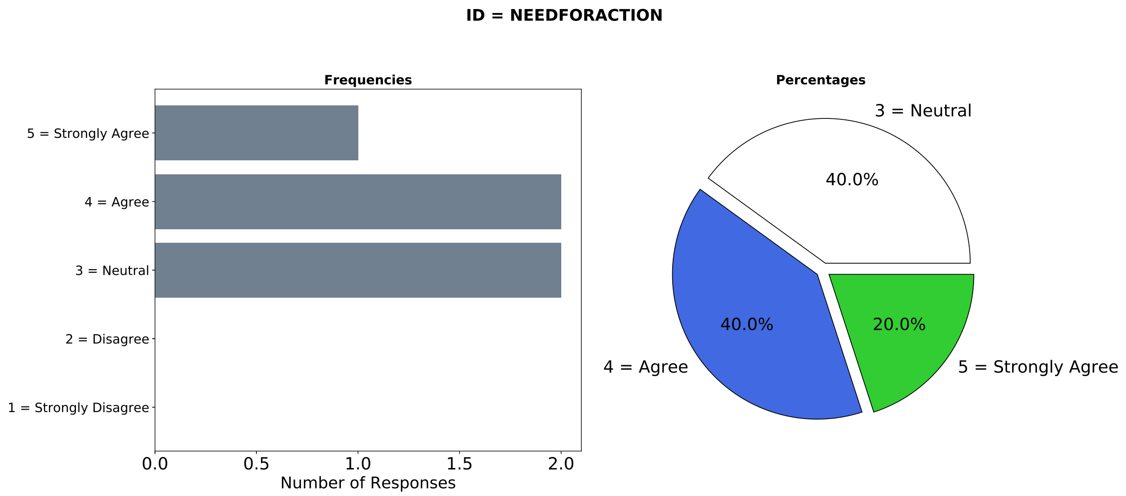
Three out of seven provided with a comment. Two out of these three agreed that input was not always clear as one did not know *"when input is required based on the visual interface"* (R06) and the other mentioned that *"the small writings that appear about [above] the icons"* (R01) *"are hardly readable"* (R01). The third respondent suggested *"using additional colours like green or red"* (R05) which *"could enhance the feedback"* (R05).

In Figure 6.21 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: NEEDFORACTION, namely:

> *"If the completed command implies the need for further crewmember action, the need for action is indicated. "*

Most respondents indicated agreement with the fact that the application informs the user in case further action is needed (Mdn=4, IQR=1). Two responses are however missing.

Two out of seven provided with a comment. One of the two mentions that *"in the demo the user created stops without following the "take photo" and "take 360" steps but it was not clear to me if the system indicated they needed to take further action"* (R06), the other suggests to improve this with *"additional 'HUD' banners"* (R03).

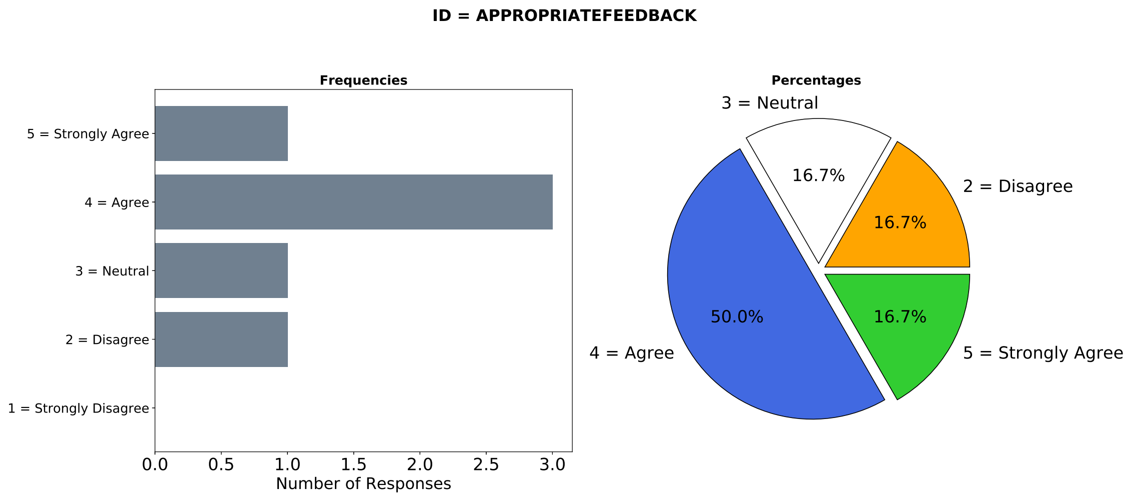


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	3.0	4.0	1.0	2.0

Figure 6.21: Frequencies and percentages for the response to ID: NEEDFORACTION, N = 7, missing response(s) = 2.

In Figure 6.22 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: APPROPRIATEFEEDBACK, namely:

"The application responds to crewmember interaction with appropriate feedback."



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	4.0	5.0	3.25	4.0	0.75	1.0

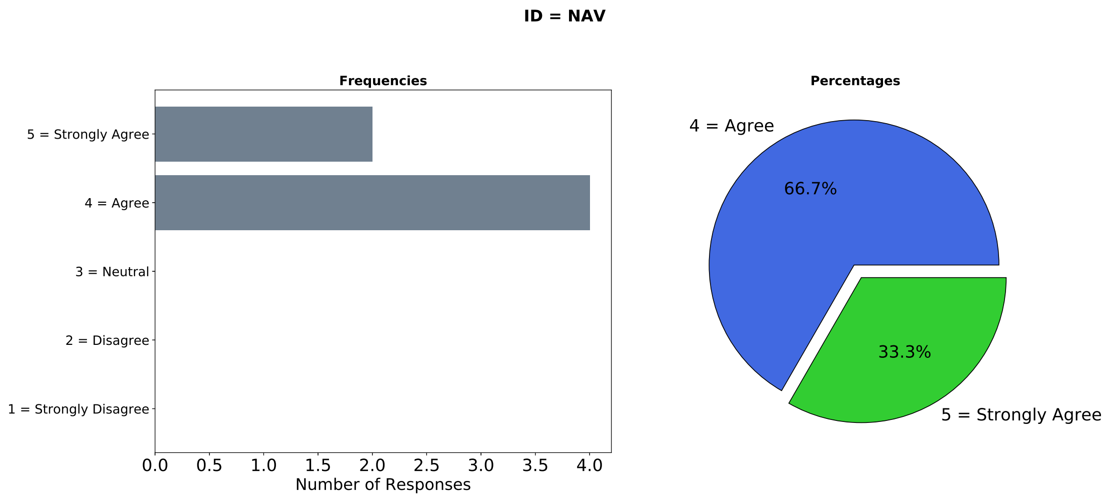
Figure 6.22: Frequencies and percentages for the response to ID: APPROPRIATEFEEDBACK, N = 7, missing response(s) = 1.

Most respondents indicated agreement with the fact that the application responds with appropriate feedback (Mdn=4, IQR=0.75). One response is missing though.

Two out of seven provided with a comment, both agree that feedback is not always appropriate as *"taking pictures does not show the picture that you take for control before sending"* (R05) also *"there is a significant delayed response time in opening some of the tabs"* (R06), both make suggestions, one of the two suggests that this *"could be mitigated if they could multitask (i.e. doing something with their hands) while navigating the application"* (R06), the other one suggests a *"validation with a green label"* (R05) instead of *"a growing light blue bar"* (R05).

In Figure 6.23 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: NAV, namely:

"Navigation through the interface is clear and intuitive."



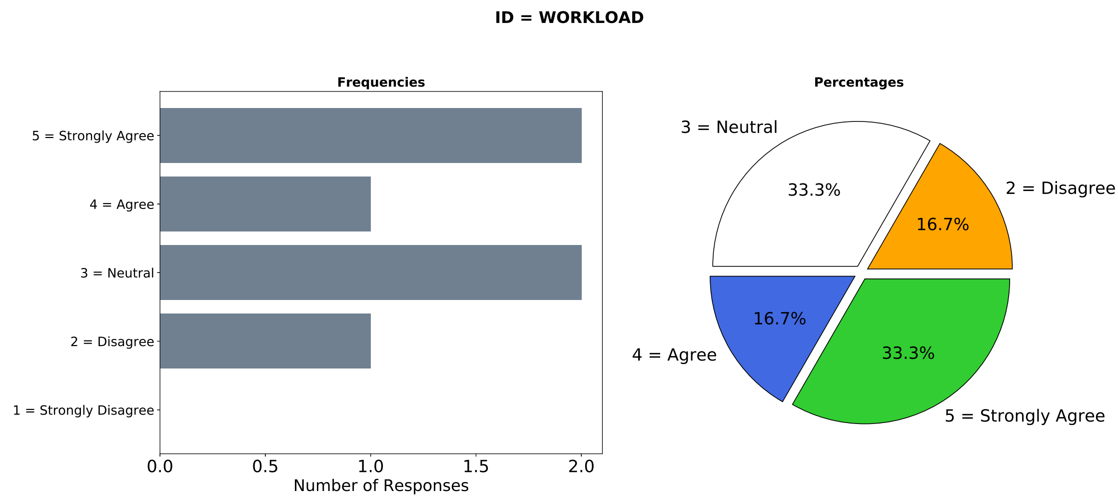
count	min	median	max	25%	75%	IQR	missing values
7.0	4.0	4.0	5.0	4.0	4.75	0.75	1.0

Figure 6.23: Frequencies and percentages for the response to ID: NAV, N = 7, missing response(s) = 1.

Most respondents indicated agreement with the fact that navigation through the interface is clear and intuitive (Mdn=4, IQR=0.75). Nonetheless, one response is missing.

In Figure 6.24 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: WORKLOAD, namely:

"The application minimizes the user's mental workload."



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	3.5	5.0	3.0	4.75	1.75	1.0

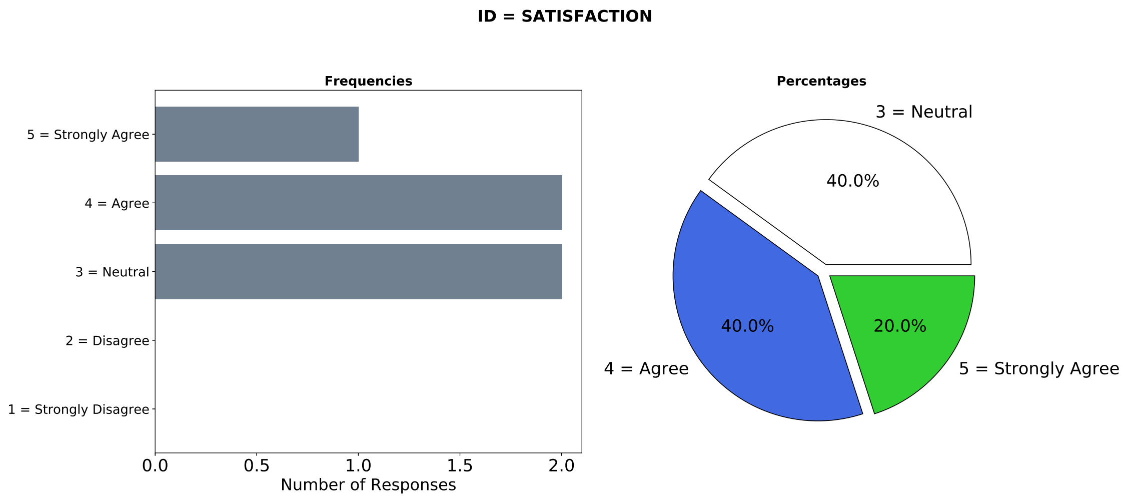
Figure 6.24: Frequencies and percentages for the response to ID: WORKLOAD, N = 7, missing response(s) = 1.

Opinions seem to be divided with regard to whether the application minimizes users' mental workload. While 33.3% (N=2) of the respondents remained neutral and 16.7% (N=1) expressed disagreement, an equal number (50%, N=3) indicated that they agreed or strongly agreed (Mdn=3.5, IQR=1.75). Note that one respondent did not rate this item.

Four out of seven added a comment. Two out of these four agree with the statement and mention positive aspects such as *"icons together with text"* (R04) being a *"great idea"* (R04) as well as the fact that *"the HoloLens reduces that [one of the most tedious but important tasks in terrestrial field geology, is taking site coordinates and detailed photo-documentation and descriptions of the sites] by having the GPS and verbal "field notebook" easily assessable in one application"* (R06). The other two are *"not convinced"* (R05) and argue that *"voicing up (like in SIRI) to a speech recognition system to confirm [an] AR action uses much less mental workload than looking for a precise point in an AR display, glazing [gazing] carefully and steadily at it during the tempo incrementation of the blue bar for 2 to 3 seconds"* (R05), moreover, *"moving an AR display or a virtual flag with the interaction of your hand in the field-of-view (like the Hololens allows it) is much easier and less mental demanding than moving your eye to transport the item attached to your glancing position"* (R05). Finally, it is argued that *"when all AR displays and sub displays are opened it is requiring quite some mental workload to look at all of them to search what you need while keeping enough situational awareness around you to complete your task efficiently and safely"* (R05). The other one of the two suggests that one *"can only judge "minimize" if you can compare with other options"* (R01).

In Figure 6.25 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: SATISFACTION, namely:

> *"The interface is satisfying for crewmembers and training members."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	3.0	4.0	1.0	2.0

Figure 6.25: Frequencies and percentages for the response to ID: SATISFACTION, N = 7, missing response(s) = 2.

Most respondents indicated agreement with the fact that the application is satisfying (Mdn=4, IQR=1). Note however that two did not respond.

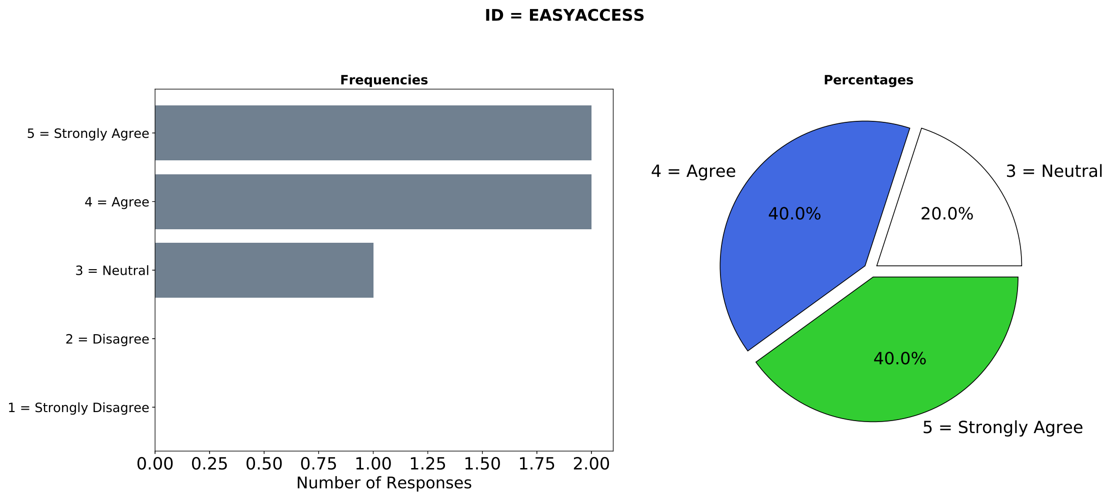
Two out of seven left a comment both agreed that *"without wearing it"* (R06) and without even being *"trained on it"* (R05), *"it is impossible to reply"* (R05).

In Figure 6.26 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: EASYACCESS, namely:

> *"The required information is easily found and accessed."*

Most respondents indicated agreement with the fact that information is easily found and accessible (Mdn=4, IQR=1). Two responses are however missing.

One out of seven left a comment mentioning that *"the displayed information is easily found"* (R05) however that there are *"easier way [ways] to access it (voice recognition, hand movement in the filed [field] of view to interact with the AR displays"* (R05).

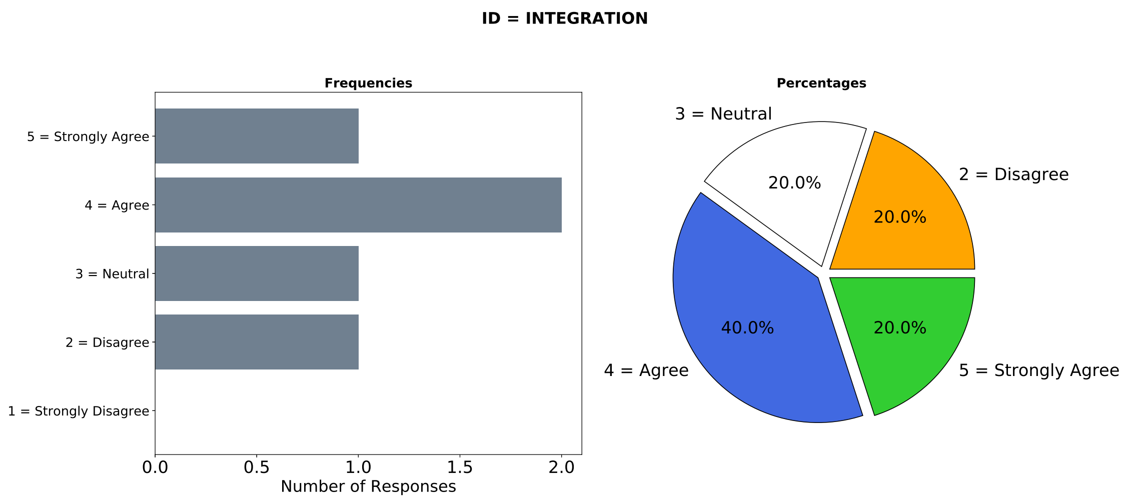


count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	5.0	1.0	2.0

Figure 6.26: Frequencies and percentages for the response to ID: EASYACCESS, N = 7, missing response(s) = 2.

In Figure 6.27 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: INTEGRATION, namely:

> *"It is possible to understand what the features of the application represent and to realize the integration of the different media."*



count	min	median	max	25%	75%	IQR	missing values
7.0	2.0	4.0	5.0	3.0	4.0	1.0	2.0

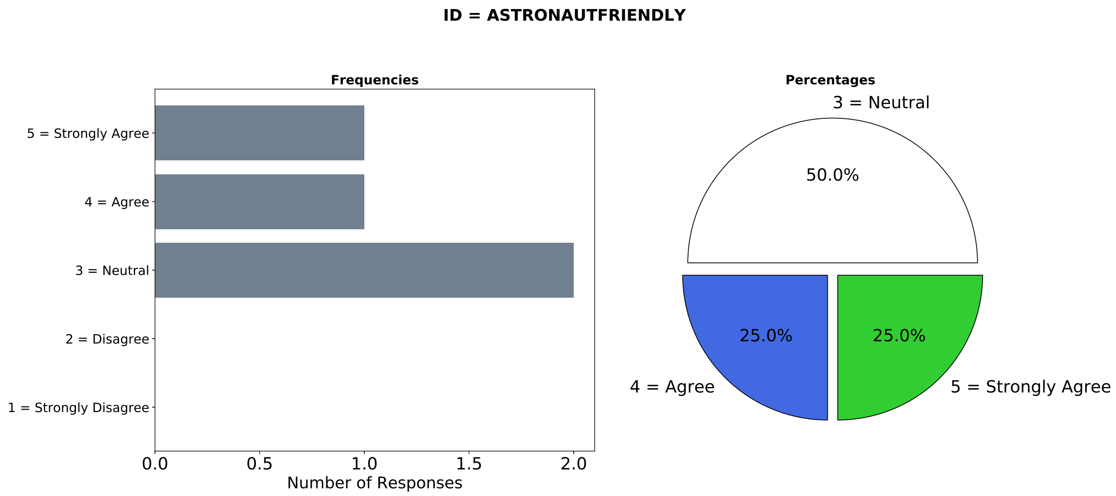
Figure 6.27: Frequencies and percentages for the response to ID: INTEGRATION, N = 7, missing response(s) = 2.

Most respondents indicated agreement with the fact that the features of the application are understandable as well as the integration of the different media (Mdn=4, IQR=1). However, it should be noted that two responses are missing.

Only one out of seven left a comment, suggesting to integrate the displaying of *"photos"* (R05) and *"data from instrument analysis"* (R05).

In Figure 6.28 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: ASTRONAUTFRIENDLY, namely:

> *"The application is "astronaut crew-friendly"."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	3.5	5.0	3.0	4.25	1.25	3.0

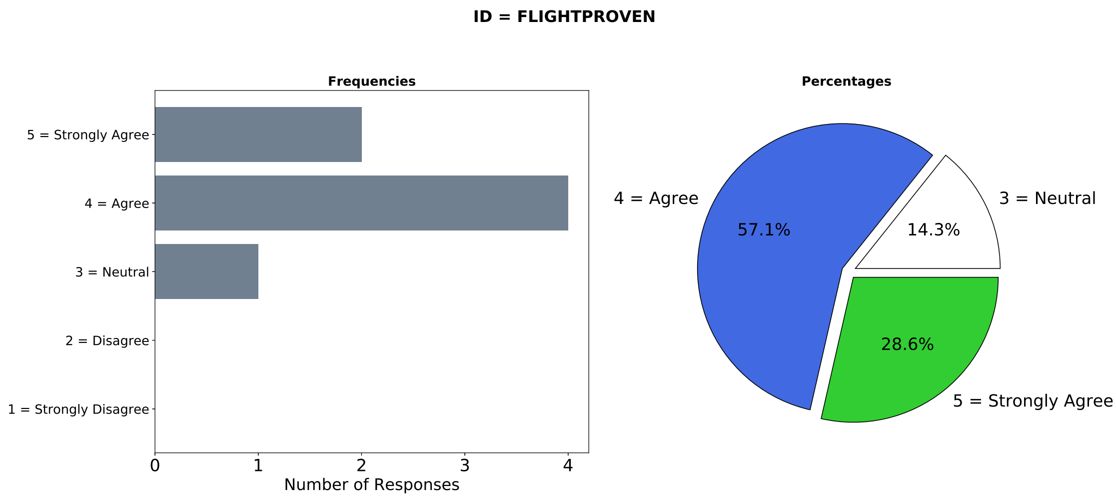
Figure 6.28: Frequencies and percentages for the response to ID: ASTRONAUTFRIENDLY, N = 7, missing response(s) = 3.

Overall opinions are rather neutral leaning towards agreement regarding the fact that the application is astronaut-friendly in fact, while three responses are missing, half of the respondents (N=2) remain neutral, the other half expresses agreement or strong agreement (Mdn=3.5, IQR=1.25). No conclusions can be made regarding this point as the percentage of missing responses is 42.8% [95].

Three out of seven left a comment. One of the three felt that the interface was astronaut-friendly as *"overall, it is very straightforward"* (R06) while another one said that it *"very much depends on how easy it is to navigate with the gaze"* (R01). Finally, the third respondent believes that *"using more voice recognition or hand gesture interaction with the AR environment would be more astronaut crew friendly"* (R05).

In Figure 6.29 the frequencies and percentages, median and IQR are shown for the questionnaire item with ID: FLIGHTPROVEN, namely:

> *"The application has the potential to become "flight proven" assuming that augmented reality technology will be an integral part of the astronauts' EVA helmet."*



count	min	median	max	25%	75%	IQR	missing values
7.0	3.0	4.0	5.0	4.0	4.5	0.5	0.0

Figure 6.29: Frequencies and percentages for the response to ID: FLIGHTPROVEN, N = 7, missing response(s) = 0.

Most respondents indicated agreement with the fact that the application has the potential to become flight proven (Mdn=4, IQR=0.5).

Five out of seven provided with feedback. Three respondents agree that it is a step *"constrained by technology limitations"* (R03) *"e.g. electronics in the 100% O_2 atmosphere within the suit"* (R02) which represent a *"fire hazard"* (R02). The other two agree, one adds that once it is able to *"cover all functions and to show that everything is easily accessible and reliable and compatible with the maintenance of a good situational awareness and work efficiency"* (R05) it will become flight-proven. The other one added: to be *"convinced that augmented reality systems will play an important role in the future of human space flight, because they reduce the amount of training required for a flight (e.g. to Mars), and they allow to work scientifically inside a spacesuit, and, generally, with systems that we haven't trained for. That's why since many years, astronauts propose to develop these systems. A very important part is that they are easy to use, i.e. quick to set up and without cables. I would have loved to be able to use such systems on my two flights to ISS"* (R07).

6.4. Astronaut and Astronaut Geological Field Training Experts Interview Results

The results from the qualitative content analysis performed on the semi-structured in-depth interviews with experts, namely ESA astronauts and astronaut geological field training experts are reported hereafter. The main extracted categories and associated subcategories as well as representative quotes are presented. A total of four interviews were conducted with two ESA astronauts and two support engineers of the Caves & Pangaea team. The semi-structured interviews differed in length ranging from 20 to 70 minutes.

6.4.1. Main Categories and Subcategories

The qualitative content analysis of the four conducted interviews led to the identification of 113 quotations. 22 quotes were assigned to two subcategories and two quotes to three subcategories. Therefore, the total number of classified quotes is 139. The 139 quotes constituting 25 subcategories were then assigned to nine main categories relevant to the experts' experiences and opinions with new technologies focusing on aspects related to usability, user interaction and implementation as well as experiences with new concepts of operations. A summary of all the extracted subcategories and identified main categories as well as related sources are shown in Table 6.2.

The main categories identified are as follows:

1. *User satisfaction*
2. *Effectiveness*
3. *Efficiency*
4. *Workload*
5. *Situational awareness*
6. *Training*
7. *User interaction*
8. *Implementation aspects*
9. *Concepts of Operations*

Table 6.2: Main categories and associated subcategories, as well as the sources related to the subcategories.

Category numbering	Categories	Subcategories	Sources
1.	User satisfaction	Annoying Complex Easy usage Helpful Information in the field-of-view Integration of different systems Simple Useful	newly created [2], [93] [2], [93] [82], [139] [15], [55], [79], [82], [139] [2], [93] [55] newly created
2.	Effectiveness	Effective	[173]
3.	Efficiency	Time consuming	[173]
4.	Workload	Hard task	newly created
5.	Situational awareness	Situational awareness	[32], [55], [123], [173]
6.	Training	Training	[103]
7.	User interaction	Hands-free Voice Button/clicker Double confirmation Interaction	[15], [55], [79], [125], [126], [146], [198], [221] [3], [55], [126], [139], [146], [154], [230] newly created newly created [96], [173]
8.	Implementation aspects	Adaptation of technology rather than human AR integrated in the helmet Easy data transfer Operationally feasible Enable remote support Work automation and/or sharing work with off-site scientists	newly created newly created [158], [197] newly created [14], [51], [98], [130], [175], [212] [135]
9.	Concepts of operation	Concepts of operation	[121], [122], [135], [173]

As mentioned in subsection 3.2.3 the selection criterion and associated level of abstraction was aimed at extracting information from the text related to usability, user interaction and implementation aspects of new technologies. Moreover, information on concepts of operations were extracted as well.

According to the ISO standard 9241-11 [92] usability refers to the extent to which a product can be used by specified users to achieve specified goals with efficiency, effectiveness, and satisfaction in a specified context of use.

Here, satisfaction is defined as the users' comfort with respect to the use of the system as well as the users' positive attitudes toward this [173]. Therefore, *annoying*, *complex*, *simple*, *easy usage*, *useful*, *helpful*, *information in the field-of-view* and *integration of different systems* have been assigned to the *user satisfaction* main category (see Table 6.2). It should be noted that some of those aspects can be found mentioned explicitly in the System Usability Scale (SUS) [2], [93] (see Appendix E) such as: *ease of use*, *integration* and *complexity*. While other aspects can be found in the Smart Glasses User Satisfaction (SGUS) questionnaire [82], [139] (see Appendix F) such as: information location and helpfulness. The importance of keeping an interface, that is meant for astronauts, simple has been found by Anandapadmanaban et al. [55] as well. *Useful* and *annoying* are newly created.

Effectiveness refers to the accuracy and completeness with which a goal is achieved. Quality of the user's solution and error rates are effectiveness indicators [173]. This attribute was mentioned by respondents and hence assigned to a distinct category named *effectiveness* accordingly.

Efficiency is defined as the relation between accuracy and completeness with which users achieve a preset goal and the resources used to achieve it. Task completion time and learning time are indicators of efficiency [173]. Efficiency as an attribute was mentioned explicitly by only one of the respondents hence it was not included, however the subcategory *time consuming* was quite often mentioned and thus assigned to the main category *efficiency* as these are strongly connected.

Workload is a measure for astronaut operations to be considered especially for EVAs. It is defined as the mental and physical effort required to meet the demands of the assigned task [173] and is considered a metric of usability [192]. *Hard task* was newly created and assigned to the *workload* category.

NASA-STD-3001, volume 2 [178] specifies workload requirements. With respect to cognitive workload it is stated in V2 5007 that:

> *"Cognitive workload shall be accommodated in the design of all system elements that interface with the crew for all levels of crew capability and all levels of task demands"*

In NASA's Human Integration Design Processes document [173] it is stated that:

> *"Display formats must provide situational awareness, reduce crew workload, and enhance crew safety by providing readily understood graphical and textual subsystem information in a timely manner."*

Situational awareness was thus defined as main category as it is considered an important aspect when designing displays for astronaut crew and can be considered a usability aspect.

The main category *training* is connected to the learnability principle, one of the 5 usability principles defined by Nielsen [103]. Learnability refers to how easy is it for users to successfully complete a basic task the first time they deal with the product, hence to how much training it requires.

The main category *user interaction* includes different modes of interaction namely: *hands-free*, *voice* e.g. speech recognition, as well as with hand gestures e.g., via a mechanical interface such as a *button/clicker*. These different modes of interaction have been investigated in some space-related AR studies as well [79], [123], [125], [139], [146], [198] [221], [230]. Hands-free interaction has been highly recommended by Apollo astronauts as well [154]. The *button/clicker* as well as the *double confirmation* subcategory have been newly created.

The category *implementation aspects* includes recommended features and concerns related to the implementation of new technologies. Some categories were newly created, others such as: *easy data transfer* is a feature that has been indirectly recommended by Apollo astronauts [197] and other AR space-related studies [158]. Moreover, *enabling remote support* has been recommended by Apollo astronauts and in many analogue studies [14], [51], [98], [130], [175], [212] while *work automation by "smart" tools and/or sharing work with off-site scientists* is an aspect that is being investigated by Caves & Pangaea analogue campaigns as well [135].

The *concepts of operations* category includes experiences related to newly tested concepts of operations. The development of concepts of operations is an important aspect when designing new technology as it gives an overall picture of the operation from the users' perspective [173]. It is also part of the user and task analysis process of the user-centered design methodology [173] described in section 3.2 hence it was considered fundamental to get additional insights on this subject.

In Table 6.3 the number of subcategory occurrences and the number of respondents per subcategory is shown in absolute numbers and percentages. It should be noted that additionally under the number of respondents it is specified in brackets whether the respondents were majorly astronauts denoted as 'astro' or support engineers/developers denoted as 'dev'.

Table 6.3: Number of subcategory occurrences and respective main category occurrences within the material, percentage of occurrences of subcategories and respective main category occurrences within the material, number of respondents (number of astronauts denoted as 'astro' and number of support engineers/tool developers denoted as 'dev') per subcategory as well as the percentage of respondents w.r.t. all respondents (N=4).

Category and corresponding subcategories	**Number of occurrences**	**% of occurrences w.r.t. all subcategories**	**Number of respondents**	**% w.r.t. all respondents**
1. User satisfaction	**45**	**32%**		
annoying	6	4%	2 (2 astro)	50%
complex	3	2%	2 (2 astro)	50%
easy usage	4	3%	3 (1 astro, 2 dev)	75%
helpful	7	5%	3 (2 astro, 1 dev)	75%
information in the field-of-view	4	3%	3 (1 astro, 2 dev)	75%
integration of different systems	7	5%	3 (1 astro, 2 dev)	75%
simple	8	6%	2 (1 astro, 1 dev)	50%
useful	6	4%	3 (1 astro, 2 dev)	75%
2. Effectiveness	**2**	**1%**		
effective	2	1%	2 (1 astro, 1 dev)	50%
3. Efficiency	**5**	**4%**		
time consuming	5	4%	2 (2 astro)	50%
4. Workload	**3**	**2%**		
hard task	3	2%	3 (2 astro, 1 dev)	75%
5. Situational awareness	**4**	**3%**		
situational awareness	4	3%	2 (1 astro, 1 dev)	50%
6. Training	**4**	**3%**		
training	4	3%	3 (2 astro, 1 dev)	75%
7. User interaction	**33**	**24%**		
button/clicker	9	6%	4 (2 astro, 2 dev)	100%
double confirmation	3	2%	2 (1 astro, 1 dev)	50%
hands-free	5	4%	2 (2 astro)	50%
interaction	3	2%	2 (2 astro)	50%
voice	13	9%	2 (2 astro)	50%
8. Implementation aspects	**29**	**21%**		
adaptation of technology rather than human	3	2%	2 (2 astro)	50%
AR integrated in the helmet	9	6%	3 (1 astro, 2 dev)	75%
easy data transfer	2	1%	2 (1 astro, 1 dev)	50%
operationally feasible	2	1%	2 (2 dev)	50%
remote support	7	5%	2 (2 astro)	50%
work automation and/or sharing work with off-site scientists	6	4%	2 (1 astro, 1 dev)	50%
9. Concepts of Operation	**14**	**10%**		
concepts of Operations	14	10%	3 (1 astro, 2 dev)	75%
$\sum$	**139**	100%	4 (2 astro, 2 dev)	-

In Figure 6.30 a Pie chart is shown for the reader to get a better overview of the number of occurrences of each subcategory and to see which of these subcategories resulted having the highest number of quotes.

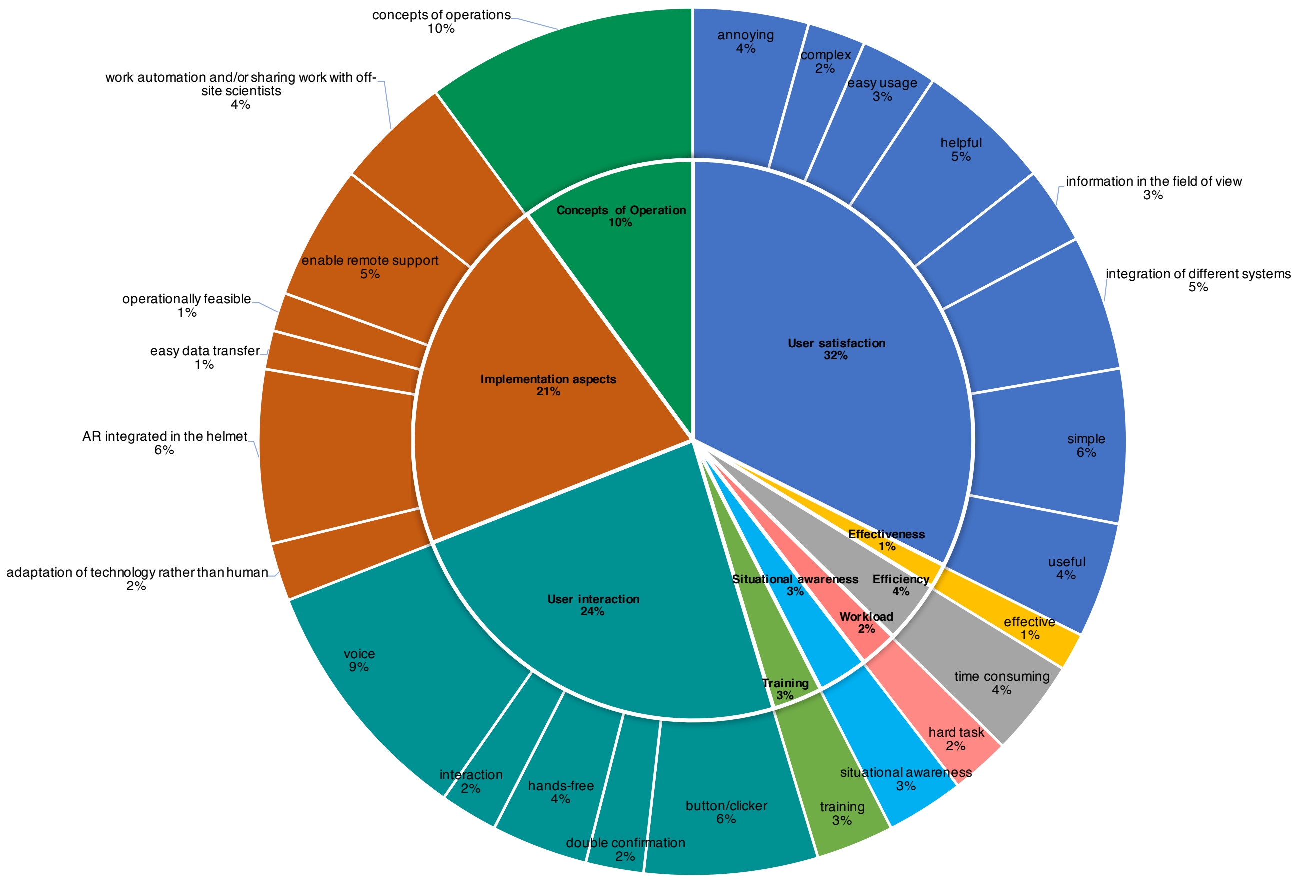


Figure 6.30: Pie chart showing the nine main categories and their associated 25 subcategories. The arc length of each sector is proportional to the number of quotes. The percentage of occurrences of each category and their subcategories are displayed on the corresponding sectors.

6.4.2. Category I: User Satisfaction

Subcategory: Annoying

Two respondents, both of them astronauts, addressed this subcategory highlighting what aspects of an AR application they would find annoying leading to poor user satisfaction.

One of them described previous experiences during analogue missions, in this case NASA's Extreme Environment Mission Operations (NEEMO), where crew were given to test different types of applications for maintenance and medical related tasks with the Microsoft HoloLens and said:

> *"it was annoying, it would get stuck or you know when you finally were able to change the step it would actually change two of them and then it would get stuck and then it wouldn't go back and then I mean it was just... so I was like if I just had a piece of paper in my hands." (R01)*

The other respondent referred to the AR-IoT tool developed for this project. Based on the video demonstration the astronaut mentioned that it is annoying when an application used for field work provides with too many options, namely fields to fill in. More specifically, the astronaut said:

> *"longitude, latitude, modify... honestly I would not need it in the field and it is just annoying to have too many options." (R02)*

Subcategory: Complex

Two respondents, both of them astronauts, mentioned this subcategory highlighting what operational aspects technology- and task-related they find complex leading to poor user satisfaction.

One of them described previous experiences during analogue missions, in this case NEEMO, where crew were given to test an AR application for logistics which involved the use of the Microsoft HoloLens and said:

> *"I thought OK, I just want to have my single white bag, usually I fire the computer, I punch in a number and I get the stowage location and here with the HoloLens I need to walk through a completely over-lengthened procedure that is way too complex, takes way too long to get me there." (R02)*

The other astronaut addressed the fact that maintenance tasks are often complex and just as humans have a hard time dealing with these, technology has a difficult time as well and said:

> *"but the real pieces of equipment where you really would need help right?! Because they are complex, because they have maybe a bunch of cables and they are very close to each other and maybe they are all tangled and that's where you as a human have a difficult time, but that's also where the technology has a difficult time." (R01)*

Subcategory: Easy usage
Three respondents referred to this subcategory, the two Pangaea & Caves experts, with the role of support engineers, and one of the astronauts. All of them gave suggestions on features that would make the developed AR-IoT tool easy to use. One of the support engineers referred to the interaction component and the double confirmation procedure used for the developed application, saying:

> *"if you make that piece like very easy to use in a suit and during the operations, that's gonna add a lot of value and I think it is a very good idea to have like this double accepting." (R04)*

While describing a complex experiment involving lengthy procedures, as pdf-file on a computer, which had to be carried out during a NEEMO mission, the astronaut mentioned that an application, potentially an AR application that could guide the astronaut step by step would ease them while using procedures, explaining:

> *"a program logic guiding you step by step I think that would be helpful to avoid steps that you easily get stuck or get lost in the text or to ease us." (R02)*

Subcategory: Helpful
Three respondents addressed this subcategory, the two astronauts and one of the support engineers. One of the astronauts first mentions the difficulties related to operations encountered during Pangaea analogue missions, namely the cooperation between different instruments which is essential to later on understand exactly what the astronauts did and then expresses their opinion on how the developed AR-IoT tool can be helpful by saying:

> *"this cooperation part of your instruments, that is really difficult currently using the EVA gloves you cannot use small buttons and you don't wanna go on the small screen like on a GoPro on the bag and then touch stuff, so there it is absolutely mandatory that clever tools and I think your HoloLens approach is really really nice to come and help the astronaut in simplifying and being way more efficient." (R02)*

On a different note, one of the support engineers stresses the importance of starting now with the development of technology that shows potential in being helpful for astronaut crew, despite the current constraints by saying:

> *"if you think that will be helpful in that case [geological field exploration] it would also be sort of helpful now given the technology constraints. So right now, you have to face with these constraints, in the future you won't, but still if it will be useful for that objective, you have anyway to start now to develop it." (R03)*

Subcategory: Information in the field-of-view
This subcategory was mentioned by three respondents, one astronaut and both support engineers. One of the support engineers stressed the importance of information being in the field-of-view (FOV) more specifically in appropriate locations for the astronaut doing field work, explaining:

"you know the helmet itself doesn't turn so even if they move their head they are gonna like see inside of the helmet, so it would be best if it's in that field-of-view." (R04)

On the other hand, one of the astronauts stressed the importance of only having necessary information in the FOV, saying:

"I mean you don't want to clutter your view you don't wanna clutter all your awareness with information that is not relevant to you." (R02)

Subcategory: Integration of different systems
Three respondents addressed this subcategory, two support engineers and one astronaut. Both support engineers based on their experience with Caves & Pangaea stressed the importance of integrating different external tools, systems and subsystems in the AR-IoT tool in the future to render it useful for astronaut crew, one of them explained:

"if you want them to use it actively of course it must be at some point capable of integration with for example the EFB or whatever other system for data transmission, ground anyway has to interact with that in a proper way." (R03)

Similarly the astronaut mentioned that:

"once you are on the Moon I also believe that we will have additional tools like the spectrometers and also tools like the camera system or like some measuring devices for distance or whatever I mean like different types of instruments that all need to be connected, the data needs to flow into the system, so that people later on understand exactly what we did." (R02)

Subcategory: Simple
Two respondents, one astronaut and one support engineer, mentioned the importance of a technology or interface meant for astronauts to be simple. In fact, one of the support engineers who develops technologies for EVAs, while narrating experiences in the development of a cart for astronaut crew for EVAs, said:

"of course because it was so small and simple they really liked it." (R04)

When the astronaut was asked what is necessary for an application to be astronaut-friendly, the response was:

"it needs to be simple!" (R02)

Subcategory: Useful
Three respondents mentioned this subcategory, one astronaut and two support engineers. The astronaut believes that only in case an AR tool is essential for the task, it is useful, else annoying, explaining:

"I have tried the HoloLens which is obviously one of the professional systems that is out there and even that was annoying, I mean of course it is useful if you don't have something else and I think some applications that are especially useful of course is when you have somebody remote seeing exactly what you are seeing and being able to project things for you in real life or pointing out things for you, I mean those are all things that you cannot do in any other way." (R01)

One of the support engineers on the other hand makes a suggestion on how to increase the value of an AR application by enhancing its usefulness. More specifically, it is explained how during Pangaea missions, crew was given zoom cameras on a long stick (as bending to reach the ground is very difficult in a spacesuit) to inspect samples and how:

"if they had a tool that could like give them a close up and that they could kind of already display in that system and I think that would be very useful as well, I mean I know it's an external tool but I am just saying that a link to those additional tools would increase the value as well." (R04)

6.4.3. Category II: Effectiveness

Two respondents mentioned this subcategory, one astronaut and one support engineer. Both stressed the importance of technology leading to effective operations. The astronaut said, referring to the developed AR-IoT tool:

> *"that's exactly what we are looking for, tools that give us more autonomy and lead us to efficient and effective operations."(R02)*

Whereas the support engineer said:

> *"the final goal, is still of course improved exploration, so improve technology to support effective sampling, scientific experiments and operations on the field." (R03)*

6.4.4. Category III: Efficiency

Subcategory: Time consuming

This subcategory was mentioned by two respondents, both astronauts. These narrated about their experiences during analogue missions, namely Pangaea and NEEMO, respectively and explained their concerns regarding tasks/operations being time consuming with the technology at hand, saying:

> *"the EFB had too many steps, they are all interlinked so it was time consuming, as to go to one action one had to go through all the prior steps." (R02)*

> *"that was a little bit of rickety type of instrument and it was quite a complicated calibration and setup procedure so that was a little bit challenging and time consuming and we were, we had this procedure on the iPad however we didn't really have a good solution to like hold the iPad while we were reading we also didn't want this iPad to be on the seafloor because if you had like sand going into the seal of its case it might just damage it and then it might have leaks and you know not have an iPad the next day." (R01)*

6.4.5. Category IV: Workload

Subcategory: Hard task

Three respondents mentioned this subcategory, two astronauts and one support engineer, highlighting activities that are challenging during astronaut geological field work:

> *"it's really hard for them to say if the sample is really interesting if they look at it from a distance, and I mean you know they are tall guys trying to look on the ground, the ground will anyway be really bright and not all samples are actually interesting for scientists." (R04)*

> *"this cooperation part of your instruments, that is really difficult currently using the EVA gloves you know you cannot use small buttons and you don't wanna go on the small screen like on a GoPro on the bag and then touch stuff." (R02)*

6.4.6. Category V: Situational Awareness

This subcategory was mentioned by two respondents, one astronaut and one support engineer. The importance of the astronaut maintaining situational awareness during operations in the field was underlined by both the astronaut and the support engineer who said, respectively:

> *"the EFB on tablets during Pangaea-X was a good tool but it was a burden, a lot of interaction was required with the tablet which would make you loose awareness of the site around you. A tool like yours is definitely something we look forward to." (R02)*

> *"humans have to collaborate together with a tool capable of easy data transfer, for sampling, for operations and coordination and situational awareness." (R03)*

6.4.7. Category VI: Training

Three respondents mentioned this subcategory, two astronauts and one support engineer. The importance of geological field training for more crew autonomy in connection with innovative support tools was mentioned by one astronaut saying:

> *"we figured out that it is hugely important that the astronaut is much better trained to be autonomous for like even scientific decisions and indeed to have also the right tools obviously we are all or most of us are not the perfect scientist." (R02)*

While the other astronaut referred more specifically to potentially beneficial effects achieved through training with specifically AR applications. The astronaut mentioned that none of the applications tested led to a satisfying usage with one of the reasons being short training times, explaining that:

> *"in neither one of those cases I had the chance to like use it [MS HoloLens] for a long time. So, the question is, is there a training effect if you use it for five hours, ten hours, twenty hours, fifty hours, does it then become second nature? And then you might say, oh my gosh it's great I never want to do anything without it, it's just that I think it takes time to verify that." (R01)*

6.4.8. Category VII: User Interaction

Subcategory: Hands-free

Two respondents addressed this subcategory, more specifically the two astronauts. They both mentioned the astronaut's need to be able to work hands-free during EVAs, one of them referred to the general need to work hands-free when in the field, saying:

> *"I think every activity when you are out in the field where you need to work hands-free." (R02)*

The other one referred specifically to a NEEMO analogue mission scenario, and more generally to EVAs on the International Space Station (ISS) explaining:

> *"what seemed to work much better was to have the IV, like the person inside just reading the steps of the procedure to us so we could work hand-free just basically what people do in EVAs on the space station nowadays, I mean they have somebody read the next step to them so they don't, you know we don't carry any written reference." (R01)*

Subcategory: Voice

This subcategory was addressed by two respondents, the two astronauts. Both agree that voice e.g., speech recognition should be a primary means of interaction with modern interfaces such as the developed AR-IoT tool, explaining:

> *"ideally of course it should be voice and maybe a button just as a backup for whatever reason voice is not working or for whatever reason there is a very noisy environment in the suit and it's interfering but I think as a baseline it should be something that you don't need your hands." (R01)*

> *"voice, voice command and voice recording is definitely the solution!" (R02)*

Subcategory: Buttons/Clicker

All four interviewees mentioned this subcategory. While one of the astronauts expressed a preference for a mechanical interface such as a button, saying:

> *"I actually prefer like a more old style mechanical interface if that is absolutely robust so like there is one button on the EVA suit as an interaction that you mentioned that definitely is I think the way that the engineers would decide." (R02)*

the other one expresses concerns about the interaction being hands-free if the mechanical interface such as a button or clicker is not integrated in such a way that would ensure hands-free interaction, explaining:

> *"if you are doing something with the both of your hands what do you do? Unless it's a button that is somehow integrated on your hand itself and you just need to move a finger or something like that." (R01)*

One of the support engineers made a different suggestion, namely using a smartphone-like interface which could ease the interaction, saying:

> *"if instead of using the "clicker" button you connect it from a smartphone and you implement two buttons on the smartphone interface because these are easy to press you can for example have a button that say create site and the other saying something else such that instead of having to go into the menu and waiting until the AR button loads to create the flag." (R03)*

Subcategory: Double Confirmation
Two respondents addressed this subcategory, an astronaut and a support engineer. The support engineer mentioned that using a double confirmation approach is correct, pointing out that different combinations of inputs could have a large impact, saying:

> *"double confirmation, no but it's good it's good, I think this is a very good idea, it should be done this way and but probably there are like many options it can be clicker I don't know what else it could be but like probably there are some other options and I think by varying those options you actually make it much much better or worse." (R04)*

The astronaut suggested an alternative combination to gazing and clicking, saying:

> *"instead of the clicker, the double confirmation: gazing plus audio [voice] should also be an option." (R02)*

Subcategory: Interaction
Two respondents addressed this subcategory both of them were astronauts stressing the importance of keeping the user interaction with a tool, interface or procedure to the minimum and simple to avoid critical consequences, one of them explained:

> *"in NEEMO we had a procedure that had more than 50 pages, it had a lot interactions and so it was a complex experiment and as soon as you made one error the entire experiment was doomed." (R02)*

6.4.9. Category VIII: Implementation Aspects

Subcategory: Adaptation of technology rather than human
Two respondents mentioned this subcategory both of them were astronauts. They explained the importance of a tool that is meant for astronauts to be adapting to the astronaut rather than the astronaut having to adapt to the tool. One astronaut says,

> *"It should be follow me and I shouldn't be adapting to the tool, it's the tool that should be adapting to me." (R01)*

Subcategory: AR integrated in the helmet
Three respondents mentioned this category, one astronaut and two support engineers. They both agree that currently there are technological constraints when it comes to the integration of AR in the spacesuit such as current devices neither being certified for vacuum nor for the spacesuits which have 100% oxygen where the use of electronics can lead to fire hazards. The astronaut mentions:

> *"having this device inside the helmet would be preferred from a designer point of view so like a pilot, like a heads-up display but currently NASA is trying to avoid this for the first Artemis missions because inside the spacesuit you have 100% oxygen, you remember Apollo 1?" (R02)*

nevertheless, the technology development has to start now in fact the support engineer says:

> *"I think anyway that NASA is how to say already struggling in having a HUD, so I don't really think that for at least 5-6 years there will be a development on AR, or a usage of AR, its good anyway to have proof of concepts and proceed the development because at some point we have to be ready to embed them in the proper technology, yes totally." (R03)*

Subcategory: Easy data transfer
This subcategory was mentioned by two respondents, one astronaut and one support engineer, both agree that easy data transfer is a crucial need of modern tools and interfaces meant to aid crew during EVAs, explaining:

"humans have to collaborate together with a tool capable of easy data transfer, for sampling, for operations and coordination and situational awareness." (R03)

"to setup all the systems and to transfer data from instrument A to the EFB that also needs to be easy." (R02)

Subcategory: Operationally feasible
This subcategory was mentioned by two respondents, both support engineers. They raised the importance of operational feasibility when designing tools and interfaces for EVAs, saying:

"if you use everything together you basically don't have enough hands I mean [the] astronaut would have to hold the tools but still click on the tablet and that just, you know didn't work all the time, so he would have to put the tool away to check something in the electronic field book or like press few buttons, then take the tool back, then to take a picture with another tool they also had to click it on the tablet and they were complaining that it's not like really feasible." (R04)

"it must be also comfortable to bring on your head for prolonged sessions, for example hours, a traverse it will last at least two hours, it can be up to four hours, six hours, you have to be with power constraints, of course you can connect batteries et cetera but you can't connect, I don't know, six kg of power banks just to power the heads-up display otherwise, I mean it's unfeasible."(R03)

Subcategory: Enable remote support
Two respondents addressed this subcategory, namely the two astronauts. One of them pointed out the importance of enabling remote support, in form of "virtual colleagues" to increase the scientific outcome of the mission, through future innovative interfaces, saying:

"so this is the main point: to increase the outcome of the mission, to get more scientific data without bringing actually the stones back and also to enable virtual colleagues to walk with you in the field so and all this needs to be integrated efficiently and shouldn't lead to an overload of the astronaut." (R02)

The other one mentioned that enabling real-time remote support through an interface such as an AR application would be an irreplaceable feature, saying:

"I have tried the HoloLens which is obviously one of the professional systems that is out there and even that was annoying, I mean of course it is useful if you don't have something else and I think some applications that are especially useful of course is when you have somebody remote seeing exactly what you are seeing and being able to project things for you in real life or pointing out things for you, I mean those are all things that you cannot do in any other way." (R01)

Subcategory: Work automation and/or sharing work with off-site scientists
This subcategory has been addressed by two respondents, one astronaut and one support engineer. Both addressed the importance of relieving the astronaut crew from work which is transferred to the monitoring scientists, in fact the support engineer describes an idea that emerged during one of the Pangaea missions, saying:

"they also had an idea, what if the astronaut takes the picture, it being sent to the scientists and they give it a name, they create that folder for [the] sample and whatever, what if they just do the necessity and then either some system does it or just scientists that do it manually. So I think your system could be like the dummy scientist who does like some of the work that is repetitive in some way." (R04)

The astronaut explains how some processes should be automated to a certain extent, saying:

"if everyone starts calling and questioning me, and then I could easily get distracted and if they say like: enter this data, enter that data, so that should not be the case, so if they wanna have additional data then it needs to be an easy work flow or it should also be maybe semi-automatic." (R02)

6.4.10. Category IX: Concepts of Operations

This subcategory was mentioned by three respondents, two support engineers and one astronaut. They all addressed the new concepts of operations being currently tested during the Pangaea analogue missions enabled by modern tools, explaining:

> *"the electronic field book for example, it allowed us to test an exploration scenario, where the two astronauts work in close vicinity but actually have two different tasks to do, then you say astronaut 1 collects rock 1 and astronaut 2 collects rock 2 so those need to document, those need to describe what they see, the context to enter the data." (R02)*

> *"what we tested is the response time of humans to certain tasks, the time to reach a point in the traverse and then maybe to split, to separate, and then sampling, each astronauts sampling its own spot and then coming back together everything of course with this data sharing." (R03)*

7

Discussions and Recommendations for Future Work

In this chapter, the results will be discussed in relation to the literature outlined in chapter 1 and chapter 2. Furthermore, potential applications of the Augmented Reality (AR) Internet-of-Things (IoT) tool, the limitations of the research and design as well as recommendations for future work are presented.

Human factors design principles, standards and guidelines specifically for displays are well established in the space sector for decades already [119], [173], [181], however with the rise of modern immersive technologies such as AR, clear standards and guidelines have yet to be defined. This research project explores which usability, user interaction and implementation aspects are important for astronauts with respect to new technologies as well as for the specific use case: geological field activities. Moreover, it analyses in-depth which of these pre-established design standards are transferable to the design of immersive technologies, more specifically AR, and which have to be newly defined.

Studies involving the development of new technology are often technology-centered instead of user-centered. Based on the survey by Swan and Gabbard [97] it could be concluded that of 1104 articles on augmented reality, only roughly 3% addressed aspects of Human Computer Interaction (HCI) and only roughly 2% described a formal user-based study [97]. Swan and Gabbard [97] highlighted that in order to create effective user interfaces for emerging technologies, such as augmented reality it is crucial to understand perceptual and cognitive characteristics of the users, hence develop AR based on performed user studies [97]. A more recent study which analysed AR studies between 2005-2014 concluded that overall on average less than 10% of all the AR papers were user study papers [7]. Additionally, analysis showed that the majority of the studies included little field testing and few heuristic evaluations. That survey however, included only one AR space-related application study, namely the one by Markov-Vetter and Staadt [160]. It should be noted that most AR studies for space applications focused on the user and gathered feedback from the users through direct observations, surveys and questionnaires and in some cases performed usability tests, nevertheless, only a few AR space-related studies have been performed so far (see Table D.1 in Appendix D).

While most AR space-related studies were user-centered and frequently mention that the assessment was performed through user feedback and interviews after the user tests (see Table D.1 in Appendix D), the methodology used for data collection and analysis for these evaluations is rarely clearly outlined.

Therefore, this study not only provides additional knowledge on user-centered design AR studies in general but also and particularly on user-centered AR space-related application studies. In fact, a user-centered design methodology, which is in line with the human-centered design methodology used by NASA [173], the principles proposed by Gould and Lewis [116] and the model suggested by Lee et al. [96] was adopted for this research project. This involved expert reviews, heuristic evaluations as well as expert requirements compliance questionnaires and in-depth interviews with astronauts and astronaut geological field training experts. A clear overview of the methodology used for data collection and analysis is presented as well. This study proposes a set of heuristics principles relevant to the design of AR applications to be adopted during heuristics evaluations in the conceptual phase, it provides with customised questionnaires based on among others established space program requirements [181] and provides with insights on the methodology adopted for the qualitative analysis of the interviews.

Finally, as mentioned earlier in chapter 1, most AR applications have been tested for procedural work on the ISS such as the applications developed and studied by Braly, Nuernberger, and Young Kim [15] and the AR application by Markov-Vetter and Staadt [160], as well as mobiPV4HoloLens [126], WEKIT [34], [128], EdcAR [52], [124], [128], WEAR [146], [221], ARAMIS [72], Sidekick [175] and T2 Augmented Reality [230]. Only a few other studies investigated the benefits of AR for other applications such as the study by Furuya et al. [79] on on-board stowage operations and logistics or the study by Karasinski et al. [99] who investigated the use of a tool combining AR and IoT for just-in-time training. Only one AR application specifically for geological field exploration has been developed so far focusing entirely on navigation and traverse planning [55]. As a result, the AR-IoT tool developed for this study is the first of its kind created to support astronaut geological site inspection activities.

Therefore, the aim of this research was ultimately to develop and investigate whether an augmented reality user interface is a usable concept able to tackle issues related to future human extraterrestrial surface exploration missions. More specifically, long communication latencies which will require more crew autonomy, as real-time ground support will be unavailable. And the purpose of the investigation was particularly to determine whether an AR interface for astronaut geological field activities, more specifically site inspections, is regarded as useful and helpful as well as operationally feasible, when compared to currently used tools and media, by astronauts and astronaut geological field activities experts.

7.1. User Interface

The expert requirements compliance questionnaire results showed that the AR-IoT prototype fulfills the majority of the set requirements outlined in subsection 3.2.3. Overall opinions were positive with respect to the developed proof of concept, in fact experts commented that the application is *"promising"* as in terms of design and development it is *"a good start"* and *"on a good track"*. One expert defined the AR application as *"easier and faster"* compared to conventional media currently used e.g., tablets. This is in line with the common perception of users that hands-free displays such as the Microsoft (MS) HoloLens are preferable over handheld devices, such as tablets, due to the user's freedom of being able to use both hands for a task while having the necessary information [79], [198].

While the in-depth interviews highlighted important usability, user interaction and implementation aspects to be considered in the development of new aiding tools based on experts' past experience, from the questionnaires and associated comments, it was possible to extract crucial aspects specifically for the AR-IoT tool. More details regarding different aspects of the user interface such as layout and capabilities are described in subsection 7.1.1 and subsection 7.1.2, respectively. While usability and user interaction aspects are detailed in subsection 7.1.3 and subsection 7.1.4, respectively.

7.1.1. Layout

Regarding the layout of the user interface (UI), it resulted that the current elements are perceived as logically grouped for the purpose of use. Moreover, experts rated the interface display elements grouping and ordering as well as the colours and feedback as consistent. Specifically the icons, labels and text have been defined as clear and relevant. In fact, the text was chosen to be bright white, a colour that resulted in optimal visibility for the Holo-SEXTANT application as well [55]. It resulted that most respondents perceived that the required information to perform a certain action or task was easy to be found and easily accessible within the interface. Results also showed that the features of the application as well as the integration of the different media are mostly understandable. The sole recommendation that was made with respect to the latter aspect was, to include the displaying of recorded media such as the photos taken and eventually data resulting from the analysis of external analytical instruments. More details on recommended features are provided in subsection 7.1.2.

While some experts make suggestions regarding additional features and information that are still required by the UI to have an even more complete set of tasks for geological activities (elucidated in subsection 7.1.2), others criticise the fact that too much information is displayed and that too many windows stay open at the same time, which can result in a loss of situational awareness and affect safety. This finding, namely the importance of keeping the user's field-of-view (FOV) unobstructed to avoid safety risks, is in line with those of the Holo-SEXTANT study [55] and the study by Karasinski et al [99]. Overall it can be concluded that the amount and type of displayed information needs to be optimised. To do that, one needs to have a clear idea of what the type and sequence of activities are during operations. The author suggests to verify and validate the feasibility of the proposed concept of operations (see chapter 5) and based on the user tests, improve the layout of information. This is in line with the methodology suggested by NASA [173]. After that, it is suggested to elaborate a strategy regarding the principle of salience compatibility for information [96]: bring irrelevant information to the background or "greying it out" to minimize the amount of obstructing open windows, and adjusting the size of elements depending on their distance from the user as recommended by the experts. The importance of adjusting interface elements to the user's visual acuity by adapting their distance and size became evident in the StowageApp study by Furuya et al. [79] as well. One of the astronauts also suggested to keep the panel background of displays to the minimum to decrease the users' feeling of being overloaded by information. In the Holo-SEXTANT study this problem was tackled by rendering the panels translucent [55]. One of the experts made a valuable suggestion to minimize information cluttering, hence situational awareness impairment, by rendering windows expandable as more information is added by the user.

Additionally, colour coding was suggested by one of the experts to increase salience compatibility of information. This is a common display design principle, also suggested by Lee et al. [96], however caution is required when it comes to colour coding for validation purposes as colours used for interfaces that are meant for astronauts need to be in line with the clear meaning associated to them (yellow: caution, red: warning or emergency, green: normal state) [173], [181]. It remains unclear whether the expert was not aware of this and if aware how the actual implementation was foreseen.

Finally, suggestions for future iterations include testing the interface colours under representative environmental conditions, namely simulated lunar conditions. It should be noted that vivid blue was chosen for the interface elements to ensure maximum visibility, based on the findings by Anandapadmanaban et al. [55], who actually tested their interface outdoors, hence visibility of the interface during training on Earth should be ensured.

7.1.2. Capabilities

From the questionnaire, it could be concluded that even though experts agree with the fact that the interface only provides important data for the accomplishment of astronaut geological site evaluation tasks and that the design is task-based, several relevant missing capabilities have been identified. It became evident that the integration of, and access to, a map is considered by several experts very important for the tool to increase its usefulness in the field, which is in line with the findings by Beaton et al. [123] and Anandapadmanaban et al. [55]. Being able to see distances between sites, horizon features' names, the location of important vital assets e.g., lander position, rover position, extravehicular activity (EVA) buddy position in AR, were additional capabilities recommended by one of the experts. The displaying of distance between sites was also recommended by Anandapadmanaban et al., in fact estimating distances in a terrain with sparse landmarks, as is the case for the Moon and Mars, is especially difficult for humans without any aiding tools [55]. Furthermore, having a calculated time allowance to reach the safe haven on the basis of remaining suit consumables, consumables consumption, distance and calculated consumables consumption was suggested and considered essential data as well. This feature was also suggested by Apollo astronauts [197], Schmitt et al. [81] and explored by Johnson et al. [19] who integrated the explorer's energy consumption as metric in their traverse planning algorithm called SEXTANT. This algorithm is the one that was later implemented in Holo-SEXTANT [55]. It should be noted that the capability of the tool to show consumable levels has also been successfully integrated into the AR-IoT tool prototype I, as described in subsection 3.1.2 and subsection 3.2.2. This capability was described but was not explicitly demonstrated in the video provided for evaluation to experts, which showed concept IV (second latest version) due to Covid-19 restrictions.

One expert pointed out that the AR-IoT tool allows for the accomplishment of one of the most tiresome yet crucial tasks in terrestrial field geology which is taking site coordinates, detailed photo-documentation and descriptions of sites through the presence of GPS and a verbal “field notebook”. With respect to the integration of the site coordinates documentation action, on the basis of interview and questionnaire data though, it resulted that the majority agrees that the action should be automatised, namely coordinates should be directly transmitted from another device and recorded automatically. The IoT client-server architecture via Message Queuing Telemetry Transport (MQTT) protocol adopted for the AR-IoT tool to display dummy consumables values can certainly be expanded for GPS coordinates transmission. Nonetheless, it was pointed out that, the integration within the interface of site lithology activities, namely reporting and/or documenting what types of rocks or regolith are present as well as the documentation of structural relationships, more specifically identifying whether there are any rocks or regolith on top of each other or in a crater, are still missing. In addition, the creation and displaying of AR sample markers as well as sample location was recommended, these fall under geological sampling activities which were deliberately considered out of the scope of this first proof of concept but are certainly valuable and potential features to be integrated in the future.

Based on the interview and questionnaire data, the importance of having clear step-by-step procedures to accomplish any type of task, especially if extensive prior training was not performed by crew, has been underlined once more. One of the experts in fact suggested integrating additional cheat sheets and guidelines. As references, or so called guidelines for the tasks that are part of geological site inspections are already integrated in the interface, it is assumed that the expert stressed the importance of additional cheat sheets in relation to geological sampling activities which are recommended for future implementations.

7.1.3. Usability

From the questionnaire results it can be concluded that only certain usability aspects have been fulfilled and improvements can be made. When referring to usability, the extent to which a product can be used by specified users to achieve specified goals with efficiency, effectiveness, and satisfaction in a specified context of use is meant [92]. Where efficiency is defined as the relation between accuracy and completeness with which users can achieve a preset goal and the resources used to achieve it, and effectiveness is defined as the accuracy and completeness with which a goal is achieved [173]. Finally, satisfaction is defined as the users' comfort with respect to the use of the system as well as the users' positive attitudes toward this [173]. Overall, as mentioned in chapter 1 the way the tool will be used and whether it allows the user to do so in an effective, simple and pleasurable way is investigated. To assess this, one needs to investigate whether the target group, which in this case are astronauts and astronaut geological field activities experts, perceive the tool as useful, helpful and operationally feasible. In fact, this will be crucial to understand whether the target group would consider adopting the tool in the first place.

Efficiency

The majority of the experts remained neutral or decided not to judge the efficiency of the tool in the questionnaire. It appears that respondents found it difficult to express their opinion, as one explicitly said that it can only be evaluated by comparison. Suggested improvements included minimizing the loading bar response time. Please note again that experts based their responses on concept IV (the second latest version) and not prototype I as desired, because of Covid-19 restrictions. In prototype I, based on the feedback of the heuristic evaluations, response time upon gazing on an interface element had been reduced significantly as items being gazed at highlight immediately instead of loading. While most experts felt that it was difficult to rate the efficiency of the tool, it was a usability aspect specifically addressed by astronauts during the interviews, who narrated their experience, with time consuming thus inefficient and not satisfactory tools, during analogue missions.

Effectiveness

From the interviews, it can be deduced that effectiveness is a determining usability aspect that experts are looking for in innovative tools to accomplish one of the primary goals during EVAs, namely effective scientific operations e.g., effective sampling, as mentioned both by support engineers and astronauts. With respect to the effectiveness of the AR-IoT tool specifically, based on the questionnaire results most experts state that the interface allows for error-free operations. Note that during the heuristic evaluations the addition of a "go back" option was suggested, which is in line with one of the experts' suggestions in the questionnaire. In concept IV, the one evaluated by the experts, this feature was not included yet, however in prototype I it is. The tool can therefore be considered as potentially effective.

User satisfaction

On the basis of the in-depth interviews conducted with astronauts and analogue mission support engineers, aspects related to user satisfaction received most mentions (32% of 139 quotes). It emerged that astronauts are looking for innovative tools that are useful and helpful while not becoming a burden. According to the gathered results from the questionnaire the AR-IoT proof of concept was considered useful and helpful for future astronaut geology field training and geological field exploration activities on the Moon. Experts mentioned that the application has potential and will definitely gain additional value once certain features, as outlined in subsection 7.1.2, are implemented. Moreover, based on the interviews and questionnaire comments, it could be concluded that both astronauts and support engineers perceive that an AR tool such as the one developed is easier and faster compared to the currently used tool, namely the Electronic Field Book (EFB) on tablet which does among other things not allow for an important aspect especially mentioned by astronauts during the interviews, which is hands-free operations, discussed later on in subsection 7.1.4.

One of the astronauts mentioned how only if an AR tool is irreplaceable, it would be perceived as useful else, under current circumstances, annoying. The same astronaut narrates how most experiences with AR tools so far have been unsatisfactory with respect to implementation and hence annoying. Moreover, the respondent remains sceptical with respect to the potential of these tools in aiding crew during complex activities such as maintenance with real pieces of equipment which involve a multitude of tangled up cables and hence require high-precision AR overlays. Complexity is another aspect specifically mentioned by astronauts in relation to user satisfaction as particularly procedures that need to be followed during operations are often overlengthened, hence perceived as complex. Thus, astronauts hope for systems that are simple and easy to use as well as capable of guiding step by step, for instance, through procedural work. It should be noted that the attribute 'simple' was stressed during the interviews (6% of 139 quotes) especially by one of the astronauts and identified as an important requirement in the Holo-SEXTANT study as well [55].

Use case related usability
With respect to the tools' use case: geological field work, based on experiences with analogue missions such as NEEMO and Caves & Pangaea, majorly support engineers but also astronauts frequently mentioned (5% of 139 quotes) the importance of integrating different systems that are capable of easy data transmission into one system, such as the developed AR-IoT tool. One expert specifically mentioned how enabling the integration of the EFB used for Caves & Pangaea missions with the AR-IoT tool would increase the possibility of an active use of the system by astronauts. One of the astronauts stressed the importance of having one system, capable of gathering and merging data from the different instruments used in the field, to increase the understanding of the performed field work by the scientists later on. The same astronaut appreciated the approach of the developed AR-IoT tool in regard to the cooperation aspect as well as the integration of different media and underlined the helpfulness of the tool in simplifying operations while ensuring efficiency.

An aspect particularly related to augmented reality that has an impact on the user's satisfaction mentioned in the interviews is: the displaying of necessary information specifically in the user's field-of-view (FOV). One of the support engineers underlined the restrictions on movement of the spacesuit helmet and hence the importance of placing information in the user's FOV. Whereas one of the astronauts emphasized the importance of showing solely important data to avoid cluttering the user's view, hence impairing situational awareness. The importance of showing solely important data that is crucial to the task is a finding that has been corroborated by the Holo-SEXTANT study [55], as well as the study by Furuya et al. [79]. In fact, several extra features had been removed from the head-up display for Holo-SEXTANT [55] while a simple head-fixed panel with solely important data for the specific task was introduced for the StowageApp [79].

Situational awareness
Situational awareness is another fundamental aspect mentioned during the interviews, both by astronauts and support engineers, and confirmed by the Holo-SEXTANT study [55], [123], that needs to be ensured when developing and planning the implementation of new tools. One astronaut mentioned that during Pangaea missions, the use of the EFB on the tablet was a burden as it required a lot of interaction causing a loss in awareness. For this reason, the astronaut stated that a tool like the AR-IoT tool is something they look forward to. Despite that, rating the level of satisfaction explicitly was challenging for most respondents, actually two respondents mentioned that without wearing the HoloLens and without any prior training with the tool, it is difficult to make a statement.

Training
The training attribute is a usability aspect that was specifically mentioned during the expert interviews, especially by astronauts but also support engineers. Two astronauts stressed the importance of training with tools, one addressed the relevance of training with the 'right' tools capable of increasing crew autonomy for scientific decision-making processes, as a matter of fact enhanced crew autonomy is a major requirement for future missions to Moon, Mars and beyond [32], [153]. Whereas the other astronaut raised the interesting point of tools potentially becoming second nature and indispensable after extensive training.

Mental workload
Providing adequate system status feedback is generally a fundamental usability principle [105] and display requirement [181]. According to the questionnaire results, experts perceived the system status feedback as visible for every user action, as the majority agreed that feedback was appropriate upon the start and end of an action, and when input was required by the user. Suggested feedback enhancements by one of the experts included the use of colours to complement status information. Though, as mentioned earlier caution is required when it comes to colour coding for validation purposes. Therefore, the use of colours for additional status information is not considered necessary for the current proof of concept. Additionally, respondents rated the navigation through the interface as clear and intuitive and as mentioned earlier information was perceived to be easily found and accessible. Despite that, the respondents' opinion was divided with respect to whether the interface allows the minimization of the user's mental workload. While some experts perceived the tool as less demanding to use, as it integrates all the necessary actions for one of the most tiring but essential tasks of terrestrial field geology, specifically taking site coordinates and detailed photo-documentation and descriptions of the sites, others mentioned that an evaluation can be made and conclusions drawn only by comparison. During the interviews, support engineers and astronauts gave insights on notably hard tasks, hence tasks that require a high workload. In particular, tasks such as geological sampling e.g., identifying interesting samples and cooperating with different tools simultaneously under the spacesuit restrictions, as reported earlier already, were described. A suggestion for future tools or implementations of the AR-IoT tool made during the interviews, by both astronauts and support engineers, was to enable an automation of repetitive tasks e.g., sorting samples.

User-friendliness
Similarly, it is concluded that expressing an opinion on the user-friendliness, more specifically on the astronaut-friendliness of the tool resulted to be troublesome for most questionnaire respondents without having the chance to personally experience and interact with the tool. Nevertheless, important aspects regarding an astronaut-friendly implementation of such an AR-IoT tool have been made during the interview by both astronauts, namely ensuring tools to be "follow me", thus adapting to and seconding the astronaut instead of the other way around. Another suggestion, frequently addressed by astronauts during the interviews (5% of 139 quotes), included the fact that tools should enable remote support, through the presence of virtual colleagues on the field or expert scientists assisting operations interactively real-time.

Operational feasibility
Finally, based on questionnaire data, the developed tool was rated by the experts as operationally feasible. The importance of this aspect was addressed by the support engineers in the interviews as well. These narrated the pertinence of rendering tools operationally feasible, among other points they mentioned specifically how decisive it is for the tools to be comfortable for the user. Tools need to be comfortable not only when wearing but also when operating them. Regarding the latter, one engineer underlined once more how using a tablet during Pangaea was unfeasible, at times, as crew had too many tools to handle simultaneously. Nevertheless, ultimately, as mentioned by one of the experts, the operational feasibility will also depend on how reliably and easily the tool can be operated within the suit. Recommendations arising from the questionnaires to increase the operational feasibility include the integration of open field mapping and navigation aids. Note that, as mentioned earlier, the prominence of a tool that aids crew in navigation and traverse planning has been identified and investigated already by Anandapadmanaban et al. [55] and Beaton et al. [123]. As mentioned already this was however considered out of the scope of this research for the time being.

7.1.4. User Interaction

Aspects related to user interaction received the second highest number of mentions during the interviews after user satisfaction attributes, namely 24% of 139 quotes. While the subcategory *voice* referring to the use of voice input received most attention (9% of 139 quotes) by the two astronauts, the use of the button/clicker was addressed by all respondents (6% of 139 quotes). More details on the most suitable means of user interaction will be outlined hereafter.

Double confirmation

From the interviews and the questionnaire data it can be concluded that the novel process of "double confirmation" is considered a good strategy in fact it ensures error-free operations and gives the user more control and freedom over their actions. The double confirmation strategy can be seen as multimodal user interaction which is considered to be best especially for AR systems [90], [143] moreover the advantage in terms of error-free operations generally for human-machine communication has been corroborated by Cohen and Oviatt [190] as well. Currently, the developed AR-IoT tool permits users to avail themselves of gaze input to pick the action that needs to be performed, the chosen action is then highlighted upon selection and requires double confirmation through one simple click performed by pressing on the, so called, clicker. Opinions seem to be divided with respect to the most effective and efficient type and combination of user inputs required to accomplish an action. While some experts appreciate the current strategy of a "few gaze and clicks" others, as deduced from the interview mentions, especially astronauts stress the effectiveness and lower user efforts when using voice e.g., speech recognition.

Voice

The use of voice e.g., speech recognition has raised concerns in the past by Apollo astronauts [154] especially because of interferences in noisy environments as is the case for a space station or spacesuit. The use of voice as user input is currently still debated, however from the interviews and questionnaire data it can be concluded that the majority of the experts believes that voice is more intuitive which is in line with the findings of Cardano et al. [146], Ravagnolo et al. [139], Anandapadmanaban et al. [55] and Byrne et al. [230]. Experts confirm that as long as the voice commands are simple, hence straightforward to remember, voice input should be considered as a primary user input. In fact, in the MobiPV4Hololens study by Helin et al. [126] one issue for users was to remember the voice commands. Moreover, experts agree that having a backup option e.g., a mechanical interface like a button, in case voice input is not working, is a good option. To overcome interferences with voice inputs in noisy environments, a recommended solution to be considered is the use of bone conduction microphones [55].

Interestingly, most non-space related AR studies focused on comparing voice and gestures as user inputs or a combination of the two without considering gaze. While Chen et al. [247] concluded that voice is more accurate than gestures, whereas gestures are faster, the study by Lee et al. [152] concluded that voice is the most effective and efficient user input compared to multimodal interaction (gestures and voice) and gestures-only. Nonetheless, results from the study by Turini et al. [65] showed that all three user interactions: gaze, gestures and voice were easy and intuitive. It should be noted again that only a few non-space related AR studies focused on the user [7] and the user interaction, therefore, similar findings on voice being the most intuitive and preferred input by users are scarce. Moreover, concerns on the limited performance of voice recognition in combination with AR in noisy environments has caused voice control, as a means of interaction, to be kept to the minimum [115] or even out of consideration [40], [41]. However, this may be the case as the technology readiness level (TRL) of voice recognition technologies was low at the time of these studies. Nevertheless, according to Microsoft, voice input for AR applications, more specifically for the HoloLens, can reduce time, minimize user effort as tasks are supposed to be more fluid and reduce cognitive demand as it is intuitive. Moreover, as the range of accuracy of gazing and gestures is limited for the HoloLens, voice might be the only trusted input by the user [166]. Similarly, Gamm and Haeb-Umbach [208] performed a study of voice controlled user interfaces for consumer electronics and referred to the unique properties voice offers: straightforwardness as no mental translations are required, hands- and eye-free operations and remote control. Moreover, Gamm and Haeb-Umbach [208] state that speech input/output is often considered the most natural interface however, man-machine communication would only be really natural if the machine could meet the capabilities of a human [208]. Voice being considered the most trustworthy user input, the least mentally demanding may be the reason why it is considered by most experts the most intuitive and hence preferred means of interaction.

Button/Clicker
Some experts expressed concerns regarding the use of a button as well, because of both the usage and the integration within the spacesuit, whereas others mentioned that they would prefer such a mechanical interface. The concerns around the use of the button rely heavily upon the integration with the spacesuit, in fact the location of integration will be crucial. Initially, the idea was to integrate the button in the spacesuit gloves such that crew would only have to move their fingers guaranteeing hands-free operations, this idea was mentioned by one of the astronauts during one of the interviews as well.

Hands-free
Hands-free operations is a user interaction aspect stressed particularly by astronauts during the interviews. In fact, it is a crucial aspect especially for this use case: EVAs. Studies by Anandapadmanaban et al. [55] and Johnson et al. [19] corroborate this. The possibility to work hands-free is an important aspect during EVAs in general as suggested by Apollo astronauts [197], but also during other astronaut operations. Several space-related AR studies aimed at tools that enable hands-free operations specifically for procedural work on-board [15], [79], [126], [146], [198]. Most non-space related AR studies also stressed the importance of using voice as user input to ensure hands-free operations [18], [210].

Enabling hands-free operations specifically through voice control is a crucial aspect when in comes to controlling devices in vehicles as well. In fact, findings by Owens et al. [101] suggest that voice control may permit drivers to keep their eyes on the road longer and track their course more regularly, moreover it appears that a lower mental demand is required by drivers when compared to manual control of these devices. The study by Carter and Graham [28] found that driver performance increases when using voice control instead of manual control moreover, users rated the voice controls with feedback as easiest, most likeable, and most efficient to use while driving. Voice control appears to be the preferred means of interaction when hands-free operations are crucial to the success of operations. Similarly to astronaut extravehicular activities, in vehicle driving, voice controls appear to be the key to enhanced performance, ensured situational awareness and ease of use.

Finally, no conclusions can be drawn on the most suitable combination of inputs as a means of interaction for the double confirmation strategy, however it can be deduced that interaction has to allow hands-free operations and has to be kept simple as well as to the minimum as has been underlined specifically by astronauts during the interviews. Some experts are satisfied with gazing and clicking on a button, others suggest voice as primary input and clicking on a button as backup whereas some others suggest gaze and voice, voice and hand gestures or eye-blinking and tongue clicking. It should be noted that especially hand gestures are usually not preferred and especially unfamiliar HoloLens users have a hard time using and recalling gestures [55], [139], [146]. Note however that Microsoft mentions that HoloLens 2 features more intuitive interaction through enhanced hand tracking [164] which could be a turning point.

To summarize, with respect to the type of user input, reported downsides of the use of a button are related to the integration within the spacesuit as it can eventually hinder hands-free operations, moreover similarly to gazing it is perceived as more mentally demanding compared to the use of voice, whereas voice input is perceived as intuitive, always enabling hands-free operations and the least mentally demanding input. While the button was rated as robust and reliable, the experts that participated in the questionnaires and interviews remained sceptical on the reliability of gaze input. Nonetheless, heuristic evaluators who actually wore the HoloLens and tried the application rated the gaze input as effective, clear and user-friendly. Hand gestures are generally not preferred as these are not perceived as intuitive and user-friendly especially by unfamiliar HoloLens users. Therefore, it is recommended to first of all have the users test different input combinations for the double confirmation strategy (multimodal interaction) ensuring their reliability and eventually allowing for user input customizability. The importance of a customizable UI was a result of the Holo-SEXTANT study by Anandapadmanaban et al. [55] as well. This was also mentioned by one of the experts during expert reviews in the conceptual design phase of the AR-IoT application.

7.2. Applications of the AR-IoT Tool

The developed AR-IoT tool is a promising asset for analogue training missions that involve the simulation of extraterrestrial geological field work and in the future for lunar geological field work. Based on the positive feedback from experts at the European Space Agency's Astronaut Centre (ESA-EAC), the author envisions this AR-IoT tool to be further tested and evaluated by crewmembers and scientists in analogue environments such as Caves & Pangaea.

Furthermore, it is recommended to render the AR-IoT tool interfaceable with some of the analytical tools used for geological sampling, to then replace the Electronic Field Book (EFB) [135] on tablet, currently used during Caves & Pangaea astronaut field training, once additional usability tests and evaluations of the tool in a realistic operational scenario have been performed. In fact, these additional assessments are foreseen to add value and further prove the benefits of the tool.

Moreover, this AR-IoT tool features several desirable and recommended capabilities identified during BASALT-2 analogue missions [122, 123], namely: incorporating means to take hands-free field notes, providing suit diagnostics, overlays of virtual traverse waypoints and annotations as well as generally aiding with geological inspection tasks. Therefore, the author envisions that the developed tool could be beneficial during future BASALT analogue missions, and thus recommends the application and assessment of it during the next BASALT mission and eventually the integration with Holo-SEXTANT.

The AR-IoT tool is not only supposed to support astronauts in performing geological site inspections more autonomously but also to aid monitoring scientists that are off-site in keeping track of the operations performed. In fact, scientists could see what the astronaut is seeing and have immediate access to the data being generated and stored by the astronaut. Scientists could then, as suggested during the in-depth interviews, already start gathering, sorting and analysing the collected data, maximising the scientific understanding and outcome.

It is believed that this AR-IoT tool proof of concept, being one of the first of its kind applications, paves the way and represents an incentive for the development of further augmented reality applications for lunar and planetary EVA training in terms of both software and hardware. Several technological constraints are still present and need to be coped with when planning the implementation into operations of AR technologies, thus for this AR-IoT tool which is considered by experts to have the potential to become flight-proven to actually become it. Ideally, of course, AR should be integrated in the astronaut's suit visor. The integration of AR in the suit was frequently addressed during the interviews as well (6% of 139 quotes) by both astronauts and support engineers. Nevertheless, as emerged from the experts' comments one of the major issues is the use of electronics in a 100% oxygen environment as is the case for spacesuits presenting risks of fire hazards. A solution for the integration of AR glasses or even head-up displays (HUD) within the suit has not been found yet.

As mentioned by Anandapadmanaban et al. [55], the MS HoloLens has been designed for indoor use, hence one has to first find custom hardware solutions consisting of a physically more robust design, more powerful sensors, and hardware that is application specific. Regarding AR display requirements a power analysis needs to be performed as well [55]. The implementation of the optics in the helmet visor represents a challenge as well, considering the distance from the retinas to the display which would be on the visor. Finally, sensors and mapping cameras would have to be implemented on the exterior of the suit to enable spatial 3D mapping for world-locked content and localization purposes [55].

Despite this, as one of the support engineers mentioned during the interviews, the development of this technology has to start now such that the potential usefulness and helpfulness can be shown and to be ready to embed this technology once the constraints are no longer present.

7.3. Limitations of the Design and Recommendations for Future Work

The developed AR-IoT tool is a proof of concept specifically for astronaut geological site inspections, nevertheless site inspections are only a subset of geological field activities. A major challenge for Apollo astronauts on the Moon was navigation [37], the fact that the tool does not provide with traverse planning and navigation capabilities is therefore a limitation. Moreover, geological sampling activities are often cumbersome for astronauts that are not geologists and experts in the field, a tool such as the developed AR-IoT tool shall include capabilities that support geological sampling e.g., the integration with different scientific instruments as suggested by experts during the interviews and based on the comments from the questionnaires. Finally, as suggested by one of the experts showing remaining consumables, consumables consumption and time remaining to reach the safe haven based on distance as well as consumables consumption and remaining consumables is a capability to be considered for future implementations. While this capability has been partially integrated in the AR-IoT tool, it can be expanded, in fact only the percentage of consumables remaining is shown for now while for a real training this feature can be expanded through simulation of data or real data depending on the equipment available.

Besides the integration of different capabilities that are currently missing and could increase the value of the application significantly, some user interface layout and design as well as interaction aspects could turn out to be design limitations. These include: a missing coherent plan for button states (for more details see section 6.2), an interface colour scheme that has not been tested outdoors or in relevant environments and a combination of user inputs that is not considered optimal by the majority of the experts.

The following design improvements for the AR-IoT tool are thus recommended:

1. As suggested and planned after the heuristic evaluations: the implementation of a coherent plan for button states.
2. Testing the interface outdoors or in a relevant analogue environment to see whether the colours and brightness actually allow for sufficient and adequate visibility, to then adapt the interface colours. This was suggested by some experts as well. Despite this, note that the colour scheme applied for Holo-SEXTANT, which has been successfully tested outdoors [55], was adopted so far.
3. Test different means of interaction and user input combinations e.g., gaze and click, gaze and voice, only voice, to decide either which combination is the most feasible or to allow for customizability.
4. Expand the interface to geological sampling activities by integrating analytical tools that can generate or retrieve the mineral spectra as well as analyse it.
5. Integrate navigation and mapping capabilities.

Item 1, 2 and 3 were part of the initial plan but could not be implemented or tested due to time constraints and restricted access to the hardware and software due to Covid-19. Feature 4 and 5 are simply recommendations for future work.

Finally, current hardware limitations such as limited battery life, limited FOV and discomfort identified by the majority of the other AR space-related studies [15], [125], [139], [198], as well as partially during the heuristic evaluations, could be overcome through the migration of the application to the MS HoloLens 2. Based on the specifications of HoloLens 2 [164] major improvements with respect to comfort and FOV size have been achieved. Moreover, it should be noted that HoloLens 1 is no longer available for purchase.

7.4. Limitations of the Research and Recommendations for Future Research

Limitations of this research should be considered as well. As the analysis of qualitative data involves interpreting the study findings, it is potentially open to subjectivity and researcher's bias [184]. Various strategies, as suggested by Mays and Pope [168], have been undertaken to protect against bias and increase the reliability of the results, such as providing with a clear explanation of the data collection and analysis methodology as well as reporting illustrative verbatim quotes. Nevertheless, to minimize this bias even further and enhance the reliability of the analysis, as suggested by Mays and Pope [168], an independent assessment of the interview transcripts should be performed by additional skilled qualitative researchers as well as a comparison for agreement between the raters.

Despite the high value of the qualitative data gathered through the questionnaires and in-depth interviews with experts in the field, including ESA astronauts, participants of analogue missions and analogue mission support engineers, including geologists, EVA experts and EVA tool developers, it is believed that the collection of quantitative data can provide with additional insights and value to this research. It is therefore highly recommended to evaluate the effectiveness, efficiency and user satisfaction of the AR-IoT proof of concept through controlled experiments. Please note that this was also the initial plan before the restrictions imposed by Covid-19 on human subject testing at ESA-EAC. Extensive research and discussions with experts at ESA-EAC have been performed regarding the setup of these experiments, described in Appendix A along with a feasible concept of operations described in chapter 5 which has been partly corroborated during the interviews by the experts. In particular, regarding the idea of testing the tool with one subject only, under the assumption that the two astronauts operate separately on the field, and then measure the time to accomplish the assigned tasks.

Moreover, assessing the usability of an AR-IoT tool through a video demonstration is certainly less intuitive and more cumbersome for a user than trying the application in real life. Unfortunately, due to Covid-19 performing usability tests with human test subjects wearing the HoloLens with the application running on the device was not possible as access to the hardware was denied and human subject testing with the device not allowed. Thus, it is recommended to perform usability tests on the application prototype as soon as it will be possible again. As initially planned, two usability tests, using the System Usability Scale (SUS) proposed by Brooke [93] and complemented by Bangor, Kortum, and Miller [2] (see Figure E.1 and Figure E.2 in Appendix E) as well as the Smart Glasses Usability Satisfaction (SGUS) questionnaire (see Figure F.1 in Appendix F) proposed for the WEKIT study [139], which also used the MS HoloLens, should be performed. Usability tests are crucial when adopting a user-centered design methodology and should be done during the dry-runs before the actual controlled experiments. It is recommended to perform the usability tests with at least two subjects from the user target group, e.g. astronauts.

Another limitation of this study was that only a small number of experts, namely four, were available for an in-depth interview while a rather larger but still low number of experts, seven in total, filled in the questionnaire. Nevertheless, as can be seen in Table D.1 in Appendix D, most space-related AR studies involved small sample sizes. In fact, it is believed that, as the participants were all experts, the study can be considered representative even if the sample size is small. Despite that, future evaluations, in the form of usability tests and controlled experiments, are recommended to be performed with a larger sample size of preferably experts to increase the validity of this research.

A further source of bias may rely in the overly positive responses of the support engineers who are enthusiastic about the tool and its potential. Nevertheless, it was noted from the interviews and the comments attached to the questionnaires, that also astronauts are generally optimistic about the usefulness and helpfulness of the AR-IoT tool. However, they are generally more cautious (as seen from the number of missing responses) in rating requirements compliance as trying the tool in real life was not possible and a comparison was not either. Nonetheless, specifically one of the astronauts was generally more sceptical about AR tools due to unsatisfactory experiences in terms of implementation of AR in tools used in the past.

Finally, as was just mentioned, with respect to the questionnaire data in some cases there was a high number of missing responses, for nine questionnaire items there was a 28% of missing responses and in two cases even 42.8%. Missing responses were reasoned by the respondents of the questionnaire in the comment section as follows: "cannot evaluate" or "cannot be assessed". A pattern of missing responses was noted in two respondents specifically, in fact these two rated totally 48% and 52% of the items in the questionnaire. It should be noted that no conclusions were drawn regarding questionnaire items, which were two in total, with 42.8% missing responses as suggested by Jakobsen et al. [95].

8
Conclusions

NASA is committed to land astronauts on the Moon by 2024. The lunar exploration program - Artemis aims at using innovative technologies and systems to explore more of the Moon than has ever been done before. The goal is to establish sustainable missions by 2028, use what is learned on and around the Moon to take the next giant leap, namely sending astronauts to Mars. Activities planned during the missions of the Artemis program, especially in early phases, involve finding and using water and other critical resources needed for long-term exploration, investigating the Moon's mysteries and acquire more knowledge on Earth and the universe by carrying out experiments. All these activities involve extensive lunar geological field work which is orders of magnitude more complex than field geology on Earth. Extravehicular activities (EVAs) will become increasingly more complicated than the tasks executed during the early Artemis missions and generally during human spaceflight missions so far. EVA systems and crewmember skills that currently do not exist will be required. Moreover, the most critical foreseen issues concerning these future missions will be long communication latencies. As a consequence, ground control will not be able to support the operations step by step, also real-time monitoring and instructions will not be possible. Thus, crew will need to be more autonomous.

Therefore, modern human-machine interfaces have to be designed to support astronauts during their deep space missions. Augmented reality (AR) has found a broad range of applications in many domains including the space industry, specifically for procedural work. Several studies have been performed on the effectiveness of AR in assisting procedure-guided tasks for both on-board and on-ground operations. While AR is currently considered one of the key disruptive technologies, it is by far not clear how AR is best applied for future human lunar and planetary surface exploration. This research strove to close this gap and identify key usability and user interaction aspects, pertinent capabilities determining the adoption of AR tools, as well as provide essential insights into future human-machine interaction and design requirements for augmented reality technologies.

While some AR space-related studies recommended combining AR and IoT to enhance sensing and control of surrounding devices other non-space related studies already demonstrated an increase in work-productivity and situational awareness through the use of an AR-IoT tool. Nevertheless, only a few human spaceflight tools that combine AR and IoT technologies to increase crew situational awareness and work-productivity have been developed while a tool for EVAs specifically has neither been developed nor tested yet.

The aim of this study was to develop and investigate the usability of an AR user interface (UI) for future human lunar and planetary surface exploration EVAs, more specifically for site evaluation activities, and its potential in providing support to the crew when real-time communication with ground control is not available.

One main research question and several subquestions were posed at the beginning of this thesis. These will be answered in relation to the theory outlined and the data gathered through the experts' evaluations in this chapter.

By combining all the observations and considerations from this study, an answer to the main research question could be formulated:

"is an augmented reality user interface a usable concept to tackle long communication latencies, hence increase autonomy of astronaut crew during future human extraterrestrial surface activities on the Moon and eventually on Mars?"

The answer to the question is that the AR-IoT tool developed is a potentially usable concept capable of enhancing crew autonomy during future human extraterrestrial surface activities on the Moon and eventually on Mars.

To be able to design a tool that is usable and relevant to the use case it is designed for, one needs to understand the type of activities that one is designing for and the type of information that is useful for the user and to be provided by the user to accomplish the given tasks. An extensive user and task analysis based on literature reviews, informal discussions with experts at the European Space Agency's Astronaut Centre (ESA-EAC) and finally through in-depth interviews with experts at ESA-EAC was carried out. **Relevant activities for the chosen use case include: providing with site descriptions, performing geolocation site screenings which includes detailed photo-documentation, performing analytical screening to identify the mineral spectras and taking site coordinates**. It should be noted that the integration of analytical tools capable of geological sampling was considered out of the scope of this project, however it is highly recommended for future implementation, which is in line with experts' suggestions. In terms of **necessary information** it could be deduced that **consumable levels, time allowance to reach a safe haven, as well as navigational data** is essential for geological site activities. While providing with navigational and traverse planning information was considered out of the scope of this project, information on consumable levels are retrievable through the developed AR interface. In fact, an IoT client-server architecture via Message Queuing Telemetry Transport (MQTT) protocol has been successfully implemented allowing the AR application to interface with other surrounding devices. It is envisioned that this setup can be expanded to the use of other devices for which an integration into the AR-IoT tool is necessary.

To ensure and enhance the usability of a designed tool, the designer needs to make sure that design principles specifically relevant to the developed user interface are not violated. As mentioned earlier, for common displays human factors design principles, standards and guidelines are well established in the space sector for decades already. Nonetheless, with the emergence of new immersive technologies, relevant and applicable design principles have to be determined first. Following a user-centered design methodology which involved expert reviews, heuristic evaluations, requirements compliance questionnaires and in-depth interviews with experts at ESA-EAC, relevant design principles that are applicable to AR were ascertained and afterwards compared with prior AR space-related studies. It was concluded that **existing principles that are still relevant are: consistent, clear and relevant interface elements e.g. text, icons, labels and colours**, as well as **logical groupings for the context of use**. Providing with **appropriate user feedback** is fundamental as well as **displaying solely important task-based data to not overload the user**. Ensuring a **clear and intuitive navigation** through the interface as well as **easy access to information and error-free operations** are key usability factors. According to the majority of the experts the above mentioned principles have been fulfilled in the design of the AR-IoT tool. **Aspects that have been specifically identified as decisive for AR interfaces and in particular for the lunar and planetary EVAs are: displaying only strictly necessary information in the field-of-view (FOV) of the astronaut** such that situational awareness, thus safety are always ensured as well as **adjusting interface elements to the visual acuity of the user by adapting their distance and size**. Moreover, as stressed by astronauts AR interfaces should be as **simple and easy to use** as possible to not become a burden. **AR tools should also be following or even seconding the astronaut instead of the other way around** and enable **remote support** capabilities. Additionally, an AR-IoT tool like the one developed shall offer **integration possibilities with external tools** e.g. analytical tools and/or navigational tools, ensure **easy data transfer** as well as be operationally feasible, namely **comfortable to wear and to use**.

A major difference between an EVA and operations inside the International Space Station (ISS) is the kind of interaction an astronaut can have with an AR system. While human-computer interaction (HCI) requirements are also well established in the space sector for decades already, again with the emergence of new technologies such as AR, requirements have yet to be defined. While several studies have been performed regarding the use of AR on-board the ISS, only a few have focused on the use of AR for EVAs, more specifically for lunar and planetary surface exploration. To ensure usability, operational feasibility needs to be ensured and hence appropriate means of interaction with the tool. While it can be concluded that the novel process of **"double confirmation" is considered a good multimodal interaction strategy** as it allows for error-free operations as well as more user control and freedom over the user's actions, **no conclusions can be drawn regarding the most suitable combination of inputs as a means of interaction**. Some experts are satisfied with the idea of gazing and clicking on a button, others suggest combining voice and clicking, more specifically voice as primary input and clicking on a button as backup whereas some others suggest gaze and voice, voice and hand gestures or when technology will allow for it even eye-blinking or tongue clicking.

Generally, with respect to the type of user input, reported **downsides of the use of a button are related to the integration within the spacesuit as it can eventually hinder hands-free operations, moreover similarly to gazing it is perceived as more mentally demanding compared to the use of voice**, whereas **voice input is perceived as intuitive, always enabling hands-free operations and the least mentally demanding input**. While the **button was rated as robust and reliable**, the experts that participated in the questionnaires and interviews remained sceptical on the reliability of gaze input. Nonetheless, heuristic evaluators who actually wore the HoloLens and tried the application rated the **gaze input** as **effective, clear and user-friendly. Hand gestures are generally not preferred as these are not perceived as intuitive and user-friendly** especially by unfamiliar users. From the interviews and questionnaire data, it can be concluded that **the majority of the experts believes that voice is generally the most intuitive user input** which may be related to the fact that it is the least mentally demanding, thanks to its straightforwardness, and the most trustworthy. Other AR space-related studies corroborate this finding, namely that voice is the most intuitive means of interaction, while similar findings remain scarce for non-space related AR studies as most of these did neither focus on usability aspects nor on voice as user input. In fact, voice was at times even discarded due to the low technology readiness level of voice recognition technologies. Nevertheless, it appears that the preferred means of interaction with other devices, in case hands- and/or eye-free interaction is crucial for successful operations, is voice control e.g., control of in-vehicle systems and consumer electronics. Finally, from the interviews and questionnaires it resulted that, **user interaction has to allow hands-free operations and has to be kept simple as well as to the minimum**, as has been underlined specifically by astronauts. It should be noted that the experts that filled in the questionnaire and participated in the interviews did not have the chance to test the interaction in person and had to base their opinion on a video demonstration of the AR-IoT tool. Having the astronauts and astronaut geological field training experts actually test different input combinations for the double confirmation strategy also referred to as multimodal interaction, while wearing the HoloLens, ensuring their reliability and eventually allowing for user input customizability is definitely a recommendation for future evaluations.

Based on the research performed on the Apollo missions and EVAs in general, it became evident that workload is high during these operations. Crew workload should be minimized based on NASA's guidelines, moreover reduced workload can lead to increased user satisfaction and therefore increased usability. **No conclusions could be drawn on whether the tool allows to minimize the user's mental workload**. While some experts perceived the tool as less demanding to use as it integrates all the necessary actions for one of the most tiring but essential tasks of terrestrial field geology, specifically taking site coordinates and detailed photo-documentation and descriptions of the sites, others mentioned that an evaluation can be made and conclusions drawn solely by comparison. Despite a concept of operation and a clear operational setup for the controlled experiment had been developed in accordance with experts at ESA-EAC, tests could unfortunately not take place due to Covid-19 restrictions.

The usability of the tool depends on several factors, such as: whether the user target group, in this case astronauts, perceives the tool as useful, helpful and operationally feasible. In fact, these factors will be decisive in the adoption of the tool in the first place. **The AR-IoT tool was perceived as helpful and useful** by both astronauts and support engineers. The tool was defined as promising for geological site inspections and **potentially easier and faster compared to the currently used tools**, such as the Electronic Field Book (EFB) on tablet which does among other things not allow for an important aspect especially mentioned by astronauts in the interview, which is hands-free operations. **The tool was also defined as operationally feasible**, especially as it has the potential to replace the EFB on a tablet which was unfeasible at times during Pangaea analogues as crew had too many tools to handle simultaneously. Nevertheless, the operational feasibility has to still be tested in the field where the integration of mapping and navigation aids will be crucial. Finally, the operational feasibility will also depend on how reliably and easily the AR-IoT tool can be operated within the suit, which is yet to be determined.

A
Controlled Experiment Design

To evaluate the effectiveness, efficiency and user satisfaction of the AR-IoT concept the initial plan, before the restrictions on human-subject testing at the European Space Agency's Astronaut Centre (ESA-EAC) due to Covid-19, was to perform experiments in a laboratory environment. The experiments were supposed to be carried out in the Training Hall at ESA-EAC, in Cologne. It is recommended to carry out these tests in a later moment in time before testing the tool during an actual Caves & Pangaea expedition or similar analogue missions.

A.1. Participants

The plan initially entailed the selection of 20 participants from ESA-EAC in Cologne for the chosen operational test scenario, for more information see chapter 5. This is the minimum number to achieve statistically significant numbers according to Nielsen [106]. It is forecast that most of the participants will not be familiar with AR nevertheless, differences between subjects have to be taken into account. Participants would consist of mostly ESA staff including astronauts, astronaut geological field activities experts, past analogue mission participants and eventually Spaceship EAC and Caves and Pangaea interns. The aim was to have 10 (50%) female and 10 (50%) male participants.

A.2. Experiment Design

The scenario chosen is simple and designed such that not too many factors are involved and a fair comparison with the conventional media and tools is possible.

The concept of operation (ConOps) or operational test scenario has been explained in detail in chapter 5. This is recommended to be reviewed and verified through additional feedback from Caves & Pangaea experts besides the literature study on several analogue missions reported in chapter 5 and the in-depth interviews with experts who addressed tested ConOps. It is recommended to compare two scenarios, one that involves the use of the AR-IoT tool and the other which involves the use of a cuff-checklist as was the case during Apollo missions. The scenario involving the HoloLens is illustrated in Figure A.1 while the scenario involving conventional media e.g. the cuff-checklist is illustrated in Figure A.2.

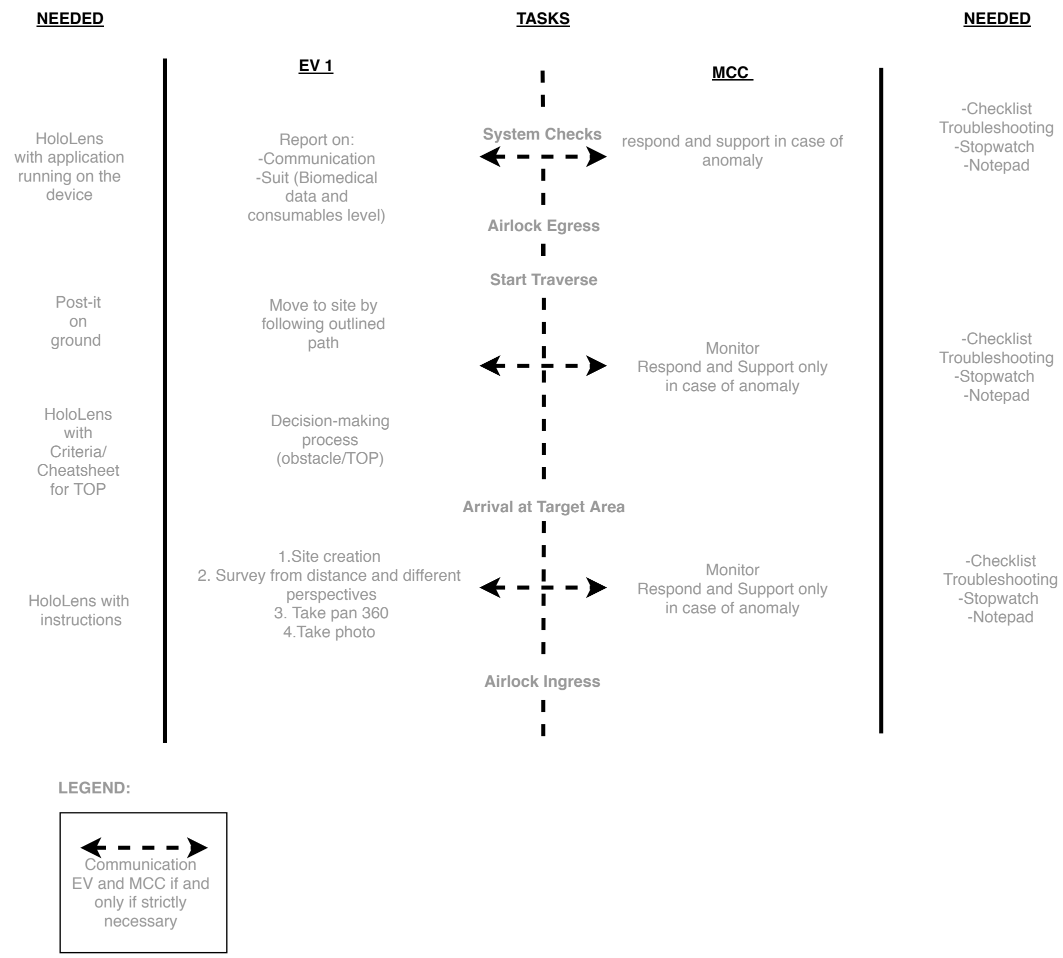


Figure A.1: Extravehicular crew (EV1) and Mission Control Centre (MCC) tasks in chronological order throughout the different EVA phases and the required instructional media respectively. It should be noted that this scenario includes the HoloLens. Note that a communication system is not displayed in this figure but required.

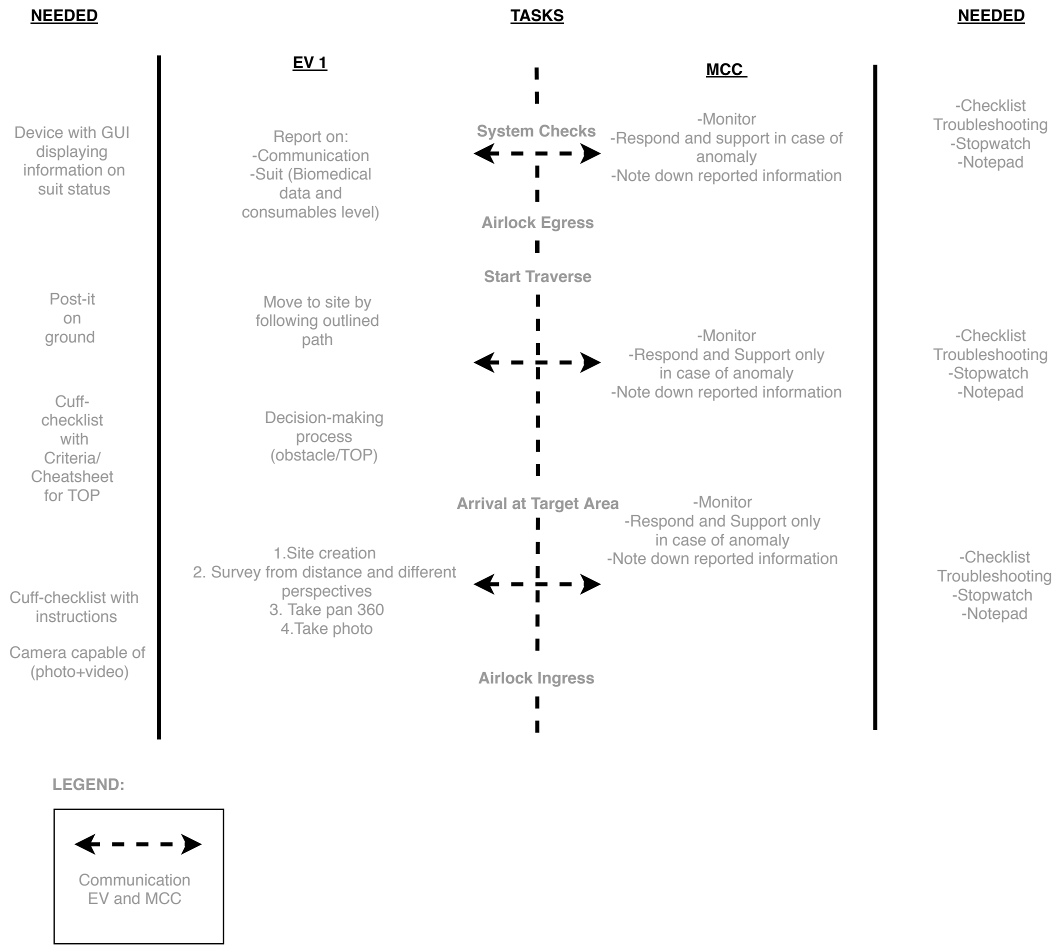


Figure A.2: Extravehicular crew (EV1) and Mission Control Centre (MCC) tasks in chronological order throughout the different EVA phases and the required instructional media respectively. It should be noted that this scenario includes the cuff-checklist. Note that a communication system is not displayed in this figure but required.

A.3. Test Equipment

The IoT network consists of a Raspberry Pi [202] which receives and manages data from several different types of sensors. It should be noted that this IoT network is modular such that different types of sensors can send data to the Raspberry Pi. In case the experiment is carried out as suggested in the Training Hall at ESA-EAC without the test subject wearing a suit, it is suggested to simulate sensor data from the IoT network, as meaningful information on for instance: suit pressure, oxygen level and water level, cannot be retrieved. The AR cues are generated using in-house software that is programmed in Unity [227] and is displayed on the Microsoft (MS) HoloLens [163]. For more information on the hardware used as well as the hardware and software integration the reader is referred to subsection 3.1.1 and subsection 3.1.2. It should be noted that in the scenario where no HoloLens is involved, a device that is capable of displaying the IoT health status data through a custom Graphical User Interface (GUI) and taking photos as well as panoramic 360 photos is required. For this it is recommended to use an Android phone as the GUI developed for these tests was specifically developed for Android devices (see subsection 3.1.2). Moreover, a communication system to talk to the simulated MCC shall be provided to all test subjects.

A.4. Instruction Methods

As mentioned earlier, the initial plan was to compare two scenarios to assess the usability of the AR-IoT tool in a realistic and feasible operational setting. For this, different instruction methods were to be provided, the innovative technology to be tested namely the AR-IoT tool and the conventional tool, more specifically the cuff-checklist. To summarise:

instruction method 1: AR-IoT tool and Mission Control Centre (MCC) as ground control.
instruction method 2: Cuff-checklist and Mission Control Centre (MCC) as ground control.

Please note that the details on the task instructions given and the guidelines are described in detail in section 5.3.

A.5. Procedure

Before the test, participants should be requested to complete a demographic questionnaire in which they are requested to mention: age, gender and whether they have experience with AR technology and/or geological field activities. After that, the test subjects should be provided with a brief task instruction described in section 5.3 and then they should be ready to start the test. Only in the case the test subject would have to perform the test using the HoloLens he or she should be briefly familiarized with the way a user can interact with the interface, namely gaze and click to confirm. They should also be informed about the fact that the application features speech recognition.

A.6. Task Instructions

Task instructions given to the test subjects should be limited. At the start of each test, participants should be told that they have to perform tasks related to geological site inspection activities for which they are given sufficient instructions to be followed. More specifically, the test subjects should be requested to:

"perform the tasks as accurate and as fast as possible while maintaining adequate situational awareness and following the instructions given by tool at hand."

In case the test involves the use of the MS HoloLens, the user should be familiarized with the tool first, especially with the hands-free interaction possibilities. A quick interaction trial by gaze and click should be performed, moreover the user should be made aware of the fact that the device avails of speech recognition. Participants should also be instructed on how to use the communication device e.g., press button A and speak.

In case the test involves the use of the cuff-checklist and communication with MCC who acts as transcriber of the information, participants should be instructed on how to use the communication device e.g., press button A and speak. Furthermore, participants should be provided with a device capable of: taking photos, panoramic 360 photos and of displaying consumable levels and biomedical data on a GUI.

After that, each participant should be requested to locate themselves at the airlock and perform system checks in communication with MCC who should provide with all required instructions to accomplish this task.

In fact, from Figure 5.1 and Figure A.2 it can be seen that the first task that crew would have to perform would be carrying out system checks. In this case crew should report to ground control whether voice communication is working nominally and in case the HoloLens is used, whether the HoloLens application is functioning nominally. These system checks do not require any knowledge by the participant as all instructions would come from ground control. Moreover, in both cases crew would be requested to report suit data consisting of biomedical data, namely suit pressure, heart rate, temperature and consumable levels more specifically water, oxygen and carbon dioxide. System checks would be performed with Mission Control Centre (MCC), this is a pre-task. It means that the time taken to accomplish this task and accuracy in carrying out this pre-task would not be considered in the results analysis.

Upon system checks completion, the test subject would be instructed to egress the airlock and start the outlined traverse by means of post-its on the ground until the marked target area would be reached where the test subject would be requested to perform all the requested site evaluation activities. All tasks should be performed with the sole use of the tools provided (HoloLens or cuff-checklist). Special attention should be taken during the traverse and interesting sites, so called targets of opportunity (TOPs) should be "logged" or reported to ground with the HoloLens or through the voice communication devices, respectively.

It should be noted that the path outlined by the post-its would have not have been visible to the test subject prior to that moment.

Once the airlock is egressed, the stopwatch would be activated. The participant would then soon reach the first TOP, quite evidently marked by a certain number of rocks placed in the middle of the path where he or she should be logging data with the HoloLens or via communication with ground, as accurately as possible, following the instructions on the HoloLens or cuff-checklist regarding the features to document. The test subject would know which option to select on the HoloLens to log the data of a target of opportunity (TOP) as it would be named accordingly. The user would have to report on:

- Number of rocks
- Colour of the rocks
- Other notable characteristics of the surrounding area

After that, the participant should continue the traverse through the outlined path to the target area where he or she was requested to perform a site evaluation activity. The participant would know which tasks to perform as they are specified under the option "site evaluation" on the HoloLens and as such in the cuff-checklist. Upon arrival at the marked target area, the participant should create a waypoint by placing a flag or a physical marker attached in the provided cuff-checklist, respectively and document the area based on the specified criteria, which are the same as for the TOP evaluation, moreover the subject would be requested to take a regular photo and a panoramic photo with the HoloLens or the given device, respectively.

The stopwatch should be stopped once the test subject signals that he or she is done with the test.

The idea behind the given task is to keep them as simple as possible to reduce the workload and task complexity for users that are completely new to the field of augmented reality.

It should be noted that the role of the IoT in this test scenario is among others marked by the presence of a predetermined traverse map, in fact in this scenario it is assumed that a lunar or planetary rover part of the internet-of-things network has previously mapped the area and the data gathered would have then been used to create a map of the area to determine interesting sites or target areas and the associated traverse. Moreover, sensors embedded in the astronaut suit and the life support system are considered part of the internet-of-things network leading to enhanced realism of the scenario. The data obtained and displayed to the participant during the test is not meant to influence the results, it is however meant to enhance realism and showcase a useful and expandable feature for later geological field training activities to the experts. In fact, during Caves

& Pangaea campaigns, carry-on "sensor boxes" containing different environmental sensors, required for instance to explore caves, are employed. It follows that integrating the information from these sensors into the AR-IoT tool would be straightforward and useful.

A.7. Guidelines

For crew members to know which kind of information is required to be collected during the different phases of the EVA, guidelines need to be established by the science team prior to the expedition.

The cuff-checklist created for the operational scenario of this research is inspired by the cuff-checklists created for the BASALT analogues [12] and Apollo cuff-checklists shown in Figure 2.3 and Figure 2.2 used by the EV crew as prompts for what descriptions they should be reporting at different stages in the EVA.

For the selected concept of operation the "dummy" observational data mentioned earlier has to be collected, namely: number of rocks, colour of the rocks and other notable characteristics of the surrounding area.

The following guidelines in form of cuff-checklist displayed in Figure A.3 are given during the test scenario employing conventional media. Note that, a physical marker was planned to be attached to the cuff-checklist. The HoloLens application avails itself of the same instructions under the site evaluation option. The same instructions of data logging of the TOPs are incorporated under the "log data" option in the AR interface. It should be noted that despite the fact the naming "log data" and "data logging of the target of opportunity" are not consistent yet, this would have been ensured before the test. A cheat sheet in this case is not necessary as the tasks requested to be performed are consciously simplified to avoid the need of prior training or briefings of inexperienced users.

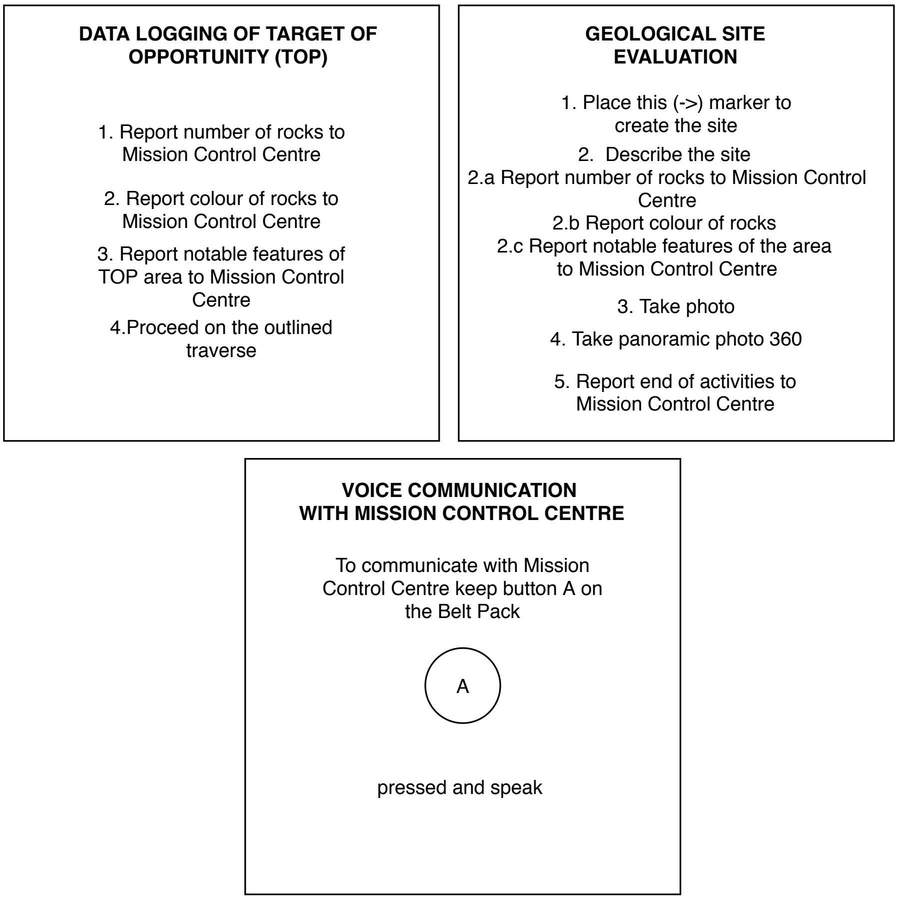


Figure A.3: Cuff-checklists provided as instructions for the TOP evaluation and the target area evaluation during the controlled experiment with conventional media.

A.8. Evaluation Metrics

From literature it was concluded that most AR space application studies focused particularly on usability, and obtained mostly qualitative results, from user feedback and direct observations, rather that quantitative results based on effectiveness and efficiency metrics such as task completion time and accuracy/error rates (see subsection G.2.2). This research was initially aimed at obtaining quantitative results as well as qualitative results. To compare the two instruction modes, objective and subjective metrics had been chosen. As mentioned in literature, objective and subjective metrics are most of the time complementary [96]. From an extensive literature study, it was concluded that to evaluate effectiveness and efficiency the most common metrics are error rates/accuracy and task completion time, respectively. For the definition of effectiveness and efficiency the reader is referred to subsection G.7.3. The target group of this application are astronauts more specifically extravehicular (EV) crew members who have to perform extravehicular activities (EVAs). As mentioned in subsection G.1.3 EVAs are of limited duration and represent a high physical and mental workload for the crew. Furthermore, the primary goal of EVAs is to maximise the scientific outcome. Therefore, time and accuracy are crucial in these activities. Hence, to assess the usability of the tool, it becomes essential to test the required time for task completion and accuracy. Finally, as part of the user-centered study method, it is crucial to evaluate user satisfaction, this is done through user rating scales. To assess the usability of the tool, the System Usability Scale (SUS) (see Figure E.1 and Figure E.2 in Appendix E) and the Smart Glasses User Satisfaction Questionnaire (SGUS) (see Figure F.1 in Appendix F) are recommended to be used after the tests. To assess workload, either the NASA TLX (see Figure G.14 in Appendix G) or the Bedford scale (see Figure G.15 in Appendix G) are recommended to be used. Workload is known to be high during EVAs (see subsection G.1.2) and should be minimized, hence it is considered a relevant metric to investigate.

A.9. Variables

The experiment's independent variables collected for the statistical analysis are as follows: subjects, instruction modes, tasks and total number of trials. 20 subjects (10 male, 10 female, mean age TBD) are recommended. There are two instruction modes, one with ground support and cuff-checklists, the other with AR-IoT tool on MS HoloLens and ground support. One task namely a geological field inspection task should be carried out. This leads to 40 trials in total. The experiment's dependent variables collected for the statistical analysis were planned as follows: task completion time in seconds, error rate/accuracy based on a wrong task completion and workload assessed through NASA TLX or Bedford scale. User rating scales were planned to be used to assess user satisfaction.

A.10. Hypothesis

The following hypothesis are to be tested:

Hypothesis 1: *When compared to conventional media, the AR-IoT tool will reduce the amount of time to complete the EVA task: a geological site inspection.*
Hypothesis 2: *When compared to conventional media, AR-IoT tool will enhance accuracy/reduce errors of the EVA task: a geological site inspection.*
Hypothesis 3: *When compared to conventional media, the AR-IoT tool will reduce cognitive load of the EVA task: a geological site inspection.*

A.11. Data Collection

The following recommendation is made for data collection, namely that during each trial demographic information such as age, gender and whether the participant has had experience with AR technology and/or geological field activities in the past is collected through the use of a questionnaire. The objective metric to assess efficiency namely task completion time is recommended to be collected through the use of a stopwatch. From the moment the test subject has completed the system checks and everything is working nominally the stopwatch should be started. The stopwatch should be stopped after the target area site inspection is completed, more specifically when the test subject clearly signals task completion. In case technical issues occur the test subject should be assisted by the Mission Control Centre (MCC) and the stopwatch should be paused until the issue is fixed and the test subject can continue nominal operations. The amount of time operations were paused should be documented. The objective metric to assess effectiveness namely accuracy/error rate is recommended to be collected in the following manner. The test subject is supposed to complete a number of tasks during the traverse and when arriving at the target area, these tasks are all specified in the AR-IoT tool and on the cuff-checklist. Every missed tasks should be accounted for as an error and every task that is not performed according to the given scheme should be accounted for as an error as well. Data on the subjective metrics such as user mental workload and user satisfaction should be collected through predefined questionnaires and user rating scales namely the Bedford scale (see Figure G.15 in Appendix G) or the NASA TLX (see Figure G.14 in Appendix G) and the System Usability Scale (SUS) (see Figure E.1 and Figure E.2 in Appendix E) as well as the Smart Glasses User Satisfaction Questionnaire (SGUS) (see Figure F.1 in Appendix F).

B
Heuristic Evaluation

Table B.1: Heuristic Evaluation Questionnaire

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree	Comments
The **system status feedback** is visible for every user action (1)						
The **response times** are appropriate to the user (1)						
The **icons** are concrete and familiar (2) [V2 10044, NASA STD-3001 vol.2]						
The user is **free to control** their actions (3)						
The user interface is **consistent** (4) [V2 10038, NASA STD-3001 vol.2]						
The user interface is **legible** (4) [V2 10039, NASA STD-3001 vol.2]						
Font size and type ensures acquisition, readability, and interpretability of the display allowing for timely and accurate processing of information. V2 10050, NASA STD-3001 vol.2>V2 10050, NASA STD-3001 vol.2						
The **prompts** are brief, unambiguous and imply the user is in control (6, 7)						
The **color and brightness** contrast is good (6)						
Instructions for use of the system are visible or easily retrievable (6, 8)						
Dialogues contain **relevant information** to the current task (8) [V2 10041, NASA STD-3001 vol.2]						
Displays and controls are **visible and easily accessible** [V2 10029, NASA STD-3001 vol.2]						
The system provides the display area to present all critical task information **within the user's Field-of-View (FoV)** [V2 10037, NASA STD-3001 vol.2]						
The displayed information is **relevant, sufficient, but not excessive**, to allow the crew to make decisions and perform the intended actions [V2 10040, NASA STD-3001 vol.2]						

Table A.1: Heuristic Evaluation Questionnaire

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree	Comments
1	The **system status feedback** is visible for every user action (1)	X					
2	The **response times** are appropriate to the user (1)					X	
3	The **icons** are concrete and familiar (2) [V2 10044, NASA STD-3001 vol.2]				X		
4	The user is **free to control** their actions (3)		X				
5	The user interface is **consistent** (4) [V2 10038, NASA STD-3001 vol.2]				X		
6	The user interface is **legible** (4) [V2 10039, NASA STD-3001 vol.2]					X	
7	**Font size and type** ensures acquisition, readability, and interpretability of the display allowing for timely and accurate processing of information. [V2 10050, NASA STD-3001 vol.2]					X	
8	The **prompts** are brief, unambiguous and imply the user is in control (6, 7)		X				
9	The **color and brightness** contrast is good (6)				X		
10	**Instructions** for use of the system are visible or easily retrievable (6, 8)		X				
11	Dialogues contain **relevant information** to the current task (8) [V2 10041, NASA STD-3001 vol.2]				X		
12	Displays and controls are **visible and easily accessible** [V2 10029, NASA STD-3001 vol.2]			X			
13	The system provides the display area to present all critical task information **within the user's Field-of-View (FoV)** [V2 10037, NASA STD-3001 vol.2]				X		
14	The displayed information is ***relevant, sufficient, but not excessive***, to allow the crew to make decisions and perform the intended actions [V2 10040, NASA STD-3001 vol.2]		X				

1. no feedback on:
- recording
- saving
- take pan: how good, when done, when saved
- how to move the flag
good feedback on the flag

2. Rapid
3. -icons good
 -change naming of the "reference"
4. - no free to control on guide references as the go backward is missing
- on where to take the photo

5. enlarge the steps of the guided procedure, introduce intermediary steps
6. OK
7. OK
8. additional ones for photo and pan 360
9. more elaborate process for references and change the name
10. OK
11. OK
12. OK
13 Loose flags, forget where
14. not always enough, e.g pan, photo

Figure B.1: Heuristic Evaluation by Evaluator 1

Table A.1: Heuristic Evaluation Questionnaire

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree	Comments
1	The **system status feedback** is visible for every user action (1)		X				
2	The **response times** are appropriate to the user (1)					X	
3	The **icons** are concrete and familiar (2) [V2 10044, NASA STD-3001 vol.2]					X	
4	The user is **free to control** their actions (3)					X	
5	The user interface is **consistent** (4) [V2 10038, NASA STD-3001 vol.2]		X				
6	The user interface is **legible** (4) [V2 10039, NASA STD-3001 vol.2]					X	
7	**Font size and type** ensures acquisition, readability, and interpretability of the display allowing for timely and accurate processing of information. [V2 10050, NASA STD-3001 vol.2]				X		
8	The **prompts** are brief, unambiguous and imply the user is in control (6, 7)		X				
9	The **color and brightness** contrast is good (6)					X	
10	**Instructions** for use of the system are visible or easily retrievable (6, 8)		X				
11	Dialogues contain **relevant information** to the current task (8) [V2 10041, NASA STD-3001 vol.2]					X	
12	Displays and controls are **visible and easily accessible** [V2 10029, NASA STD-3001 vol.2]					X	
13	The system provides the display area to present all critical task information **within the user's Field-of-View (FoV)** [V2 10037, NASA STD-3001 vol.2]			X			
14	The displayed information is ***relevant, sufficient, but not excessive***, to allow the crew to make decisions and perform the intended actions [V2 10040, NASA STD-3001 vol.2]				X		

1. -pan 360 feedback, when starts, when done missing
   - microphone feedback, when recording, when not recording missing
   - feedback on flag, drag and drop
2. OK
3.OK simple, coherent, clear
4.doing what desired
5. not 100% consistent due to loading bar. get rid of the animation
6. OK
7. feedback message "photo taken" font size bigger
8. Ok
9. OK
10. drag and drop placement instructions
11. OK
12. OK, good display proximity
13. no comment
14. not enough, info on drag and drop flag missing

Figure B.2: Heuristic Evaluation by Evaluator 2

Table A.1: Heuristic Evaluation Questionnaire

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree	Comments
1	The **system status feedback** is visible for every user action (1)					X	
2	The **response times** are appropriate to the user (1)				X		
3	The **icons** are concrete and familiar (2) [V2 10044, NASA STD-3001 vol.2]				X		
4	The user is **free to control** their actions (3)				X		
5	The user interface is **consistent** (4) [V2 10038, NASA STD-3001 vol.2]		X				
6	The user interface is **legible** (4) [V2 10039, NASA STD-3001 vol.2]					X	
7	**Font size and type** ensures acquisition, readability, and interpretability of the display allowing for timely and accurate processing of information. [V2 10050, NASA STD-3001 vol.2]					X	
8	The **prompts** are brief, unambiguous and imply the user is in control (6, 7)				X		
9	The **color and brightness** contrast is good (6)					X	
10	**Instructions** for use of the system are visible or easily retrievable (6, 8)					X	
11	Dialogues contain **relevant information** to the current task (8) [V2 10041, NASA STD-3001 vol.2]					X	
12	Displays and controls are **visible and easily accessible** [V2 10029, NASA STD-3001 vol.2]					X	
13	The system provides the display area to present all critical task information **within the user's Field-of-View (FoV)** [V2 10037, NASA STD-3001 vol.2]				X		
14	The displayed information is ***relevant, sufficient, but not excessive***, to allow the crew to make decisions and perform the intended actions [V2 10040, NASA STD-3001 vol.2]				X		

1. -No action message required "click to confirm", once the loading animation is gone, on hover change color, on click trigger action, text requires too much cognitive effort.
- have an active and a passive button state
- feedback for photo taken, replace text with other feedback, color/sound
-microphone state feedback insufficient
-what have I opened, visual reference
2. - get rid of loading animation
3. -Save icon
   - Delete icon replace with clear and put trash can
4. -References are restrictive, missing an overview of available options
5. -consistency low due to inconsistent loading animation (e.g photo)
   - font size
6. OK
7. OK
8. ambiguous prompts, e.g photo and microphone. Prompt for deletion is required to confirm the action
9. OK
10. OK
11. OK
12. OK
13. References not in FOV
14. written text not always necessary. Differentiate navigation and action type of buttons

Figure B.3: Heuristic Evaluation by Evaluator 3

C
Expert Requirements Compliance Questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Comments/Feedback Suggestions
1. The application is **useful and helpful** for future astronaut geology field training and geology field exploration activities on the Moon.						
2. The application is **operationally feasible** for future astronaut geology field training and geology field exploration activities on the Moon.						
3. Only data that is **important to mission success and significant** in terms of crew interface is provided.						
4. The overall display design is based on the geological site inspection tasks that will be performed with the display.						
4. a Specific data shown, the **display layout and groupings,** and the choice of display elements is **driven by operational requirements.**						
4. b Information is **logically grouped** according to purpose, function, or sequence of use (e.g., either a left-to-right or top-to-bottom orientation).						
4. c The display follows operational flows and allows **task completion with the minimum number of actions**.						
5. The display is **consistent** when grouping/ordering display elements.						
6. The interface elements, colours and provided feedback are **consistent.**						
7. The interface elements (e.g., text, icons, labels, objects) are **clear and relevant**.						
8. **Information density** is held to a **minimum** in displays used for						

critical geological site inspection tasks.						
9. Primary information required for performing a geological site inspection task is on a **summary display**.						
10. **Information layering**, via secondary displays or dialog boxes, is implemented to provide supplemental information in support of the primary display (e.g., specify options available to the crewmember or to provide details).						
11. The interface is designed for **efficient use of crew time** and to **minimize** crew and flight controller **training time.**						
12. The number of **user inputs** e.g. gestures/voice/gaze needed to perform simple or routine functions is **minimized.***						
13. A **help function** is accessible to the crewmembers.						
14. Display design facilitates **error-free operations.**						
15. **Data** is **protected** from inadvertent errors and hardware failures e.g. frequent saves.						
16. When a process is initiated or completed, crew member **feedback** is provided.						
17. When an input is required, an indication is provided to the crewmember, e.g., a cursor change.						

18. If the completed command implies the need for further crewmember action, the **need for action** is indicated.						
19. The application responds to crewmember interaction with **appropriate feedback.***						
20. **Navigation** through the interface is clear and intuitive.*						
21. The application minimizes the **user's mental workload.***						
22. The interface is **satisfying** for crewmembers and training members.						
23. The required information is **easily found and accessed**.						
24. It is possible to understand what the features of the application represent and to realize the **integration** of the different media.						
25. The application is "**astronaut crew-friendly"**.						
26. The application has the **potential to become "flight proven"** assuming that augmented reality technology will be an integral part of the astronauts' EVA helmet.						

*Despite the fact it is difficult to rate these aspects without wearing the HoloLens with the application running on the device, please try to rate the questionnaire statement based on your understanding of the material provided: video demo, screenshots, additional information.

D

Review of Space-related AR Studies

Table D.1: Summary of space-related AR studies and projects.
(*Please note that O and S stand for objective and subjective, respectively. N/A indicates that it is not applicable to the study/project.*)

Study/Project	Topic	Focus	Data type	AR device used	Alternative method	Study Design	Dependent measures	User rating scales	Participants
Braly, Nuernberger and Young Kim [15]	Procedural work on ISS	User	O + S	MS HoloLens	Paper manuals	Between-subjects	Time, accuracy/error rate, perceived mental workload	NASA TLX, SUS, questionnaires	20 (35% female, 65% male)
Markov-Vetter and Staadt [160]	Procedural work on ISS	User	O + S	Vuzix, WRAP920	PDF	Within-subjects	Time	NASA RTLX	10 (30% female, 70% male)
Furuya et al. [79]	Stowage operations on ISS	User	O + S	MS HoloLens	Apple iPad	Between-subjects (pilot study) Within-subjects (user study)	Time, accuracy/error rate, perceived mental workload	NASA TLX	25 (pilot study) 9 (user study)
Karasinski et al. [99]	Just-in-time training	User	O + S	MS HoloLens	N/A	N/A	Time, perceived workload	N/A	5
MOON [114]	Assembly	N/A	N/A	Handheld device	N/A	N/A	time	N/A	N/A
OnSight [175]	Rover operations	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EdcAR [223]	Procedural work	User/Technology	S	Epson Moverio BT-2000)	N/A	N/A	N/A	direct observations, user feedback	
WEKIT [128]	Procedural work	User	S	MS HoloLens	N/A	N/A	N/A	TAMARA, SGUS, SUS, QUIS, interviews	147
MobiPV [3], [126]	Procedural work on ISS	User	S	iPad/Google Glass/HoloLens	N/A	N/A	N/A	direct observations, user feedback	at least 4 astronauts
WEAR [146], [221]	Procedural work on ISS	User	S	custom built HMD	N/A	N/A	N/A	direct observations, user feedback	at least 1 astronaut
ARAMIS [72]	Maintenance, Inventory, Stowage	N/A	N/A	iPad	N/A	N/A	N/A	direct observations, user feedback	1 astronaut
Sidekick [175]	Procedural work on ISS	User	S	MS HoloLens	N/A	N/A	N/A	direct observations, user feedback	at least 1 astronaut
T2 Augmented Reality [230]	Procedural work for Maintenance	User	S	MS HoloLens	N/A	N/A	N/A	direct observations, user feedback	around 8
ARPASS [198]	Product Assurance and Safety	User/Technology	N/A	MS HoloLens	N/A	N/A	N/A	surveys, interviews	N/A
Holo-SEXTANT [55]	Planetary EVA Navigation Interface	User/Technology	N/A	MS HoloLens	N/A	N/A	N/A	user interviews and feedback (after field tests)	3

E

System Usability Scale

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	4	5
2. I found the system unnecessarily complex	1	2	3	4	5
3. I thought the system was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in this system	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8. I found the system very cumbersome to use	1	2	3	4	5
9. I felt very confident using the system	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

Figure E.1: System Usability Scale [93]

11. Overall, I would rate the user-friendliness of this product as:

☐ Worst Imaginable	☐ Awful	☐ Poor	☐ OK	☐ Good	☐ Excellent	☐ Best Imaginable

Figure E.2: Adjective Rating Scale added to the System Usability Scale [2]

F

Smart Glasses User Satisfaction (SGUS) Questionnaire

GL1 With AR-glasses I could access information at the most appropriate place and moment.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL2 Content displayed on the AR-glasses made sense in the context I used it.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL3 AR-glasses provided me with the most suitable amount of information.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL4 AR-glasses allowed a natural way to interact with information displayed.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL5 I had a good conception of what is real and what is augmented when using AR-glasses.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL6 The interaction with content on AR-glasses captivated my attention in a positive way.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL7 The instructions given by AR-glasses helped me to accomplish the task.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL8 I understood what is expected from me in each phase of the task with the help of AR-glasses.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL9 Performing the task with the help of AR-glasses was natural to me.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL10 While using AR-glasses, I was aware of the phase of the task at all times during the execution of the task.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

GL11 While using AR-glasses, I was able to pay attention to the essential aspects of the task all the time.

strongly disagree (1) | disagree (2) | somewhat disagree (3) | neither agree or disagree (4) | somewhat agree (5) | Agree (6) | strongly agree (7)

Figure F.1: Smart Glasses User Satisfaction (SGUS) Questionnaire [35]

G

Literature Study

In this chapter, first an overview of human space exploration activities is given in section G.1, then the state-of-the-art of augmented reality (AR) technology is presented in section G.2 as well as the state-of-the-art of the internet-of-things (IoT) technology in section G.3. After that, methods currently used to integrate AR and IoT technologies are investigated and presented in section G.4. Furthermore, user interface design approaches are presented in section G.5. Then, augmented reality interface design considerations are presented in section G.6. After that, evaluation and testing methodologies inherent to the user-centred design methodology are presented in section G.7.

G.1. Human Space Exploration Activities

In this section, first a general introduction on Extravehicular Activities, so called EVAs is given. Then, EVA design considerations are elaborated upon in. Third, an overview of the EVAs that have been performed on the lunar surface during the Apollo missions is given. Consequently, the information tools used during the Apollo mission EVAs, including recommendations from Apollo crewmembers, are presented. Finally, predicted issues related to the new era of human space exploration, where astronauts will go to Moon, Mars and beyond, will be elaborated upon.

G.1.1. Extravehicular Activities

EVAs are defined as any activity carried out by a pressure-suited astronaut crewmember in an unpressurized or space environment. These activities can be carried out on a space station or on the surface of another celestial body (e.g. Lunar or Martian surface). EVAs can be classified in three classes [31]:

- Scheduled
- Unscheduled
- Contingency

Scheduled activities include tasks that need to be carried out to support specified mission operations, these are activities that are part of the baseline scheduled timeline and therefore planned. An unscheduled EVA is an unplanned activity that is not included in the baseline timeline but is needed to reach mission success, mission enhancement, or to repair or override failed systems. The contingency EVA is conducted to ensure the safety of the vehicle or crew or to enable the safe return of the vehicle [31].

A further classification of EVAs can be made based on the level of complexity of the EVA, there are basic, moderately difficult, or difficult activities. Basic EVAs require the use of standard tools, restraints, or mobility aids, no special training is needed, there is little coordination between astronaut crewmembers, they involve easily accessible work sites and they do not expose crewmembers to unique hazards [31]. Moderately difficult EVAs are simple procedure-wise, however they require additional tools or equipment and skills that are not frequently used. During moderately difficult EVAs more coordination between EVA crewmembers is necessary, adjustments of existing procedures and techniques might be required as well as more in-depth training

[31]. Difficult EVAs involve a unique skillset, as they require specialized tools or mobility aids and pose access or restraint problems. Moreover, intense and/or time critical coordination between the EVA crew, the Intra-Vehicular Activity (IVA) crew and the application of a remote manipulator system operator are needed [31].

Typical EVAs in space and/or on a celestial bodies' surface include [31]:

- Mechanical support to accomplish the most important mission objectives. These include activities such as berthing, assembly, capture, deployment, positioning, and connecting or disconnecting utilities or large space structures or satellites. Furthermore, removing launch restraints and covers, deploying antennas, sensors, cameras as well as fastening (bolting or latching) together structural components.
- Maintenance and support activities which include: inspecting and replacing equipment modules or orbital replacement units, activating or deactivating experiments, retrieving samples, resupplying propellant or fluids and repairing damages.
- Transfer which involves the movement of cargo, equipment, and personnel, including the transfer of disabled crewmembers. Robotic aids can be used for these tasks.
- Experimentation with new hardware or techniques in an extravehicular (EV) environment.
- Inspection of hardware through observations or photographs.

G.1.2. Extravehicular Activity Design Considerations

There are several design considerations to be made when creating hardware for astronaut crewmembers assigned to EVAs. Guidelines to be considered when designing EVA systems and planning EVA tasks to maintain the EVA crewmembers' physiological well-being have been set. Tactile limits, eye/hand coordination, reaction time, strength capabilities, workload, food and drinking water, body waste management, medical monitoring, and atmospheric conditions are all aspects to be considered. Analysing all topics is out of the scope of this research, nevertheless relevant aspects for this research are addressed. The design considerations as outlined in NASA's "Man-System Integration Standards" include [178]:

- EVA vision is influenced by the changes in atmospheric attenuation, transmission of light via the helmet and visors, and requirements set for the visual display.
- Eye and hand coordination for an EVA crew wearing a spacesuit is altered by the limits of the specific pressure suit.
- Sensory perception and reaction time change in response to the stimuli in the space environment and the pressure suit hindrances.
- EVA workload needs to be addressed. EVAs require a high mental and physical effort, to minimize crew workload and maximize crew efficiency, planned EVAs should be performed from a preset location. Task familiarization as well as training help in reducing the workload.
- EVA medical monitoring systems should be comfortable for the EVA crew and not hamper operations. A real-time physiological monitoring system shall be provided for each EVA crewmember to measure physiological parameters. Crew should be able to detect physiological stress and/or excess while performing an EVA. Parameters to be monitored include: oxygen consumption, heart rate and Electrocardiography (EKG) signal, suit pressure and carbon dioxide pressure.
- Crew should be able to detect physiological stress and/or excess while performing an EVA. Caution and warning alarms shall be incorporated in the suit and the provision of caution and warning alarms for IVA crewmember support shall be available.
- EVA display type and location shall conform to the IVA display requirements. Labeling and color coding at EVA workstations shall conform to the IVA labeling and coding requirements and they need to be mechanically or permanently attached to the mounting surface.
- The type of EVA display is required to be suited for the task to be carried out.

- EVA displays should be able to cope with crew-imposed contact loads in all directions or be protected from these. They should be placed within the Field-of-View (FOV) allowed by the pressure suit and restraint system.
- Display readability over the range of lighting extremes expected shall be ensured. Color coding should be employed only in case adequate white illumination is available.

G.1.3. Apollo Surface Activities

The first and only time EVAs were performed on the surface of another celestial body was during the Apollo missions. An overview of the EVA duration, tasks and challenges the Apollo astronauts had to face on the lunar surface will be given.

During the first EVA of the Apollo 11 mission, on the 21st of July, in 1969, EVA crew Buzz Aldrin and Neil Armstrong spent around 2.5 hours on the lunar surface. The Apollo astronauts had to perform many different tasks in that short amount of time, including lunar sample collection, deployment of several experiments, lunar surface inspections and photographing. The first experiment was the Laser Ranging Retroreflector, followed by the Passive Seismic Experiment. The third experiment, the Lunar Dust Detector, was then mounted on the Passive Seismic Experiment. The experiments themselves were deployed without complications. Nevertheless, finding a level spot for deployment resulted in difficulties. All scientific activities were completed satisfactorily, all instruments were deployed, and samples were collected [118].

During the two EVA periods of the Apollo 12 mission, EVA crew had to perform similar tasks compared to the previous mission. Astronaut Charles Conrad and Alan L. Bean spent around 7.5 hours on the lunar surface. The Apollo 12 lunar module landed close to Surveyor 3 spacecraft, more specifically 160 meters. The Surveyor 3 had been exposed to the lunar environment 31 months before the Apollo 12 mission. The crew was assigned to collect different pieces of the Surveyor, including the TV camera and connected electrical cables, two pieces of aluminum tubing as well as the sample scoop. The aim was to analyze the spacecraft components when back on Earth to determine how they were affected by exposure to the lunar environment. The experiments this time included a cold cathode gauge, a lunar surface magnetometer, a passive seismometer, a solar wind spectrometer, a dust detector, and a suprathermal ion detector. A geological traverse, within a radius of 0.5 kilometers of the landing site was carried out as well during this mission [118].

During Apollo 14, besides playing golf, the astronauts had to perform similar activities compared to the previous Apollo missions (e.g. deploying different kind of payloads) [118]. EVA crew managed to cover a total traverse distance of 3.5 kilometers and spent a total of 9 hours on the lunar surface divided over two EVAs.

The Apollo 15 mission involved drilling activities for the emplacement of heat flow experiments (see Figure G.1 [118]) and geological traverses, assisted for the first time by the Lunar Roving Vehicle (LRV) [118]. EVA crew carried out three EVAs totaling around 18.5 hours, in fact thanks to the enhanced Portable Life Support System (PLSS) used with the astronauts' space suits, EVA crew was now able to perform longer EVAs (up to 7.5 hours).

During Apollo 16, crew spent 71 hours on the Moon and conducted three EVAs totaling around 20.3 hours on the lunar surface. These EVAs included performing lunar rover traverses of around 26.7 kilometers, collecting lunar samples at 11 different sites, deploying or performing nine experiments, and examining and photographing the lunar surface. Furthermore, astronauts had to perform repairs on the LRV, as a hammer got underneath the fender during the traverse and a part of the fender broke. The broken fender made the rover kick up dust plumes while driving, creating difficulties to the astronauts. Finally, astronaut crew managed to repair the fender during the second EVA employing clamps from the optical alignment telescope lamp and lunar maps [118].

During Apollo 17, astronaut crew spent 75 hours on the Moon, three EVAs were conducted totaling 22 hours. Surface activities included 36 kilometers of lunar rover traversing, collecting lunar samples in the Taurus-Littrow Valley at 22 different locations and deploying as well as conducting 10 scientific experiments [37]. Finally surface inspections had to be carried out and photographs had to be taken [57].

Several experiments were deployed or carried out on the lunar surface during the six Apollo lunar landing

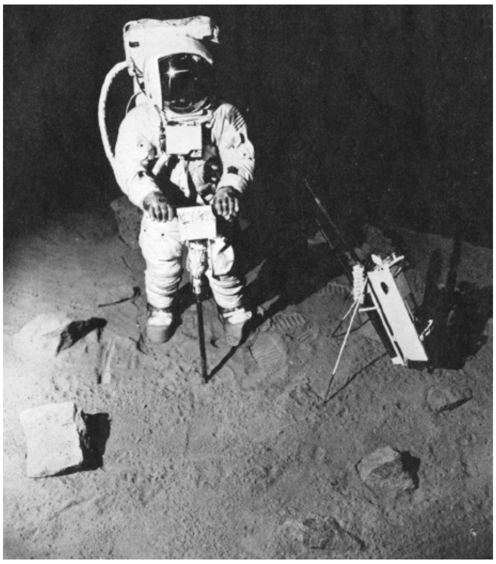


Figure G.1: Apollo Lunar surface drill [118]

missions as mentioned earlier. To minimize volume, weight and power requirements many experiments were integrated into one system, the so called Apollo Lunar Surface Experiment Package (ALSEP) (see Figure G.2 [118]). The experiments contained in the ALSEP varied over the different missions.

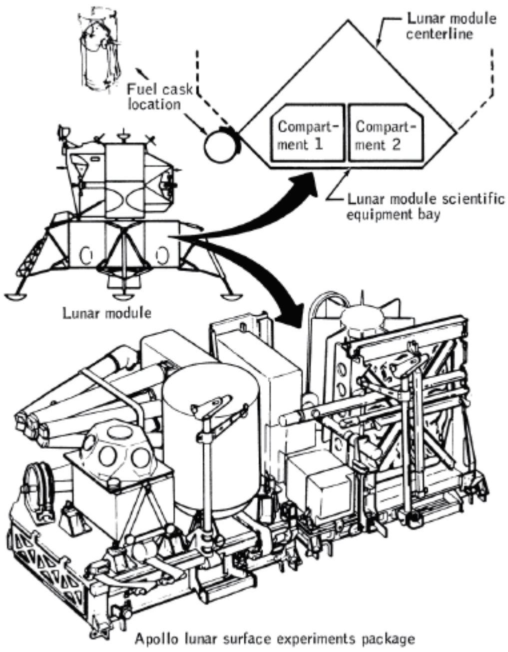


Figure G.2: Apollo Lunar Surface Experiment Package and stowage in the Lunar Module [118]

Several tools were required to carry out the surface operations during the different Apollo missions. An example of a set of tools used during the Apollo missions to carry out geological tasks on the lunar surface (see Figure G.3 [118]).

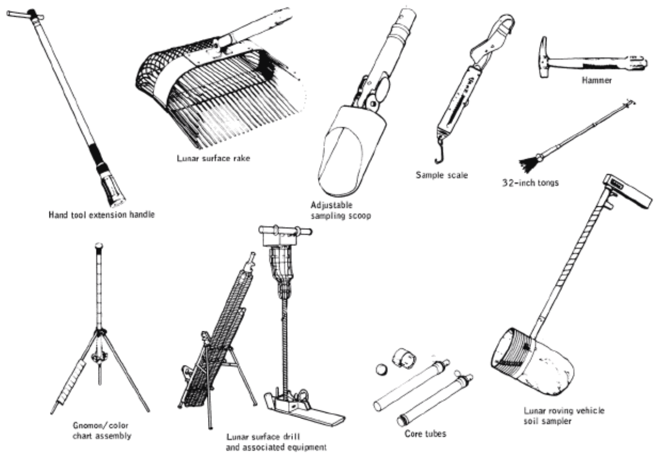


Figure G.3: Soil mechanics and lunar geology tools [118]

It can be concluded that the accomplishments Apollo astronauts reached in just a couple of days was incredible. Apollo astronauts worked at very high pace, the number of different tasks they had to carry out was very high, navigation was difficult and working with the provided tools was very exhausting [37], therefore these working conditions could neither be sustained for weeks nor for months.

G.1.4. Apollo Surface Activity Instructional Media and Recommendations

During the Apollo missions, operations and activity management were guided by procedures displayed on wrist checklists stitched on the suit (see Apollo 11 checklist in Figure G.4, [120]) and on cuff checklists (see Apollo 16 checklist in Figure G.5 [219]) on paper, furthermore guidance would be provided by ground control and other crewmembers in the crew base.

Figure G.4: Wrist Checklist Apollo 11 [120]

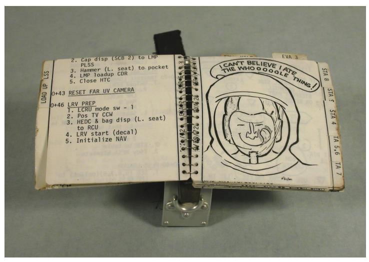


Figure G.5: Cuff Checklist EVA 2 & 3 Apollo 16 [219]

CapCom, the Mission Control Centre (MCC) in Houston, played a major role when it comes to instructional guidance, tracking, logging, monitoring, advising and scheduling. From the voice transcripts of Apollo 17, it could be concluded that EVA crew interacted with CapCom during their EVA when [37]:

- logging (reading out information)
- asking about the location of materials
- reporting descriptions on the surroundings or on the equipment status
- suggesting, requesting or documenting EVA equipment settings and usage

CapCom managed the EVA actively by [37]:

- communicating elapsed time, remaining time at a specific site, warnings when turning around
- communicating revised plans for reprioritizing, skipping or substituting work
- communicating navigation advice and identifying sites the crew is seeing

To give an example of EVA crew and CapCom communication, an excerpt of voice transcript of the Apollo 17 mission is reported below [167]. During this mission, astronaut Eugene A. Cernan was the CDR (Commander) while Robert A. Parker (Bob) was the CC (Capsule Communicator, CapCom).

> ***146:30:19*** *Cernan: Say, Bob, where can I get a new set of bags?*
> ***146:30:23*** *Parker: Okay, you want new bags...They'll be under Jack's seat.*

In fact, MCC (Mission Control Centre) could monitor EVA crew activities through the Lunar Surface TV which was either connected to the LM (Lander Module) by cable or mounted on the ground-controlled television assembly on the LRV (Lunar Roving Vehicle) and could be controlled remotely [142]. As mentioned earlier, the LRV was first used during Apollo 15 mission [142].

Nevertheless, some issues in communication with CapCom occurred with respect to verbal exchange, tracking and planning [37]. Identified issues were [37]:

- Crew mishearing or not listening to CapCom or other crewmembers
- Crew not understanding who is talking to whom
- Communication breakdowns
- Unnecessary remarks from CapCom leading to disruption
- Non-optimal scheduling decisions made by CapCom, as EVA crew was seeing opportunities CapCom could not see
- Misunderstandings between rover indicator and checklist

On the LRV, the rover used by the astronauts to extend the range of their surface EVAs, information on the heading, pitch, speed, power and temperature levels was provided through control and display modules located in front of the handle. Navigation was obtained by continuously recording direction and distance using a directional gyro and odometer and inputting this data into a computer that would check the overall direction and distance back to the lander module [50].

From interviews with Apollo astronauts that carried out EVAs, several recommendations regarding the displaying of information were retrieved. Apollo astronauts suggested that the information provided should be simple and limited to the minimum required. Most importantly, it was suggested to display only relevant information to the current operational risks and safety-related status information, which should be retrievable on a call-up basis [154].

In case of emergencies or other crucial events, alarms are preferred. Nevertheless, to relieve EVA crew and enhance their focusing capabilities, it is preferred that ground or the crew base actively monitor the situation and call issues to the attention of EVA crew. [154].

Other recommendations included visual displays to support operational tasks, including aural displays for alarms. Astronauts expressed interest in voice-activated displays, however they also expressed concerns since their use could interfere with other voice communications as mentioned earlier. The importance of having good visual and aural communication links with both the ground and the crew base was underlined [154].

Interviews were conducted with Apollo astronauts, concerning medical operations, to retrieve recommendations to improve crew health and performance for future exploration missions and lunar surface operations. Recommendations were extracted from the medical operations recommendations report concerning information displays and operations [197].

Suggestions from Apollo astronauts include the development of a reliable Heads-up Display (HUD) which displays navigation and position data on demand, consumables information and limited biomedical data. Crew suggested to display operational information only when required at any time. Crew mentioned that they did not consider necessary to know their heart rate, metabolic rate, or other physiological information during an EVA as in a continuous display, despite that being able to access this physiological information on demand would be useful [197].

Crew suggested having a navigation system available on the suit HUD for surface operations, as the Moon presents ambiguities such terrain, slopes, sun shadows, and bland environments. During Apollo EVAs it was easy to lose known points of reference and geographic orientation due to the undulating terrain. A HUD with a navigation system could save crew time [197].

G.1.5. Future Human Space Exploration Surface Activities

NASA is committed to land astronauts on the Moon by 2024 [179]. The lunar exploration program, Artemis, aims at using innovative new technologies and systems to explore more of the Moon than has ever been done before. The goal is to establish sustainable missions by 2028, use what is learned on and around the Moon to take the next giant leap, namely sending astronauts to Mars [179]. It becomes evident that the relatively simple tasks carried out during the Apollo missions of deploying experiments and collecting samples require a rather involved EVA. Lunar and planetary EVAs that involve building crew bases, laboratories and facilities are orders of magnitude more complicated than the Apollo tasks described in subsection G.1.4 [48]. EVA systems and crewmember skills that currently do not exist will be required. Activities planned during Artemis involve finding and using water and other critical resources needed for long-term exploration, investigating the Moon's mysteries and acquire more knowledge on Earth and the universe by carrying out experiments [179]. Furthermore, it will be crucial to learn how to live and operate on the surface of another celestial body where crew is days or even months away from home. Activities will also involve proving the technologies required prior to sending astronauts to Mars, where a round-trip can last up to three years [179].

G.1.6. Predicted Issues during Future Extravehicular Activities

All the challenges related to the environment of space such as vacuum, radiation, lighting conditions and thermal extremes are combined with system design constraints for EVA systems, more specifically, volume, weight, power consumption, oxygen compatibility, mobility, dexterity, and tactility restrictions. Information support and management options for EVA astronauts was restricted by all these issues during past missions. However, through a combination of training and extensive support from the Earth and from Intravehicular (IV) personnel using voice communications, these challenges were manageable. Nevertheless, future missions will render these approaches less effective and less desirable. Information transfer and management requirements will increase while opportunities for ground support and the effectiveness of training before the mission are lower [147].

One of the most critical foreseen issues regarding future EVAs will be long communication latencies. Time-delays in communication of up to 40 minutes in a two-way conversation when considering Mars will surely be challenging to cope with [37]. As a consequence, ground control will not be able to support the operations step by step, real-time monitoring and instructions will not be possible.

EVA crew's ability to take part directly in more complex mission stayed limited and relatively rudimentary until now. Information is provided to the astronauts in real-time through written checklists, verbal communication and limited capacity text displays. For future missions, a more robust capability of digital transfer and flexible data presentation, information displaying on graphic displays and a wide range of data management and command interaction is required and envisioned [147]. The communication system on the next generation interplanetary exploration suit will surely require wireless technology, with high processing speeds and good software compatibility [147].

More autonomy will be required from EVA crew to manage both systems performance and mission execution. The crew residing in the crew base will play an increasingly important role in assisting crew during EVAs, taking on activities previously performed by ground control. Hence, astronauts in the crew base need the tools and the skills to perform the EVA monitoring. Ensuring the health of the crew, efficient workflows and critical events will all be additional challenges to face. While technology can aid in coping with all these issues, technological challenges need to be taken into account as well.

G.2. Augmented Reality State-Of-The-Art

In this section, an overview of Augmented Reality (AR) and its applications will be given. After that, projects and studies involving AR specifically for space applications will be presented. Finally, future challenges related to augmented reality and its application generally and in the space industry will be discussed.

G.2.1. Overview of Augmented Reality and its Applications

Augmented Reality (AR) is a technology that superimposes computer-generated data, audio and other sensory enhancements on a user's view, in such a way that the virtual content is aligned with real-world objects, and it is possible to view and interact with it in real-time [201]. The virtual content shown comprises information that users are not able to detect directly with their own senses and it helps them performing real-world task [201]. AR is therefore able to enhance a user's perception of the surrounding world as well as the interaction with it. The first reference to augmented reality was coined by Caudell and Mizell in 1990 [36] when applying a heads-up display technology to help workers at Boeing while putting together wiring harness. Research and development in the field of augmented reality (AR) has advanced greatly in the last few decades, moving from research laboratories to consumer devices that are widely available. The progress in advanced and portable hardware, the enhanced graphics quality, the increased registration accuracy and acceptable device sizes have led to an increased adoption of this innovative technology [128].

Many industries from gaming and entertainment [5], [240], manufacturing [20], [226], education [8, 43, 44], health care [69] to the space industry [128], [160], [52] have benefited from the technological advances in augmented reality, simulation, visualization and interaction in various application areas.

The focus of this research is specifically on the space industry, however the state-of-the-art in the assembly, maintenance & repair, procedural work and training is also investigated as developments and findings in these areas are relevant and applicable for human space operations as well as became clear in section G.1.

Augmented reality has received an increasing amount of attention by researchers in the manufacturing community, as AR can be applied to solve a large range of issues throughout the assembly phase in the lifecycle of a product including planning, design, ergonomics assessment, operation guidance and training [243]. Researchers have identified the potential of AR to tackle problems arising from the extremely competitive business environment, which is pushing for innovative products at diminished time-to-market, as well as the rising trend in collaborative manufacturing environments, which requires real-time information exchanges between different stakeholders involved in the product development. Wang, Ong, and Nee [243] performed a review on AR-based assembly systems reported between 2005-2015 and found that a great amount of research has been done in the following three main areas of application: assembly guidance, AR assembly training and AR assembly, design, simulation & planning.

In the field of AR assembly training, research has largely involved procedural work. It should be noted that procedural work is the basis of all human space operations, hence special focus will be put on advancements in this field. Procedural tasks are common for installation, maintenance work and assembly, where regulations require that procedural instructions are shown on a paper medium [215], [226]. The instructions given are complex and include redundant information in different formats such as descriptive text, photos, or diagrams [110, 137, 215]. This results in the operators spending a great amount of time in studying paper instructions [209]. Therefore, several studies have been performed aimed at evaluating the effectiveness of augmented reality for procedural tasks.

Henderson and Feiner showcased potential benefits of AR applications in procedural tasks for the maintenance of armored vehicles [209]. A within-subject controlled user study investigated the use of AR for professional military mechanics. Different tasks comprising the installation and removal of fasteners, indicator lights, and connecting cables inside an armored personnel carrier turret had to be performed. Two conditions, one using AR with an untracked head-worn display with text and graphics and one using a fixed display showing laptop-based documentation, were compared during the above mentioned tasks. The condition involving AR allowed mechanics to locate tasks more quickly than when using the conventional laptop-based documentation on a fixed display. For certain tasks the AR condition resulted in reduced head movement as well [209].

Furthermore, overall the AR tool was perceived as intuitive by the users.

A similar study by Henderson and Feiner investigated the potential benefits of AR applications in procedural tasks for aircraft engine combustion chamber assembly [211]. Results showed that the AR condition led to reduced completion times, higher accuracy and user satisfaction [211].

In the study of Neumann and Majoros [226] distinct cognitive areas that AR media can complement in manufacturing and maintenance tasks were presented. Potential benefits of AR for users referring to effects in cognitive psychology in the fields of error likelihood reduction, information access, increased motivation as well as concurrent training and performance were identified. These findings have also been corroborated by others [155].

A Spatial Augmented Reality (SAR) prototype intended to be used for manual working stations of future smart factories was developed by Uva et al. [9]. Results showed that operators' performance with respect to paper manuals increases and that users accept SAR. The main advantage observed was a reduction of error rate, furthermore the effectiveness of SAR was mainly seen in difficult tasks rather than simple ones [9].

An analogous study has been done on the comparative effectiveness of augmented reality for object assembly tasks in which alternative instructional media such as a printed manual, computer assisted instructions (CAI) using a monitor-based display and CAI using a head-mounted display (HMD) were compared [18]. The aim of the study was to understand whether AR improves human performance in assembly tasks with respect to other media and what a theoretical basis for how AR interfaces might provide cognitive support and augmentation could be. The results of the study showed an 82% reduction in error rate for assembly task when applying CAI using a HMD, specifically a decreased cumulative error rate, characterised by errors due to previous assembly mistakes and a reduction in mental effort [18].

Another study confirming the benefits of AR regarding task completion times and accuracy was carried out by Iowa state University in collaboration with Boeing. In this case, a randomized study in which participants had to assemble a complex object for the first time using AR on tablets was performed [222]. Results showed that for a 46-step task of variable complexity, ranging from proper part selection to the proper alignment and fastening of bolts through various parts, the use of AR led to task completion time reduction of 30% and an error reduction that led to quality improvements of 90% [222].

It can be concluded that a great effort has been put on investigating how AR might enhance procedural work in the field of assembly, manufacturing, maintenance & repair (for surveys see [215], [243], [86]). The potential benefits regarding increased efficiency and effectiveness have been corroborated. It should be noted that in this case accuracy and task completion time are measures for efficiency and effectiveness, respectively. Additionally, AR instructions resulted in lower levels of mental workload compared with paper instructions [18].

Other interesting and relevant studies connected to augmented reality investigated the idea of 2D windows in a 3D world. Feiner et al. [205] showed the placement of 2D windows with respect to the user's head, body and objects in the nearby environment, whereas Billinghurst et al. [144] demonstrated that spatialization of information on a wearable display enhances user performance related to search time. Further work based on this concept has been performed and applied to several use cases [25]. A showcasing of a 3D navigation application which guides the user through an unknown building to a specific location demonstrated 3D signpost methods along with 2D overview maps in AR [53]. Additional work regarding other navigation possibilities including animated walkthroughs and landmark presenting has been done by Butz et al. [4]. Foyle, Andre, and Hooey [39] demonstrated methods that allow the reduction of cognitive tunneling by placing critical but non-pathfinding far from pathfinding visualizations.

Additionally, studies focusing on effective conveyal of instructive information have been performed. Tversky, Morrison, and Betrancourt [27] demonstrated that interactive animated graphics present an increased effectiveness compared to static graphics for complex system information, while Booher [78] showed that text instruction remain important when conveying procedural instructions even though pictorial cues lead to faster informational comprehension.

G.2.2. Augmented Reality Space Applications

NASA's Annual Mishap Report showed that the direct cost of NASA spaceflight accidents was almost 500 million US dollars between 2010 and 2017 [15]. Interestingly, it was found that **half of the accidents reported between 1996 and 2005 were caused by human errors and a significant portion resulted from the incorrect execution of procedures [15]**. Procedural tasks in human spaceflight operations present several challenges that could take advantage of the potential benefits of augmented reality [15]. Several augmented reality studies and projects related to human spaceflight have been carried out in the past decade, most of them were carried out in collaboration with the space agencies, NASA and ESA. To best of my knowledge all pilot studies and projects regarding AR space applications are presented hereafter.

Pilot Studies

Pilot studies that involved the design and testing of AR for space applications are presented hereafter.

A recent study carried out at the Jet Propulsion Laboratory (JPL) by Braly, Nuernberger, and Young Kim [15] was specifically aimed at investigating the benefits of AR to cope with the challenges presented by procedural work in spaceflight operations on the International Space Station (ISS). Results showed that the AR instruction method resulted in faster task completion times and lower levels of mental and temporal demand compared with paper instructions. Furthermore, when participants used the AR instruction method prior the paper instruction method, a transfer of training that improved a subsequent procedure using the paper instruction method was observed [15]. Additionally, the study concluded that the Microsoft (MS) HoloLens [163], an off-the-shelf augmented reality head-mounted display, can enhance procedural work involving simple operational tasks [15].

A similar comparative pilot study was performed in 2013, at the European Astronaut Centre (EAC), in Cologne. The pilot study compared a mobile AR system with the common method CCS (Crew Commanding Station), a laptop containing all the required information in PDF, during spaceflight procedures for intra-vehicular (IV) payload activities. The study showed that AR was preferred and decreased subjective workload [160].

A comparative ground study involving augmented reality for task guidance for stowage operations on the ISS was carried out to analyse the impact of AR on flight performance [79]. The study pointed out that stowage operations, involving the transfer of cargo items to and from different spacecraft, can take up to 60 hours while being guided real-time by ground. For this study, an application was created and deployed on the MS Hololens [163]. A between-subject study involving 25 participants was carried out in several mockups of ISS modules at NASA. Each participant had to accomplish a set of stowage tasks in a predefined time. Objective measures selected included task completion time and number of errors. An unweighted NASA TLX survey and free-form exit interviews were carried out. Inventory and Stowage Officers (ISOs) supporting flight crew were included as participants, more specifically 18 (36%). Preliminary results showed that the AR condition produced a mean task completion time of 12.1 min compared to current touchscreen instructions of around 36.75 minutes in flight (24 minutes on ground). It was concluded that results did not prove significant differences in task completion time, errors committed, or Task Load Index (TLX) responses when using the application on the MS HoloLens compared to the tablet handheld device. Nevertheless, the majority of the participants (98%) preferred the AR application over the handheld device [79].

An interesting study combining augmented reality and internet-of-things technologies for just-in-time astronaut training was performed by Karasinski et al. [99] together with NASA Ames Research Centre. A mobile prototype consisting of the MS Hololens and a network of custom internet-of-things sensors was developed. The prototype was able to show operational procedures to the user and allowed the user to interact with them in real-time. Furthermore, it allowed users to easily locate tools through precision navigation. The aim of the prototype was to enable just-in-time training. Five qualitative user tests were performed to get insights and learn about best practices that lead to a reduction in cognitive load and time saving [99]. It was concluded that augmented reality and embedded IoT sensors were an effective combination of tools capable of enhancing productivity and reducing procedural execution times. User testing results did not include numerical results or statistical data analysis [99].

A study involving the development of a Graphical-User Interface utilizing AR (GUI-AR) for space exploration scenarios in order to comprehend relevant interaction methods was performed. The positive results obtained during user evaluation sessions of this system proved that GUI-AR can be promising for other applications such as Assembly, Integration and Testing (AIT) and Assembly, Integration and Verification (AIV), on-orbit scenarios, malfunction and recovery tasks. EdcAR is a project, commissioned by ESA, connected to this study. It stands for Engineering data in cross platform AR and was aimed at using AR in different operational scenarios [52].

Projects
Projects that involved the design and testing of AR for space applications are presented hereafter.

MOON
Project MOON [114] (asseMbly Oriented authOring augmeNted reality) was developed in 2011 by Airbus Military with the aim of demonstrating that the use of AR to develop and deliver work instructions for the A400M wiring harness assembly processes leads to noticeable time reductions. For the creation of work instructions, time reductions were up to 90% (3 instead of 30 hours), work instruction consulting time reduced by 50% (1 instead of 2 hours), work instruction maintenance time was also reduced by 90% (1 instead of 10 hours) [114]. The great reduction in time is caused by the possibility of reusing information from the digital models that have been created during different design stages.

OnSight
NASA JPL and Microsoft showcased Mixed Reality technology with real-time Mars rover data. The OnSight project [175], a partnership between NASA JPL and Microsoft, recently showed how rover data combined with specially-designed motion capture systems focusing on remote experts can be combined to create a 3-D simulation of the Martian environment. Wearing the HoloLens while using OnSight, it is possible for mission scientists to meet through augmented tele-presence in order to discuss rover operations [175].

EdcAR and WEKIT
A recent study on augmented reality systems in support of astronauts' manual work support has been performed by Helin et al. [128]. This study included an evaluation study with the goal of proving that reasonable user experience of augmented reality can lead to a reduction in performance errors while executing a procedure, increasing memorability, improving cost and training time efficiency. The study includes two phases, where the first phase consists of the EdcAR-Augmented Reality for Assembly, Integration, Testing and Verification, and Operations project, led by VTT Technical Research Centre of Finland [235] through 2016 and 2017 [124]. The main objective of the EdcAR project was to verify whether AR and VR could be eligible technologies and productive tools in the aerospace industry. Further objectives included, the identification of critical technology areas in building effective AR systems (e.g. tracking, registration), creating the architecture and design of the EdcAR solution, as well as the implementation of a Proof-Of-Concept demonstrator. Identified use cases were [223]:

1. AR supported Telecom Payload Coax Cables Assembly
2. AR based On-board Training and Remote Support of Centralized Cabin Filter Replacement in ISS
3. AR based Remote Support during On-board training and Remote Support

Regarding use case 1. the following functionalities were aimed for with respect to AR:

- Semi-automatic content creation from 3D model and CVS files
- Visualizing 3D models of installed cables
- Visualizing meta-data annotated to cables
- Updating of traceability document

For use case 2, the following functionalities were aimed for with respect to AR:

- Executing lesson scripts
- Visualizing textual/graphic indications for task execution
- Visualizing dynamic graphics showing correct operations to perform

- Displaying systems/devices real-time telemetry in AR
- Checking of spare part code for correct part replacement

For use case 3, the following functionalities were aimed for with respect to AR:

- Remote working area video streaming
- Audio communication between user and remote support personnel
- Capability for the Ground Control to display in AR to the ISS crew textual/graphic indications of the task to be executed

The aim of the first phase was to observe and gather feedback from the AR system and the user. The main outcomes of the project were that the AR-googles (Epson Moverio BT-2000) used featured a processing power much lower than expected and the Android-like operation system presented issues. Additionally, the usability of the glasses was quite low, field-of-view (FOV) and latency were the main issues. Point cloud based tracking was working really well with a good tracking stability. It was concluded that using a fully Unity3D [227] based system, for instance, could show better features of AR and would also enhance system performance.

Recommendations from this study include: the visualization in AR of instructions which could be delivered by the MobiPV [3] application which will be presented later on. Concerning procedural steps to be executed, the EdcAR system could provide the crew with different types of additional information relevant to the task to be performed. Another recommendation was to connect the system to real devices and thus exploit telemetry (TM) or similar data in AR via IoT and ROS interfaces. This EdcAR capability can therefore be exploited very conveniently during the execution of Operation Data File (ODF) procedures when the operator has to monitor parameters while performing the task. Note that the ODF is the collection of procedures and reference information supporting ISS operations.

The EdcAR system could also provide real-time values of monitored parameters of nearby monitored devices so that the user does not have to jump back and forth from the setting place to the valve pressure regulator to the device displaying the telemetry for instance. This will save crew time and prevent errors during procedural task execution. Moreover, a wearable GUI is suggested to display not only telemetry but also means to set values and send commands pertinent to the ODF step in execution.

A final recommendation is that the EdcAR system should be tested in a real environment with real end-users such that ergonomic evaluations and refinements can be performed and to finally close the loop for a correct human centred design approach.

The second phase of the study included an in-situ trial and evaluation process which was developed and evaluated under the so called WEKIT-Wearable Experience for Knowledge Intensive Training project [34]. The final trial for the space pilot case entailed testing the hardware and software prototype on an actual procedure employed for astronaut training purposes. The test involved operations related to a Mars rover in the test facility, so called Mars Moon Terrain Demonstrator of ALTEC, in Torino, Italy (see Figure G.6) [139].

Figure G.6: WEKIT test subject conducting a trial [139]

The number of test participants of this study amounted to 147. Several questionnaires were used for evaluation purposes. The SSQ (Simulator Sickness Questionnaire), the TAMARA (Technology Acceptance Model for AR/WT) developed in [82], the SGUS (Smart Glasses Usability Satisfaction) developed for WEKIT, the QUIS (Questionnaire for User Interaction Satisfaction) and the TM (Transfer Mechanisms) questionnaires were used [139]. The study concluded that overall user experience was satisfactory as results of questionnaires showed that the users defined the system as interesting, easy, useful, comfortable and engaging. A System Usability Scale (SUS) was used to understand the user experience and results showed that it was acceptable with an average score of 68. Improvements have to be made to cope with the narrow FOV provided by the MS HoloLens 1, the AR device used for this project. Suggestions on how to cope with the issues related to the work space perception include: enhancing visual guidance and providing instructions and clear orientation to the task [128].

MobiPV

mobiPV [3], a mobile procedure viewer was developed for astronauts with the aim of improving efficiency and reducing the time lost to consult instructions displayed on laptop, tablet or paper. The project involved custom software run on a smartphone, connected to a wearable head-mounted camera, an audio headset and a mobile tablet when required. Astronauts were able to view procedures through either voice commands or the smartphone/tablet touchscreen. Ground control was able to monitor the astronauts performing tasks through the camera. mobiPV was successfully tested during the NEEMO 19, 20 and 21 underwater 'aquanaut' missions [170] (see Figure G.7).

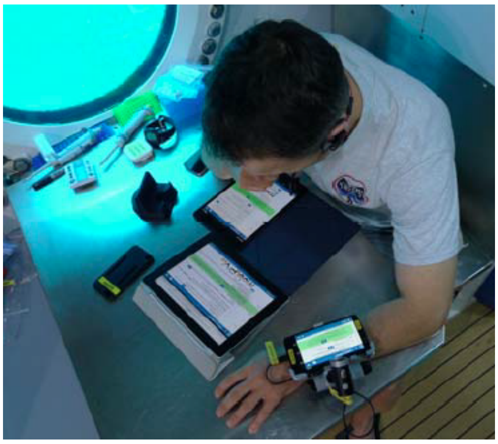


Figure G.7: Astronaut Matthias Maurer using MobiPV during NEEMO 21 [3]

NEEMO (NASA Extreme Environment Mission Operations) is a NASA analog mission which sends astronauts, engineers and scientists for up to three weeks in a row in preparation for future space exploration, to live in the Aquarius underwater laboratory, an undersea research station [170]. Additionally, it was tested during the International Space Station (ISS) mission iriss with astronaut Andreas Mogensen [158], during the Vita mission with astronaut Paolo Nespoli [61] and the Horizon mission with astronaut Alexander Gerst [60]. Tests resulted in improved work efficiency and led to saved task execution time [3].

Recommendations for future work included the capability to adapt task instructions real-time based on events and issues that occur during execution. The system could for instance inject Fault Detection, Isolation, and Recovery (FDIR) instructions into the pre-planned procedure. Another suggestion included the support in direct commanding and monitoring of small infrastructure and simple payload items through a direct interface. That capability would also enable an alternative download path to be employed for unforeseen contingencies, using internet-of-things (IoT) technologies to sense and control experiments and infrastructure [158]. The main advantage of IoT is that devices can be accessed wirelessly.

Furthermore, it is envisioned that in a future Lunar or Martian exploration scenario where ground support is limited, the use of Near Field Communication (NFC) technologies to enable tracking of items, supplies and tools will be necessary. Incorporating Artificial Intelligence (AI) is envisioned as a mean to produce tailored procedures adapted to specific situations. For instance, in case a tool is broken or missing, AI could propose an alternative tool, similarly in the case of a defect or missing system component, AI could produce tailored procedures to ease astronauts when dealing with modified configurations. Additionally, devices could contain their own procedures, instruction manuals as well as their status description and the information could be used by the AI [158]. IoT and AI technologies could enable the proposed enhancements.

Currently, the goal of new versions of mobiPV is to create a tool that allows astronauts to display procedures everywhere, at any time, on any device and in a collaborative manner.

MobiPV4Hololens [125] [127] is a MobiPV version for the HoloLens that has recently been tested and demonstrated. It had the following features:

- hands-free
- linked to mobiPV server
- working with voice commands
- using the camera on the HoloLens to capture images and videos
- capable of text-to-speech
- capable of pinning information
- featuring a collaboration mode

User tests with 5 users showed that the main benefits were the support of procedures that require both hands, the ability to see images next to the working area and the support during procedure execution from text-to-speech. However, usability/comfort issues with the HoloLens were mentioned by the users as well as the narrow FOV which cut images. Additionally, users mentioned that reading text was difficult from time to time. Demos at different VTT [235] facilities with approximately 40 users showed that the tool is working together with voice commands which were hard to remember however. The narrow FOV and usability issues were again major drawbacks pointed out.

WEAR
WEAR (WEarable Augmented Reality) is a project which was jointly funded by Space Application Services and Belgium Science Policy Office through a support technology program of ESA. The project objectives of WEAR were to support hands-free operations and offer a wearable AR system, location-based and context-sensitive information, visualization and management capabilities. WEAR is a speech interface technology (stereo 3D vision and sound), with voice commanded interface, as well as synthesized voice output and automatic item identification [146]. Crew can use voice recognition to interact with the tool and intuitive visualizations can be

displayed. Furthermore, object recognition and tracking are also possible [146].

The use cases identified were hands-free procedure viewing and manipulation, audio/video note recording as well as reference material consultation, barcode recognition, textual and graphical visualizations of information on item or subsystem on the user's field of view. A demonstration case of the WEAR tool was executed on board the ISS, crew Frank De Winne performed a real maintenance operation which was based on SODF procedures. Questionnaires, logfiles and a post-flight debriefing were used to assess the results [220].

The results obtained showed that the tool is good for a demonstrator but too bulky for an operational tool, usability in bright conditions could be improved. Nevertheless, the system offered sufficient capabilities to demonstrate the features of WEAR. Speech recognition was intuitive and reliable despite the background noise. Barcode reading was impractical but functioning, note taking, note playback and reference data visualization functioned as expected. Procedure navigation and viewing was intuitive, GUI navigation to cautions, warnings and reference data was clear. Recommendations to fully use the potential of such a tool included creating a variety of potential applications such as a smart procedure viewer, guided tour helper and training tool [220].

ARAMIS

Augmented reality technology advances the development of enhanced planning and execution methods for crew operations on-board the International Space Station (ISS) and future spacecraft. The Augmented Reality Application for Maintenance, Inventory and Stowage (ARAMIS) project had the scope of demonstrating the use of augmented reality technology to improve efficiency of operations aboard the space station leading to saved crew time that could be used for scientific research instead [176], [72]. To the best of my knowledge neither user evaluation studies were performed and nor relevant results on the effectiveness of the concept were obtained.

Sidekick

Project Sidekick [175], a collaboration between NASA and Microsoft had similar objectives. The experiment involved the use of the MS HoloLens [163]. The goal of Sidekick was to enable station crews with assistance at any time and place required. This new capability has the potential of reducing crew training requirements and increasing the efficiency at which astronauts can work in space on tasks such as science experiments, maintenance and operations. Project Sidekick was also tested during NEEMO 20 and 21 with the astronauts Luca Parmitano and Matthias Maurer, respectively. Based on video footage of Luca Parmitano during NEEMO 20, it can be seen that operations were carried successfully [109]. The task to be carried out was to retrieve an oxygen mask and manage two valves to allow for oxygen flow in the mask through external guidance. The simulated ground control would be able to see what the astronaut sees and display relevant information on the HoloLens accordingly. Neither user evaluation studies nor relevant results on the effectiveness of the concept were found regarding this project.

T2 Augmented Reality

T2 Augmented Reality [230] is a NASA Ames payload aimed at guiding crew through monthly and quarterly treadmill maintenance procedures via the Sidekick device mentioned earlier, in this case the HoloLens. Procedures were divided into single steps and overlaid onto the crew's FOV with additional assistance from guidance markers and 3D model overlays. Ground testing results showed that subjects were excited and supportive with respect to the use of AR, however it was mentioned that generally they would not wear the HoloLens if not extremely necessary. Furthermore, test subjects were worried of becoming over-reliant and tunnel focused.

Results showed that all types of overlays were effective, nevertheless every subject had unique preferences. When a choice was given between detailed and basic help to be obtained, subjects chose the detailed help. However, unnecessary details should be removed. Subjects pointed out that arrows were good for general area navigation but not for detailed/precise placement, different color overlays to identify details are suggested. Voice commands were used more frequently than gestures [230].

All the reported studies and projects so far focused on the potential benefits of AR for human space operations procedural work, specifically for ISS operations. The following project reported is not strictly related to human spaceflight operations, nevertheless addressed to the space industry [230].

ARPASS
The ARPASS project (Augmented Reality in Product Assurance & Safety Study) involves a recent study commissioned by ESA aimed at assessing the maturity and business value of augmented reality with respect to space product assurance and safety activities applications [198]. As the majority of the errors in industrial manufacturing as well as maintenance activities are due to procedural errors and several errors are due to negligence [226] it is believed that augmented reality has the potential of providing the necessary information with adequate quality to the required location and on time to prevent these errors [75]. Interviews and surveys were conducted involving product assurance and safety experts as well as a review of requirements of space industry and component technology readiness levels (TRLs). The feasibility study concluded that the TRL of certain AR components does not yet fully satisfy requirements set by the space industry however, AR presents great potential for long-term benefits [198]. The study moreover did not include use cases involving extraterrestrial exploration, nevertheless it is forecast that AR could provide great support for maintenance operations on Mars such that preparatory training time can be reduced and communication delays can be mitigated allowing a great autonomy for astronauts [198].

Holo-SEXTANT
Holo-SEXTANT is an AR tool that aims at helping extravehicular (EV) crewmembers navigating a planned traverse during the extravehicular activities (EVAs). SEXTANT stands for Surface Exploration Traverse Analysis and Navigational Tool and it brings together three components: an environmental model, more specifically terrain data, an explorer model with energetic cost functions and a solver or path planner [108]. This tool was developed and iterated upon by Massachusetts Institute of Technology (MIT) graduate students [108] till it was integrated in the MS HoloLens during the Biologic Analog Science Associated with Lava Terrains (BASALT) exploration field campaign which took place in November 2017, at Hawai'i Volcanoes National Park [55].
Holo-SEXTANT presents a modern approach that permits traverse plans to be superimposed on the terrain in the display, facilitating the user to see their surroundings while always being aware of the location of the path. It computes the most efficient route between user-specified waypoints across a terrain surface through the minimization of a cost function that is based on different factors such as terrain slope or projected metabolic expenditure [123]. The system relies on geographic coordinates of the path as an input, allowing for the display of any arbitrary traverse path and hence providing with geolocation awareness. The user can interact with the tool by voice control and the tool comprises real-time information displays pertinent to the user's location. During the BASALT campaign, the system was tested in hazardous terrain where natural or visually evident paths did not exist, rendering the Holo-SEXTANT crucial. It is argued that Holo-SEXTANT is a viable solution for navigation.

For the Holo-SEXTANT BASALT campaign use case, the user had to perform crucial tasks with their hands, implying that there was a need for a hands-free interface. Therefore, Holo-SEXTANT uses voice commands, this allowed the user to actively provide input to the tool without having to worry neither about focusing on cursor position accuracy nor on turning their head. The hypothesis that voice commands are more intuitive to use was enforced during the BASALT campaign.

From testing the tool during the BASALT campaign it was concluded that regarding the HoloLens limitations in tracking, mapping, and robustness, the device had sufficient hardware capabilities for this proof of concept navigation assistant. While other limitations, more specifically re-calibration, drift, and error did hinder the EVA experience [55].

During the BASALT-3 campaign, a set of subjective and objective technology impact metrics was considered [123]: simulation quality, capability assessment, acceptability, usability, workload and the quantitative usage of data. The simulation quality represents a measure of simulation fidelity with respect to the implementation of the capability, capability assessment represents a measure of the capability's potential to improve and enable future extravehicular activities (EVAs). Acceptability is a measure for the capability's suitability for future EVAs. Usability evaluates user efficiency, effectiveness, and satisfaction. Workload assesses the user's physical and mental effort. Finally, the quantitative usage data aimed at assessing which features were most- and least-used.

Finding key metrics to test the efficacy of a navigation interface was found to be cumbersome. In fact, paths could not be repeated by the same user as the user would remember the traverse. Additionally, the speed of crew members traversing showed a large spread. An obvious choice might appear to be a comparison between the planned path with the traversed path, then again following the path exactly is unfeasible and unnecessary to the goal of the EVA. Moreover, one of the goals of Holo-SEXTANT is to enable EVA crew to safely and willfully deviate from the path to explore in case it is desired and to return to the path easily. It was concluded that further investigation would be needed to identify the best metrics to test such an AR navigation interface.

To test the interface user interviews were performed to gather subjective feedback. Following the different needs and preferences of each EV crewmember depending on the tasks they were assigned, it became crucial to develop a highly customizable user interface (UI). Any UI component can be hidden if the user wishes so. All the UI components were directly manipulatable allowing the user to design their own "workspace" on the go.

Overall, it was concluded that the tool is significantly enhancing EVAs. Recommendations for future applications included an AR assistant able to aid during sample identification, providing suit diagnostics, aiding with tasks, taking notes, communicating with IV or mission control and more. Moreover, such an EVA tool could have integrated intelligence going beyond simple voice commands. Finally, an artificially intelligent assistant could be included to aid in decision making processes.

Studies on the required road-map, requirements, system capabilities and available technology to allow humans to embark on a new era of human space exploration aimed at sending crewed spacecraft beyond Low Earth Orbit have been performed. Humans will be facing new environments with high-levels of autonomy and human-machine interfaces including information management applications together with advanced displays and vision systems will be present. The goal is that these systems allow crew to access and understand the data from the on-board systems, obtain the capability to provide an overview and control of autonomous, automated and manual functions, while keeping cognitive overload during off-nominal or critical events to the minimum and situational awareness to the maximum [32]. On-board training, critical skills enhancement and space-based maintenance will be critical capabilities as well [32]. Based on the results of the studies and projects performed it is believed that immersive technologies, more specifically AR could ensure the fulfillment of these requirements and capabilities.

An overview of all space-related AR studies can be found in Table D.1.

G.2.3. Challenges of Augmented Reality Applications

Despite the great potential of AR, there are technical issues to be solved. One great feature of augmented reality is the ability to place digital objects in the 3D real-world space. The final goal of several AR applications is marker-less tracking. Marker-less AR is technologically challenging as it requires a complex understanding of real-world 3D space (please refer to section G.4 for more information). Even on systems capable of the processing required, many AR devices suffer from tracking latency.

As mentioned by several studies which involved different AR devices (e.g Google glasses, Microsoft HoloLens etc.), the FOV or the space in which digital holograms can appear is still limited and can hinder task execution. A headset with a narrow FOV has more difficulties offering the same level of immersion as a headset with a larger FOV. Additionally, visuals in current AR headsets are not able to meet the high-resolution demands consumers would like to receive.

Battery life is also still limited for mobile AR devices (e.g. MS HoloLens) and would not be adequate for the entire duration of an EVA, which can last up to 8 hours [141].

Furthermore, the augmented reality headsets, as for instance the MS HoloLens, are not very ergonomic as reported from several users during the studies mentioned earlier. It would certainly not be comfortable for an astronaut to wear an immersive headset for the entire duration of an EVA. Surely, the headsets will evolve and become more ergonomic and lightweight. A possibility would be integrating AR into the spacesuit helmet.

Finally, certifications of AR devices, such as the Goggle glasses or the Microsoft HoloLens, for space applications have represented a major challenge in the past.

G.3. Internet-of-Things State-Of-The-Art

In this section, an overview of internet-of-things (IoT) applications and studies performed will be given. IoT sensors and their application will be presented as well. Additionally, studies and developments of IoT for space applications will be presented. Finally, IoT related technological challenges will be discussed.

G.3.1. Internet-of-Things State-of-the-art Applications

The internet-of-things (IoT) is defined as a system of interrelated computing devices, mechanical and digital machines able to transfer data over a network independently. It presents a way to connect, monitor, optimize and automate distributed systems. The idea of the IoT is to enable network-connected devices to exchange information with each other, allowing the creation of large, completely autonomous systems as well as smaller scale home-automation applications [70].

The difference between IoT services and other IT-related techniques within an industrial or home office context lies in the ubiquitous and embedded features of the IoT services [67]. Despite the fact that the main IoT architecture(s) is yet to be defined, their underlying pattern provides with a way to gather one or more "things" (e.g sensors and interactive devices) employing distinct communication interfaces capable of sharing data and communicating with the surrounding environment via specifically designed network gateways [195]. The outcome is analogous to service-oriented architectures (SOA) which can undergo service orchestration for instance via Application Programming Interfaces (APIs). After being processed, the events and information is presented as output of the application to a human's interaction or employed to perform autonomic decisions. Additionally, the IoT concept enables applications to interact with things, machines and/or humans via actuators to exert new control [33].

The potential offered by internet-of-things technologies make the development of a large number of applications possible. IoT is already having a major impact in the following domains: healthcare [156], transportation and logistics [74], industry, farming [189], smart environments (e.g. home, office) [133], disaster management [234] and cities [21]. It is believed that IoT will be an enabling technology for future human space operations as well.

Some example application in different domains, judged as relevant to this research and applicable for human spaceflight operations, will be elaborated upon. In the logistics domain, thanks to real-time information processing technologies such as Radio-frequency Identification (RFID) and Near-Field Communication (NFC) monitoring the entire supply chain by obtaining product related information in a quick and accurate manner enabling a prompt response to intricate and changeable markets becomes possible [133]. Additionally, tracking possibilities that allow continuous inventory location tracking are enabled through IoT as well as monitoring environmental parameters which can enhance the efficiency of the food supply chain [133] as well as the work-productivity of farmers for instance [189]. IoT plays a key role when it comes to smart industrial plants as automation can be improved via the deployment of RFID tags for example. Events can be generated by reading these tags on the production parts, then necessary information about these can be stored allowing machines and robots to handle parts more efficiently after matching data with the enterprise system and the tag [133]. Another notable application involves robotics [33]. IoT technologies have recently shown potential in facilitating human-robot interaction at functional and social levels, and for tele-interaction purposes [191].

In the last decade, several survey papers on the IoT technologies from different aspects have been published. These surveys help to get a general overview on IoT application areas as well as developments and related challenges. The survey by Atzori, Iera, and Morabito [133] gave an overview of the enabling communication technologies and different visions of IoT. The survey performed by Al-Fuqaha et al. [1] provided with an overview of the enabling technologies, possible IoT applications, protocols and the key IoT challenges. Issues concerning security and privacy in IoT have been addressed in the survey done by Andrea et al. [22]. The survey work done by Sha et al. [241] presented challenges issues and opportunities in IoT.

G.3.2. Internet-of-Things Sensors and their Application

A classification of existing sensors and their application is presented in the research by De Morais, Sadok, and Kelner [33]. The following sensors could be applicable to IoT networks of human spaceflight applications [33]:

- ambient (temperature, humidity, pressure)

- motion (accelerometers, gyroscopes)
- biosensors (ECG, EEG, heartbeat, breath)
- identification (RFID, NFC tags and readers)
- position (GPS, magnetometers or fixed wireless access points through the processing of received signal strength (RSS) information)
- chemical (pH sensors, gas sensors)
- interaction (sensors that are human activated, e.g buttons)
- object information (different types of sensors that provide specific information e.g. sensors inside a machine)

G.3.3. Internet-of-Things Space Applications

Innovations in the domain of IoT will certainly have a major influence on space-related research and development activities. Many research projects, especially at universities, have been investigating this powerful emerging technology. Projects on IoT for space applications will be elaborated upon.

Companies such as Hiber [83], the former Magnitude Space, are already planning to employ networks of satellites to cover remote areas and to provide global coverage for IoT devices even in densely populated cities. The company envisions to use many satellites that communicate with their proprietary sensor nodes connected to IoT devices on ground, and upload data to the cloud through the proposed system "Low-power Global Area Network" [83].

NASA has come up with several new ideas for the future, targeting small satellites, such as NASA's Cube Quest Challenge [177] to design small satellites. Others such as OneWeb [183] envision global Internet service with small satellites in 2020, Surrey space center [200] proposes a concept of PCBSat and Satellite on Chip. Moreover, there is the Starshot program by Stephen Hawking [10], [136] a project that envisions shooting a satellite, in the form of a wafer-sized chip, to far away stars at ultra-high-speeds from Earth, through highly concentrated laser beams fired from Earth.

Advancements in low power miniaturized electronics and wireless communication technologies open up new possibilities such as Sensor Wireless Actuator Networks in Space (SWANS) and Space Pixels [218].

Satellites comprise many different sensors and actuators such as magnetometers, sun and star sensors, gyroscopes and reaction wheels, which are all connected via wires, so called harness which increases size and mass and therefore costs. The idea of SWANS, a project by TU Delft, is to replace the harness with wireless without impairing performance [218]. Finally, an external communication system could extend SWANS to comprise a group of satellites to enable multiple applications together with swarms of satellites.

It is envisioned that swarms of autonomous small satellites or wireless smart sensor networks deployed around other celestial bodies, such as Moon or Mars, are able to retrieve an increased amount of data regarding deep space exploration when compared to an observatory on Earth. Precise measurements of aerosol and smoke concentration in the atmosphere could be easier and more effective when using a swarm of small satellites in Low Earth Orbit (LEO) [218]. Finally, it is forecast that swarms of robots could improve space robotics to control robots from Earth with much lower time delays for targeted autonomous applications compared to the state-of-the-art [218].

Work with IoT satellites known as sprites or ChipSats has already proven how small technologies could spur future space exploration advancements [129]. ChipSats might contain specialized electronics for particular missions, investigate weather patterns or gather data about wildlife migrations. Moreover, these IoT satellites could send communications and data-collection tools to other planets, such as Mars [129].

A new concept based on a cyber-physical system extending across space, ground and air, called the Internet of Space Things/CubeSats (IoST) has been introduced [85]. IoST extends the functionalities of the conventional

IoT, by providing an always-available satellite backhaul network and also by contributing real-time satellite-captured information. Additionally, IoST enables new applications by integrating on-ground data and satellite information. The study identified open issues critical to the full deployment, namely system performance evaluations. A study focusing on the performance evaluation of software defined networks (SDN), specifically the Internet of Space Things (IoST) has therefore been performed [13]. The focus lied on the IoST Hub-CubeSat links, the next-hop metrics, the Inter-CubeSat links and end-to-end system operation, latencies, link rates, lifetimes and next-hop availabilities were the metrics covered [13].

Similarly, the use of CubeSats to provide global coverage for the IoT has been proposed by Almonacid and Franck [228]. Practical issues related to energy-efficient communications have been identified and a solution that takes advantage of the delay tolerant nature of the network to enhance the performance of the MAC layer has been proposed.

Since the creation of SAT4M2M [84], the data connectivity provider has been building upon TELDASAT, a low cost Machine to Machine (M2M) communication service using existing space-based infrastructure to provide IoT communications, to establish the IoT space segment. The IoTEE project is an integrant part of that plan, more specifically it focuses on the development of a multi connectivity chip implementing a new communication protocol on the same frequency bands as Sigfox and Lora. The project used the ISS as a testbed for the first Low Power Wide Area (LPWA) satellite. The aim is to complement the terrestrial LPWA and create a global coverage, a network able to send several IoT messages per day using solely the ISS [84].

An IoT space-related project with a different application compared to the studies mentioned earlier, is WEKIT. The project WEKIT involved an IoT sensor network, more specifically a biosensor harness (BH). The BH involved sensors that were able to measure blood volume pulse (BVP), galvanic skin response (GSR), bodily and environmental temperature and humidity. Additionally, two sets of inertial measurement sensors including an accelerometer, gyroscope and magnetometer were used. The measurements were carried out to assess the participants' physiology to analyze potential "stress" indicators which could outline unclear or difficult procedure steps, despite the provided user information on the MS HoloLens.

G.3.4. Challenges of Internet-of-Things Applications

Identified challenges to be addressed when implementing IoT include: distributed systems, in fact the goal is that all the independently built and assembled modules work toward a common goal making up a system that is reliable, adaptive to the space environment and robust [218]. Moreover, miniaturization poses challenges with respect to power supply and hardware limitations (e.g. radiation mitigation) [218]. The utilisation of wireless poses challenges with respect to the reliability of wireless links, problems regarding RF interference and higher power consumption due to the new wireless devices need to be tackled. Wireless technologies such as BLE (Bluetooth Low Energy) and ZigBee have proven to be operational in space, yet complete satellites for instance have not been built employing these wireless technologies [218]. Handling large amount of data is another challenge especially when high resolution data is required and when dealing with miniaturised systems. Synchronization of the swarm satellites or swarms of other IoT devices is a key requirement. In Earth orbit, synchronization is achieved utilising GPS signals, however GPS is not available in non-Earth orbits hence clock synchronization has to be used. Accuracy and clock drifts present a major challenge in this case. Making the swarms of satellites programmable and transformable requires flexibility when designing and executing software modules, this poses challenges with respect to software defined system [218]. Localization of space objects that do not have position sensors like GPS and travel at extremely high speeds while covering a large area, presents another major challenge as well [218].

The IoT, has physical challenges, such as distance limitations, battery life, and durability that become even more pronounced in the extreme conditions found in very hot or cold locations such as future space mission destinations. Tespack, a finnish startup, is attempting to tackle some of these problems by developing solar-powered backpacks with IoT and connectivity capabilities, to take energy-generation to another planet [224]. Furthermore, fast and secure networks that cannot be hacked or damaged, devices and machines able to react in real-time are challenges that need to be tackled. Further challenges currently faced by industries working with IoT are concerning implementation methods, data fusion & handling and cyber-security.

G.4. The Integration of Augmented Reality and the Internet-of-Things

In this section, an overview of the state-of-the-art of augmented reality in the integrative internet-of-things will be given. It is a relatively new field of research, hence various types of applications across different industries will be presented to get an understanding of the potential of these two emerging technologies. The few identified space applications combining AR and the IoT will be discussed as well. Additionally, augmented reality techniques to represent, transmit and interact with information of physical objects, part of an internet-of-things network are provided.

G.4.1. Augmented Reality in the Integrative Internet-of-Things Environment State-of-the-Art

Multiple research projects have been exploring different aspects of the internet-of-things, namely IoT technological challenges, the creation of applications and development of enablers for IoT systems as well as the design of models for interaction. An essential part is represented by technology enablers [133], momentarily several research groups in the IoT community are addressing enablers for computational intelligence and context-aware computing, which is part of the physical as well as the digital worlds [70]. Similarly to AR, IoT systems include physical and digital spaces, therefore, it seems natural to use AR as a user interface (UI) for IoT interaction [70].

One project by the MIT Media Lab, that began with exploiting the two concepts, namely ubiquitous computing and augmented reality, is the Reality Editor [231]. The Reality Editor is a system that enables editing the behavior and interfaces of “smarter objects”, namely objects or devices with an incorporated embedded processor and communication capability. This system is able to map graphical elements on top of the tangible interfaces on physical objects, such as push buttons or knobs using augmented reality techniques [231]. Ledo et al. [47] proposed a system that interacts within the internet-of-things, namely the proxemic-aware controls. This system takes advantage of the spatial relationships between the user’s handheld device and all the devices in the proximity [47]. A completely different approach to control devices has been proposed by Chen et al. [245], this uses infrared targeting to control surrounding devices through a head-worn computing device.

Augmented reality is seen as one of the emerging technologies that is most feasible to interact with internet-of-things objects [70]. Initially, a common issue encountered when using a smartphone enabling AR is the so called "keyhole" problem, where users interact with their surrounding environment via a screen, employing a camera to recreate reality [70]. This led to a disconnection between the user and the surroundings and therefore limited usability. Recently however, several head-mounted displays have appeared, such as Microsoft HoloLens and the Meta 2, these glassed-based AR systems certainly create new opportunities to experiment with IoT interaction [70].

In the meanwhile, architectures combining AR and IoT to improve the user’s experience through enhanced interaction have been proposed across many sectors from farming [189] to retail [46] to remanufacturing [151]. A recently conducted study for precision farming argued that conventional IoT data visualization is mostly carried out in offsite and textual environments, hence not stimulating a user’s sensorial perception and interaction. The so called AR-IoT system overcame this issue by overlaying IoT data onto real-world objects directly, enhancing object interaction. It demonstrated that using AR-IoT technology was less error prone and more promising than traditional visualization methods [189]. A great way of taking advantage of the IoT system is real-time visualisation combined with AR, where the AR device detects the IoT devices in the screen and overlays the acquired data at their location. In this precision farming project, a WSN (Wireless Sensor Network) was set up to visualize the virtual contents.

Other concepts, that allow intuitive interaction between user and systems and simple data interpretation, have been developed by Park, Yun, and Kim [246] and Jo and Kim [117]. VisIoT developed by Jo and Kim [117] is a system able to track the location of a wireless transmitter in internet-of-things devices and display it on the screen of an AR device. Useful applications that have been identified by this study are IoT for: smart factories, agriculture and smart homes. It is envisioned that in smart factories information on the status of different IoT machines and sensors can be easily retrieved through an AR device [117]. The data from IoT sensors and machines are overlaid on the screen of the AR device at the location of the IoT devices. This way it becomes possible to identify the exact location that needs more attention, for instance an inventory that is out of stock or a device with exceeding its temperature threshold. The application VisIoT [117] identified smart farming as a potential use case similarly to the precision farming study [189], it is argued that through IoT data visualisation, the identification of spots that need more fertilizer or pest control, while working in the field, could enhance work-productivity. For smart homes, the combination of AR and IoT could ease the interaction with different devices such as IoT-enabled lights and the monitoring of other home IoT devices that do not have a screen [117].

Concepts combining IoT and AR within a smart city maintenance service, to enhance the accessibility of sensor and actuator devices, where responsiveness is crucial, have been investigated [249]. White et al. [77] discusses the use of augmented reality in IoT to provide contextual information to service users and providers. Applications to enhance the user experience through personalized ad-hoc information for commuters, tourists and farmers have been analysed as well [77].

To the best of my knowledge there is only a few studies that investigated the combination of AR and IoT specifically for space applications, these will be presented hereafter.

An interesting study that developed a system that combines AR and IoT technologies for just-in-time astronaut training was performed by Karasinski et al. [99]. A mobile prototype consisting of the MS Hololens and a network of custom internet-of-things sensors was developed. The prototype was able to show operational procedures to the user and allowed the user to interact with them in real-time. Furthermore, it allowed users to easily locate tools through precision navigation. The aim of the prototype was to enable just-in-time training. Five qualitative user tests were performed to get insights and learn about best practices that lead to a reduction in cognitive load and time saving [99]. It was concluded that augmented reality and embedded IoT sensors were an effective combination of tools capable of enhancing productivity and reducing procedural execution times. User testing results did not include numerical results or statistical data analysis [99].

Another study that investigated an IoT setup in combination with AR was the space-related AR study comprising EdcAR and WEKIT aimed at supporting manual work of astronauts [128]. It involved the so called IoT demo box which was aimed at demonstrating and testing IoT features in AR. This box included a BeagleBone microcomputer which run the MQTT server and the WLAN router simultaneously. The AR system could connect via the MQTT IoT standard and AR-visualizations were showing information based on the IoT data [128]. As mentioned earlier, project WEKIT, described in subsection G.2.2, involved an IoT sensor network, namely a biosensor harness (BH) [139]. The BH involved sensors that were able to measure blood volume pulse (BVP), galvanic skin response (GSR), bodily and environmental temperature and humidity. Additionally, two sets of inertial measurement sensors including an accelerometer, gyroscope and magnetometer were used. The measurements were carried out to assess the participants' physiology to analyze potential "stress" indicators, which could outline unclear or difficult procedure steps, despite the provided user information on the Microsoft HoloLens [139].

G.4.2. Augmented Reality Information Representation for Physical Objects

Augmented reality services have to manage information for their augmentation targets, namely physical objects part of the IoT network. A review on the currently used approaches of data representation for AR use is performed. Initially, AR systems included all the assets and content that had to be visualised, using Vuforia [236] and the ARToolKit [23] for instance. The development of location-based geographical and AR services however, implied the need of content separation and format specifications and the means to support the notion of "everywhere" content and service, as well as a unified management of content on the server. Markup languages such as HTML, KML (Keyhole Markup Language) [131] and ARML [182] (Augmented Reality Markup Language) serve this purpose. ARML was specifically proposed to define geographical points or landmarks of interest and to associate GPS coordinates and simple augmentation content, namely text, logos, and images. A standard content format to represent different forms of AR services does not exist yet, it will be crucial however once thousands of objects will be able to communicate with computers and connect with each other [46]. A so called "webization of things" is forecast in the near future where every component of the IoT network will have their unique URL, show its content in standard format (HTML or Javascript) and be unified and accessible through one Web framework [46]. A proof-of-concept implementation applied to a "digital-analog" shopping experience was developed by Jo and Kim [46]. This used an AR client that received the datasets and contents from each discovered IoT item in the user proximity in standard format to allow for recognition, visualization and interaction [117]. It should be noted that the digi-log concept by Jo and Kim [46] contained information directly in the IoT device, hence it was not accessed and retrieved from an external server.

G.4.3. Information Transmission Techniques between Internet-of-Things Network and Augmented Reality Device

The MQTT (Message Queuing Telemetry Transport) protocol is frequently used to send information to a specific topic and a device or sensor belonging to the internet-of-things network [128], [77], [1] [33]. MQTT is a machine-to-machine (M2M)/"Internet of Things" connectivity protocol, it is a lightweight publish/subscribe messaging transport, useful for connections with remote locations where a small code footprint is required and/or network bandwidth is at a premium [203].

There are several solutions to enable the transfer of information of IoT devices or sensors to and from an AR device, e.g. Microsoft HoloLens. One solution would be that the IoT component includes an Arduino onboard with Wi-Fi module which is subscribed to a topic on the MQTT protocol. The Raspberry pi 3 for instance can run a Mosquitto server and then both the Microsoft Hololens and the IoT sensor/device can be connected to that broker/server and retrieve or send information from or to the topic, respectively. The integration of the MQTT functionality into the HoloLens can be accomplished using an MQTT client library. The M2MQTT library [150] is a possibility in this case, as it supports all .Net platforms and WinRT platforms for Internet-of-Things and machine-to-machine (M2M) communications [150]. The only problem is that it is not possible to use the same plugin for the HoloLens and Unity [227]. This problem can however be solved [150].

Another possibility to enable communication between an IoT device or sensor would be installing an MQTT client on the device or sensor and use the AR support equipment developed for the HoloLens during the General Support Technology Program (GSTP) AR study for ESA-EAC. This study comprised the development of an AR support equipment added to the MS HoloLens to overcome its limited processing resources and to provide an interface with rover-related systems, such as broker or Meteron Robotic System (MRS). It is employed to maintain the connection with rover-related systems and it serves to translate between the HoloLens and these systems. The AR components required to support operations are, the Message-oriented Middleware (MoM) which is an ActiveMQ broker instance to which the AR application and the METERON Operations Environment, MOE Adapter connect. The broker is started via the so called bowerick [199] wrapper to ease deployment and start-up. bowerick is a wrapper around ActiveMQ and additional supporting libraries aiming on easing simple MoM deployments and tasks [199]. It provides a simple way to start a multi-protocol broker plus a command line client which can aid during debugging. An adapter able to translate between the MoM and the rover-related software (broker and MRS) is required. Originally, the MOE Adapter was used as support, however this adapter was specifically developed for the MARVIN project, which involved other parties apart from EAC, hence cannot be applied for other purposes. An embedded MQTT client on the IoT component which includes a JSON library can replace this adapter. An adapter for the MARVIN project more specifically for the rover has been developed and successfully tested [66].

Another alternative is to use the Internet-of-Things network developed by a previous intern which includes MQTT and uses IO Broker. Sending IoT component information via MQTT protocol to a Mosquitto broker is presumed possible, it has however not been tested yet. An MQTT client on the HoloLens that connects to the Mosquitto broker would be required in this case to enable communication. It should be noted that the network currently consists entirely of accelerometers.

G.4.4. Augmented Reality Information Storage, Management and Indexing for Physical Objects

Currently, the most common method of visualization and interaction with digital objects is via GUI-based interfaces [46]. Augmented reality provides additional capabilities with target object recognition and identification [111]. Despite the fact that a server-based approach to ubiquitous AR services with everyday physical objects is possible, AR services for everyday spaces have been cumbersome to accomplish because of the object recognition process, complicated feature matching, and content look-up for multiple objects [46]. Nevertheless, cloud computing services featuring high performance are available for the fast object matching process and for speeding up the associated content retrieval in providing an AR service [248]. It remains difficult however to support the scalability to reach ubiquity. Alternatively, a solution would be to connect to a singular areal server that manages only a restricted number of objects [46]. No solution has been found so far to scale AR services and efficiently manage the large amount of data from the surrounding environment. The "unified" Web-based solution presented by Jo and Kim [46] relies on the central network server architecture which causes serious performance issues. A few studies are trying to obtain datasets directly from objects close to the user in the surroundings [117]. In the concept proposed by Jo and Kim [46] the AR client can easily and rapidly identify, track and retrieve content from surrounding objects, as only a limited amount of objects are involved.

G.4.5. AR Interaction with Physical Objects

Augmented reality offers different means of interaction, in situ object control methods are already available [117] together with improved control and simulation features [46]. CyberCode [112] was one of the first proposed means of interaction using 2D barcodes for object identification, detection and manipulation. As mentioned earlier Heun, Hobin, and Maes [231] proposed the Reality Editor, a concept that allows for direct mapping of operation via embedded processors and communication capabilities within so called "smarter objects" [231]. GuideMe, an application proposed by Müller, Aslan, and Krüßen [138] allows users to access instructions for appliances. The interaction with the appliances' instruction in this case was possible through marker-based detection. Other interaction types were proposed and compared by Alce et al. [70], namely floating icons, a floating menu and World in Miniature (WIM). The floating icons concept is based on the idea that interactive holographic icons are placed in proximity of the device they represent. The floating menu on the other hand reproduces a more traditional approach to IoT interaction while WIM, is based on the idea that the environment in which a user wants to control IoT devices is modeled as an interactive hologram. It remains unclear how a consistent and coherent AR interaction framework has to be established as most AR studies are designed in an ad-hoc manner.

There are several methods to detect and track objects, more specifically devices or sensors that are part of an internet-of-things network. AR offers three different types of detection and tracking: marker-based, marker-less and location-based tracking. In a marker-based AR application the images that have to be recognized are provided beforehand. In this scenario one knows exactly what the application will search for while acquiring camera data. It should be noted that it is much easier to detect things that are hard-coded in the application beforehand. A marker-less AR application is capable of recognizing things that were not directly provided to the application beforehand. This is more difficult to implement as the recognition algorithm running in the AR application is required to identify patterns, colors and other features that could be present in camera frames. Location-based augmented reality ties augmented reality content to a specific location. A well-known example of location-based application is the Pokemon-Go application [194].

Marker-based Detection and Tracking

When using marker-based detection and tracking, the digital world is anchored to the real world. In this case, the device must first recognise the object the user is looking at from the live camera view on the AR device. This can be accomplished by placing a distinctive picture or shape on the object which can then be recognised and the animation can start instantly, tracked to the appropriate place on the device. The user can also move the physical object around and see the virtual world attached to the real surface of the page.

Augmented reality markers are visual cues that trigger the display of the virtual information. Markers can be normal images or small objects which are trained in advance such that they can be recognized later in the camera stream. Once a marker is detected and recognized, its scale, position and rotation are derived from visual cues and transferred to the virtual information. More details on augmented reality marker types and software development kits supporting marker-based tracking are elaborated upon below.

Augmented Reality Markers

There are several types of augmented reality markers: framework, NFT (Natural Feature Tracking), GPS and object markers. Framework markers are markers that are square and have a significant black and sometimes white border. Barcodes and QR codes are part of this 2D marker category. NFT markers are images that do not need the black border, instead they are characterised by distinguishable natural features. GPS markers allow to place objects based on their known position. Object markers are similar to the 2D markers, framework and NFT markers mentioned earlier, however these are 3D markers and have the advantage that the viewing angle does not matter during the recognition phase. Finally, there are activation markers, these are replaced by the "activation" of an application, RFID tags used for short distance wireless communication and sound-markers, such as speech commands, can be used.

Augmented Reality Marker-based Application SDKs
There are several SDKs (Software Development Kits) available for image and object recognition that allow for marker-less tracking and detection, viable options are Vuforia image and object recognition SDK, ARToolKit object recognition SDK and VisionLib.

Vuforia offers a tool called VuMarks [239] which includes customized markers that can encode a range of data formats and support unique identification as well as tracking for AR applications. Vuforia Advanced Model 360 object recognition [237] offers instantaneous recognition and tracking of physical objects, the viewing position does not matter and digital content can be augmented realistically on an object by detecting the edges of the object from all angles. The Vuforia Advanced Model Target 360 works through a database which is trained in the MTG desktop tool, the Model Target database is uploaded to the Vuforia Cloud and there it is trained by a deep learning process [237]. While Object Reco [238], also an object recognition tool of Vuforia, detects the object based off the texture features, the object needs to be scanned in using the scanner tool.

The difference between object recognition and model targets is that the latter are detected based off the geometrical data of the target and require the 3D/CAD model of the object.

VisionLib [233], a multi-platform library for augmented reality applications, offers highly accurate 3D object tracking, furthermore setting up the model tracker is easy, no prior registering targets or surroundings is needed, one can simply use their own 3D models and CAD data to create trackers, as they are perfect references for physical 3D objects. Finally, robustness can be enhanced by extending the model tracker with marker-less feature tracking.

Wikitude AR SDK for HoloLens [242] offers valuable augmented reality functionalities for image, object, barcode and QR code recognition as well.

Marker-less based Detection and Tracking
Marker-less tracking is a positional tracking method. This method as the name suggests does not rely on optical markers hence it is more flexible. Marker-less tracking solely uses sensors to calculate orientation and position of the AR device camera. A model-based approach represents a way to determine the placement of virtual objects with real ones. For this method, the model of the real object can be encoded as a (CAD model) [244]. Hence, marker-less AR tracking would continuously look for the image it receives with the known 3D model and compare it [244].

As mentioned earlier VisionLib extended the model tracker with marker-less feature tracking [233]. HoloLens offers marker-less tracking as well thanks to its spatial mapping capabilities based on SLAM (Simultaneous Localisation And Mapping). The MixedReality Toolkit is comprised of a number of components and scripts that are intended to accelerate the development of augmented reality applications designed to target the Microsoft HoloLens [161], this toolkit could therefore be valuable to develop marker-less AR applications.

Location-based Detection and Tracking
Location-based systems consider the user's location when processing and displaying information to the user [76]. Location-based AR applications work if geo-positioning and augmented reality technologies are correctly implemented and provide precise data. This type of AR application uses so called geo-markers, also called points of interest (POI). The AR application in this case has to identify POIs and determine the exact position of the employed AR device. GPS is one of the famous technologies to accomplish this, nevertheless, there are several options, including those that allow for indoor positioning [16]. An AR application often uses data from a digital compass and accelerometers together with GPS data. When the application identifies a specific POI, it is able to trigger augmented reality content which is displayed on top of the real environment.

G.5. User Interface Design Approaches

In this section different design approaches including models and principles used for the development of user interfaces are presented.

Designing user interfaces that are effective is a major challenge when it comes to emerging technologies that present new ways of user interaction and perception and do not have established design guidelines [100]. Augmented reality is one of these emerging technologies as it has the potential of changing the way humans perceive the world. A literature survey performed in 2004 showed that user-based studies have been rarely utilized for augmented reality user interface design as engineering challenges were the main focus. Swan and Gabbard [97] identified the need to develop AR systems from a technology-centric medium to a user-centric medium. From the survey it could be concluded that of 1104 articles on augmented reality, only roughly 3% addressed aspects of Human Computer Interaction (HCI) and only roughly 2% described a formal user-based study [97]. Swan and Gabbard [97] state that to create effective user interfaces for emerging technologies, such as augmented reality it is crucial to understand perceptual and cognitive characteristics of the users, hence develop AR based on performed user studies [97]. A more recent study which analysed AR studies between 2005-2014 concluded that overall on average less than 10% of all the AR papers were user study papers [7]. Additionally, analysis showed that the majority of the studies included little field testing and few heuristic evaluations. It is believed that it heavily depends on the user satisfaction whether AR devices turn into a market success. Therefore, a study on the user satisfaction of AR applied to different use cases was performed [82]. AR user satisfaction was divided into two categories: interaction satisfaction and satisfaction with the AR device. 142 participants from the aeronautics, medicine, and astronautics sector were chosen. The study investigated factors such as age and gender, level of Internet knowledge, level of education and the participant role. Despite the fact the users had no experience with the Microsoft HoloLens, general computer knowledge resulted having a positive effect on user satisfaction. Additionally, user satisfaction of both learning and teaching were acceptable [82].

Different design processes that have been developed for software and user interface design, including the waterfall model, the spiral model, the star life cycle model, the Vee model, the user engineering life cycle model and the understand, create and evaluate model have been investigated.
It was considered important to get an overview to be able to finally select the most feasible design process for the proposed research project. Note that only the relevant investigated models are presented hereafter.

Mayhew [45] presents a user engineering life cycle model (see Figure G.8) which is iterative and focused on integrating users throughout the entire development process. The cycle is based on screen design standards, that are evaluated and updated in an iterative manner.

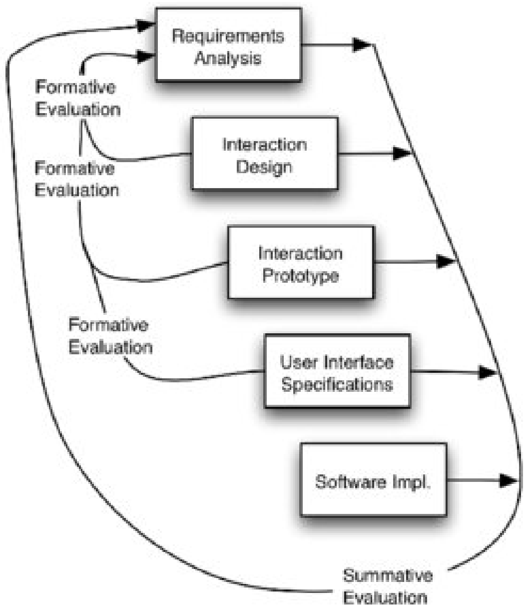


Figure G.8: User Engineering Life Cycle Model [45]

User-centred design is a framework that emerged in the mid 1980s and is nowadays one of the guiding principles to develop and design usable technologies, it is also an ISO standard, namely ISO 9241 (see Figure G.9 [92]).

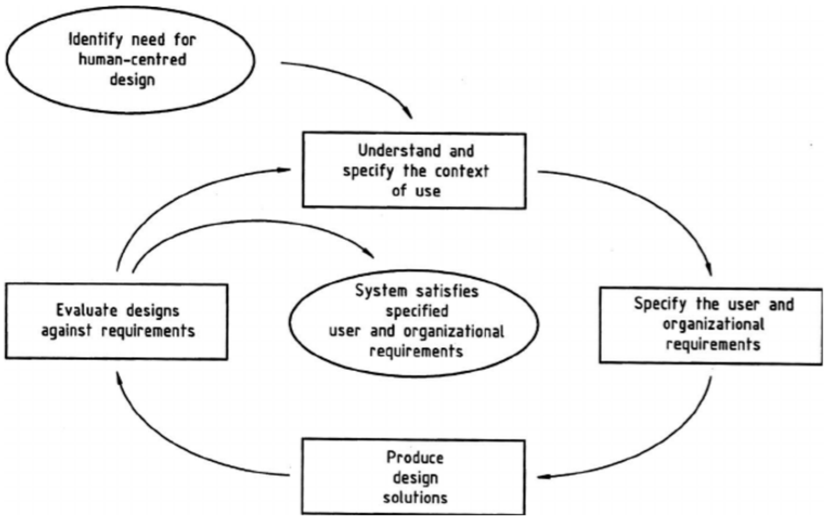


Figure G.9: User-Centered Design Process - ISO 9241 [92]

The design framework can be applied to both software and hardware development [149]. The principle behind this design process is that the user is at the centre of the development. The focus is on the people that will actually use the new product and on how they will use it. The system developed needs to consider the skills and judgement ability of the users and aims at supporting the user during their work [149]. Several sets of principles have been proposed. Gould and Lewis proposed the following three design principles [116]:

1. Early Focus On Users and Tasks
2. Empirical Measurements
3. Iterative Process

The first principle implies that the designer has to understand who the users are, which involves understanding their cognitive, behavioral, anthropometric and attitudinal characteristics. Additionally, the task that the users will carry out needs to be studied. The second principle involves testing the developed prototypes on the users that have to carry out real work and then analyse their performance and reactions based on observations made. The third implies that there should be a cycle of design, which includes performing user testing and measurements, redesigning and repeating this if necessary [116].

The set of principles of Norman [49] and Shneiderman's eight rules [26] serve to understand the design philosophy of user-centred design, however as suggested by Haklay and Nivala [149] and Lee et al. [96] the framework suggested by Gould and Lewis [116] is more comprehensive and hence applied as basis of implementation of user-centred design in design projects.

Lee et al. proposes a so called, understand, create and evaluate cycle (see Figure G.10) based on the three major human-centered design process phases [54] in line with the framework suggested by Gould and Lewis [116] and the set of principles by NASA [173] presented hereafter. The cycle is based on first understanding the user through observations and task analysis and second creating a prototype that combines the previously acquired knowledge on the user and tasks with principles, guidelines and human characteristics. Finally, evaluations based on heuristics and usability tests with papers prototypes and product prototypes are performed in an iterative manner to arrive at the final step of the product release [96].

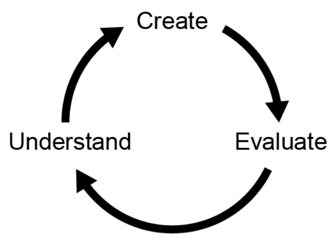


Figure G.10: Understand, Create, Evaluate cycle [96]

The human-centered design principles proposed by NASA [173] are:

- Involvement of the users and understanding of tasks and users
- Clear function allocation between the technology ad the users
- Iterative design
- Multidisciplinary design

Active involvement of the users is required and beneficial, as valuable knowledge on tasks, context of use and performance can be retrieved from the users. Involving the users allows to better understand the needs and receive feedback on how the systems will be used, what the task demands are and consequently formulate requirements that lead to improved design decisions [173]. Task analysis is an essential methodology which identifies system level and subsystem level tasks. The focus is again on the human and on how the task is performed. Cognitive and physical capabilities and limitations as well as critical crew tasks have to be identified. Critical crew tasks can occur nominally or off-nominally and they comprise tasks that are essential to crew health. If these tasks are not done in the correct manner, they can lead to loss of mission, loss of crew or undesired vehicle states, hence they have to be identified to reduce the probability of mishaps or errors occurring and to avoid errors and improve safety. The task analysis should yield understanding of inter-task interactions, critical crew and systems task to develop procedures, make human reliability assessments and perform verifications [173].

The capabilities and limitations of the human against the technology need to be weighted and a decision has to be made on which function should be carried out by whom. Factors to be considered are: speed, accuracy, reliability, flexibility of response, financial costs, users well-being and strength. It is important that it is not the tasks that cannot be carried out by the machine that are assigned to the human, a meaningful set of tasks should be allocated instead [173].

The process is highly iterative, tests, models and analysis with active user involvement should be performed and the results used to progressively refine the solutions [173].

In Figure G.11 the activities revolving around the human-centered design method proposed by NASA [173] are shown.

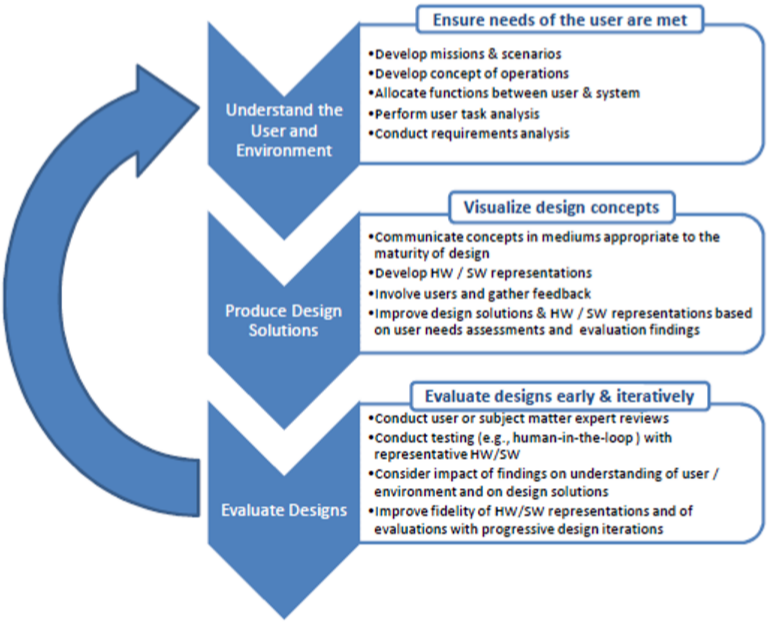


Figure G.11: Human-centered design activities [173]

The methods applied during the various steps of a user-centred design process vary depending on the development time available. There are several methods one can apply to perform a context, user and task analysis. To collect task data one can perform several activities [96], these are listed below.

- Observe the users interact with current product versions
- Perform retrospective protocols such as video recordings or verbal protocols as both retrospective and prospective protocols can be used
- Perform interviews
- Perform surveys and questionnaires
- Use automatic data recording techniques

After data collection and interpretation one can then proceed to the innovation step, consisting of a specific analysis of the data and draw conclusions to improve existing systems or to create new systems. This step is characterised by a user task analysis. There are different types of task analysis [96] listed below.

- Scenarios and use case analysis in which situations and tasks relevant to the use of the product are defined
- Environment and context analysis in which the product is used are defined
- Workload analysis in which one needs to assess the mental and physical demands of the product use on the user

- Safety and hazard analyses in which the safety implications of the product are defined
- Function allocation analysis in which the task distribution between user and technology is defined.

Prototype development follows the user task analysis. There are several types to be used these are listed below together with their features.

- Paper prototypes aid in understanding the users needs early in the conceptual design phase
- Wireframes show the grouping of content and content location, they also allow the communication of software design ideas
- Mockups render ideas more concrete
- Prototypes support user evaluation as one can use and react to it

The iterative evaluation and testing phase, which follows the creation of prototypes is elaborated upon in section G.7.

Several different models applied for user interface design have been presented in this chapter to get a better overview and be able to choose the most feasible option for this research. As suggested by literature (see section G.5), when designing for people, especially when designing an interface, a user-centered design method is highly recommended. As stated by Swan and Gabbard [97] it is crucial to understand perceptual and cognitive characteristics of the users, hence develop AR based on performed user studies [97]. It was found that in the past AR studies did not often adopt user-centered design methodologies, in fact less than 10% of all the AR papers between 2005-2014 were user study papers [7]. Additionally, analysis showed that the majority of the studies included little field testing and few heuristic evaluations. That survey however, included only one AR space application related study, namely the one by Markov-Vetter and Staadt [160]. It should be noted that most AR studies for space applications focused on the user and gathered feedback from the users through direct observations, surveys and questionnaires and in some cases performed usability tests. Based on the review of the different design models, principles and standards (see section G.5 and section G.6) it was concluded that a user-centered design method is most feasible for this research. This method includes an initial user and task analysis, followed by prototype development and evaluations. This design process is highly iterative. Task data collection can be performed via different methods e.g. direct observations, interviews, questionnaires and surveys. After collecting task data, a user task analysis that can include a workload, context and use case analysis is performed. Consequently, different types of prototypes can be designed depending on the application. Finally, it is concluded that a user-centered model which is in line with the human-centered design model used by NASA [173] and the principles proposed by Gould and Lewis [116] suggested by Lee et al. [96] is the most feasible one for this research.

G.6. Augmented Reality Interface Design Considerations

In this section, first important design aspects, criteria and principles when designing human-machine interfaces, focusing on displays will be presented. Then, the design of specifically integrative displays such as Head-Mounted Displays (HMD) will be analysed. Finally, human-computer interaction design principles are mentioned.

G.6.1. Display Design Principles

In this section, display design principles including attentional, perceptual, memory and mental model principles are presented. Further, principles and criteria for alerts, icons and labels relevant to this research are given.

Lee et al. define displays as artifacts designed to guide attention to system information relevant to the situation and to then aid in the perception and interpretation of this information [96]. In cockpits, pilots use software displays as their primary interface to command subsystems and monitor their health and status hence it is essential that these displays provide situational awareness, enhance safety providing textual and graphical information on time and reduce the workload [173]. There is certain standards set to ensure a common design framework for all flight and system interfaces as well as easy learning, enhanced crew productivity and mission safety through consistent user environments [173], these standards also exist for spaceflight system design and will be elaborated upon.

Displays and user interfaces are closely linked concepts, in fact user interfaces do not only display information but also accept inputs to alter the display and control the system. Displays show information regarding the requested action or state of the system considered. Furthermore, they relate the person to the system overcoming the so called gulf of evaluation, which is the difference between the person's understanding of the system and the actual state of the system. To design an effective display it is important to consider the compatibility of the displayed system information with the perceptual and cognitive properties of the users following the user-centered design principles mentioned in section G.5. There are 15 human factors principles to follow in order to assure this compatibility, these will however change depending on the task considered and require adaptation. The first step to take when designing a display is to define the task and the information needed by the user to carry out the task, this is in line with the user-centered design principles mentioned in section G.5.

During task analysis, scenario development and the development of an operational concepts, questions to be addressed include [173]:

1. What kind of interaction with the display formats will the user have and what possibilities are there (e.g voice command, eye gazing, gestures)?
2. Are several instances of the display formats possible? If it is the case, how will real-time data updates and commands be handled?
3. What are the insights in terms of system status and faults the user can get?
4. What are the cautions and warnings expected and how will they be handled?

Attentional, Perceptual, Memory and Mental Model Principles
Display design principles can be categorized into attentional principles, perceptual principles, those that relate to memory and finally those that can be traced to the concept of the mental model [96].

Attentional principles: include salience compatibility, minimization of information access costs, proximity compatibility and avoidance of resource competition. Urgent and crucial information needs to attract attention, features that can aid in that are: contrast, colour and flashing. Auditory alerts are highly salient as well, however interference with voice communication loops need to be considered in case of EVA astronaut applications, as mentioned in subsection G.1.4. Furthermore, it is important to relate salience with solely important information to avoid alarm fatigue and annoyance. Relevant information should be kept easily accessible reducing the accessibility costs. Information should be arranged such that their mental proximity is related to the display proximity creating high proximity compatibility [29]. Display proximity can be achieved in different ways by using the same colours, similar formats, configuring them in a pattern as well as linking them with lines for instance. Finally information should be divide into several sources and presented stimulating different cognitive systems (e.g. auditive system, visual sensory system) [96].

Perceptual principles include: making a display legible, avoiding absolute judgements limits, supporting top-down processing, assuring discriminability and exploiting redundancy gain. Legibility as well as audibility is an essential need of a good display. The user should not be required to judge the level of a variable on the basis of a single sensory variable that contain several different levels as it leads to errors of judgement. As people have expectations based on past experience, signals and information are perceived according to that, which makes it thus important to support the user enhancing saliency of the information that is contrary to what is expected. Using different physical forms reduces the chance that factors degrading one form have the ability to degrade another form as well. Information that is similar creates confusion hence it is advised to avoid unnecessary similar features and highlight dissimilar ones [96].

Memory principles include: presenting knowledge in the world, supporting visual momentum, providing predictive aids and being consistent. Presenting knowledge in the world to aid humans is necessary as the memory of humans is vulnerable and has a limited capacity. Humans are often required to retain information from one display to another when displays consist of different elements for instance sequentially viewed displays.

Methods to enhance visual momentum reduces memory load and information integration capabilities. Proactive behaviour is desired rather than reactive behaviour, displays that can predict what will happen and aid in proactivity lead to better human performance. Consistency is important when designing a display, it is advised to design displays that are consistent with other displays the user is required to work with or has been exposed to before. Interface designers should be consistent when it comes to colour coding for instance, it is important to use colours consistently to avoid confusion. To give an example, red should be consistently used for states to be avoided or even for danger [96]. More information about colour coding can be found later on in this section.

Mental model principles include: the principle of pictorial realism, namely the fact that a display should resemble the variable that it represents [217] and the principle of the moving part, namely that the moving parts of any display of dynamic information is required to move in a spatial pattern and direction that is in line with the mental model the user has on how the represented element actually moves in the physical system [217].

Alerts
During surface operations, alerts such as warnings, cautions and advisories will occur as mentioned in subsection G.1.3. A display can aid in the detection of these alerts. To effectively design alerting displays considerations have to be made such as salience compatibility mentioned earlier. To avoid confusion of alerting severity the aviation sector has set up a number of guidelines to follow, these include consistent colour coding for instance [96]. Colours should be used to complement status information and not as the sole indicator as different lighting conditions or crew abilities perception issues can affect the information transfer.
Example conventions include [173]:

- Yellow: caution
- Red: warning or emergency
- Blue: advisory
- Grey: unavailable function
- White: available dynamic function
- Green: normal state

Icons and Labels
Icons are an effective way to grab the users attention and convey a message while labels can aid in the interpretation of an icon and avoid ambiguity. Labels are essential for a user interface as they provide information on the identification of and the instructions for certain activities [173]. Hazard, caution, warning and emergency labeling are meant to convey critical information.

Icons are frequently used in the aerospace sector as well and astronauts are very familiar with a certain set of icons. The SSP 50313 - Display and Graphics Commonality Standard [181] contains all icons and labels used for the International Space Station Program. These icons can certainly be adopted for the HMD design of the AR application when necessary. However, for future space exploration scenarios, new icons will be required, thus looking into the design criteria for icons might be useful. Design criteria include: visibility, meaningfulness, location and discriminability [96].

Finally, it should be noted that preferred fonts ensuring good readability include: Helvetica and Arial [173]. For further details on guidelines for display design (e.g. layout, text content and size, labeling, spacing, grouping, use of colours and graphics) it is advised to consult the NASA/SP-2010-3407/REV1 Human Integration Design Handbook (HIDH) [119].

G.6.2. Head-Mounted Displays

One of the most technologically powerful AR devices currently on the market is the Microsoft Hololens [163] which is essentially a Head-Mounted Display, HMD. These displays help the user to reduce the access costs and keep their hands free to carry out other activities. For pilots this is beneficial as it helps retaining the full overview of flight instruments and at the same time keep an eye on the outside world [39]. For assembly workers it is beneficial as information on how to carry out the work can be directly displayed on the HMD easing the integration of information sources [243]. One of the problematics with HMDs relates to the efforts to place conformal imagery on the displays due to image updating delays. These conformal displays are augmented reality displays as they show spatial positions in the outside world and have to be updated when the user rotates their head relative to the world. In case the image is not updated fast enough upon the head rotation of the user, this can cause motion sickness and disorientation [96], else the user might adopt unnatural behaviours, compensating this issue by moving their head slower [68]. Generally, it is yet unclear whether it is more advantageous to use a head-up HMD or hand-down handheld display [159], however flight control performance is higher when the critical information on flight instruments is shown head-up [204], [30]. For astronaut EVA applications head-up HMD seem to be more advantageous as it can be used hands-free.

Head-mounted displays fall under the category of wearable technology. There are additional design considerations to be made to ensure these wearable technologies enhance productivity, user satisfaction and safety [96].

The wearable technology should be comfortable for the user to wear. The MS Hololens 1 [163] used for several AR application studies for instance WEKIT [34] and Sidekick [175] does not offer good ergonomic features, nevertheless improvements are implemented in the new version of MS Hololens 2 [164], as it is lighter and has a more balanced center of gravity.

Information should be prioritized, it is important to identify which information should be pushed forward as the spaces are constrained [96]. Information should be visible with a single glance, it should be avoided that one has to page through multiple screens. Non-visual cues should be considered. Complex tasks should be simplified and divided in smaller ones. Signs and indicators should be provided to allow the user to activate or deactivate surrounding devices detection if such a feature is available in the first place.

G.6.3. Human-Computer Interaction Design Principles

Head-mounted AR displays as for instance the MS HoloLens allow for and require user interaction. Thus design considerations with respect to the interaction aspect have to be made in addition to the display design considerations. Similar attention, perception, memory and mental model principles as for the display design apply however, response selection and interaction principles have to be accounted for as well [96].

Response selection principles include: choosing the appropriate defaults, simplifying and structuring task sequences. Ensuring that default values are useful and practical is important as well since the user might not remember how to change every setting of an interface that has many options and a lot of information. Furthermore, carrying out a complex task is often easier when the task is split in into smaller steps [96].

Interaction principles include: showing the system state clearly and supporting flexibility, efficiency and personalization. Prompt feedback is necessary with respect to the status of an executed action. The application should match the users needs. Depending on the tasks performed, the interface should provide accelerators or shortcuts for tasks that are frequently performed for instance. Additionally, having the possibility of personalizing the interface based on the users needs is advised [96].

G.7. Evaluation and Testing Methods

In section G.5 the design processes were presented and it could be concluded that the processes follow a cycle which includes understanding, creating and evaluating. In this section, the focus will be on the evaluation aspect. Please note that the evaluation process is an iterative process. It is important to involve the users throughout the design process such that they can participate and understand the design decisions, only this way user satisfaction can be achieved. Especially in the case where the human has critical control responsibilities over a system or when the interaction is crucial to achieving the mission goals [173]. First, available types of evaluation will be shown, second available designs for studies are presented, after that data analysis techniques are given.

G.7.1. Type of Evaluation

The types of evaluation include the following options [96]:

- Literature Reviews, Heuristic Evaluations and Cognitive Walkthrough

- Usability Testing
- Comprehensive Evaluations and Controlled Experiments
- In-service Evaluation
- Expert Reviews

Literature reviews serve to identify how people behave in related situations, this phase involves reading topic related research papers, books and reports. During heuristic evaluations a small set of evaluators analyze every aspect of an interface to assess whether it fulfils the usability requirements. Nevertheless, other requirements such as user satisfaction and safety need to be fulfilled as well and can be assessed during heuristic evaluations. Since evaluators tend to miss certain aspects and usually different evaluators identify different aspects, Nielsen [104] suggests using a minimum of three, preferable five evaluators to perform heuristic evaluations. Nielsen identified ten Usability Heuristics for User Interface Design, these are listed below [104].

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

Cognitive walkthroughs are used for interaction design mostly, while heuristic evaluations are used for interface design. Cognitive walkthrough is a more structured approach and it aims at analyzing every task associated with a system interaction and poses a set of questions aimed at addressing potential issues. Design questions have been proposed to guide this process [145] and include the following:

- Is the action the user is supposed to perform sufficiently clear?
- Is it likely that the user will carry out the task correctly?
- Will the user know when the following task can be carried out?
- Is feedback implemented such that the user knows when the task has been completed successfully?
- Does the user know how to carry out the given tasks?

Usability testing is a fundamental element of human-centered design (HCD) and it has been demonstrated that it increases efficiency, effectiveness and user satisfaction [173]. Effectiveness is defined as the completeness and accuracy with which users accomplish the task, quality of the user's solution and error rates are metrics for effectiveness [173]. Efficiency is defined as the relation between accuracy and completeness with which users accomplish a task and resources required to do so, task completion time and learning time are measures for efficiency. User satisfaction is based on how comfortable the system is for the user and on how positive the user's attitude is toward the use of the system [173]. Usability testing investigates how the design of a system can be improved based on the following dimensions [103]:

- Learnability
- Efficiency of Use

- Memorability
- Few and Noncatastrophic Errors
- Subjective Satisfaction

The system designed should be easy to learn and efficient to use by enabling high work-productivity. The steps required to be taken by the system should be easy to remember, moreover the system should allow the user to commit a low amount of errors and enable quick recovery. Finally, the user should be sufficiently satisfied with the system [96]. At least two usability tests have to be performed, ideally five. After each test, a design iteration is necessary and the system is refined, after that another test follows with new users [96]. Nielsen states that at least three to five iterations need to be performed and usually 25% to 40% of the improvements can be observed after one iteration [102]. Once the system design is close to being finalised, comprehensive tests and evaluations can be performed. Taking iterative steps avoids late design changes and hence costly impacts [173]. Designs with good usability can reduce errors, training time, fatigue and life cycle costs.

Comprehensive evaluations and controlled experiments are more cumbersome as they resemble real experiments. Comprehensive evaluations are more summative than usability tests. During controlled experiments the idea is to change one independent variable only, while controlling the other variables and see the effect on the dependent variables [96].

In-service evaluations are performed once the system has been released. Usually controlled experiments are difficult to perform at that stage hence descriptive evaluations are critical. A-B testing is a common in-service evaluation, where version A is compared to version B, these evaluations can involve thousands of people [96].

Expert reviews are useful as human factors specialists can identify design strengths and weaknesses and recommend improvements, these reviews essentially combine the two methods: heuristic evaluations with cognitive walkthrough. The review can be performed early on, on conceptual sketches and working prototypes [173].

G.7.2. Study Design

During an experiment the relationship between the independent variables and the changes in one or multiple dependent variables are investigated. Dependent variables include performance, situational awareness, workload and preference for instance. An experimental design includes the identification of independent variables controlled in the study and the measurement and observation of dependent variables or results of interest. It is important that so called confounding variables are accounted for. In fact, confounding variables can influence the causal relationship of interest and have to be controlled [96]. Experimental designs include [96]:

- One-factor design
- Multiple-factor design
- Between-subjects design
- Within-subjects design
- Mixed Designs

A very common type of experiment involves the comparison of two conditions, this is a so called one factor design. The two conditions often involve a control and a treatment, for instance driving a rover only (control) and driving a rover with AR support (treatment). Else two levels of treatment might be involved like driving a rover with voice command or with a hand-controller.

Multi-factor experiments on the other hand involve the evaluation of multiple factors set at multiple levels. A design that investigates two or more independent variables combining the different levels of each variable is a so called factorial design. This design allows to evaluate the effect of each independent variable but also how these interact with each other. This method has the advantage that it captures more options and hence renders it more realistic as the results are more general and it is possible to capture interactions. Selecting one compared to the other depends on what the type of study and how many factors are required to be tested [96].

Between-subjects study designs are such that different subjects test each condition, so that each person is only exposed to one condition or user interface. Between-subjects experiments minimize the learning effects and transfer of knowledge across conditions. The test sessions for between-subjects are shorter compared to within-subjects experiments for the subjects. Between-subject tests are easier to set up, especially when there are multiple independent variables. Randomization can be used to avoid order effects more easily with between-subjects than with within-subjects designs, however it is essential for both. Randomization counteracts order, learning and transfer effects [96].

Within-subjects study designs are such that the same subject tests all the conditions. This type of study design has the advantage that not many participants are required to render the results significant. A similar between-subjects designs requires twice the amount of participants compared to within-subjects studies. A within-subject design can help reducing errors related to individual subject differences. In a between-subject design individuals are randomly assigned to the independent variable or condition, therefore there is the possibility that fundamental differences between the groups impacts the experiment's results. In a within-subject design, individuals are exposed to all conditions, hence individual differences do not distort the results. Disadvantages of within-subjects tests are fatigue and boredom induced in the subject after too many tests. Results on subject performance might be negatively affected by due practice effects, as the subject becomes more skilled after taking the test multiple times [96].

Mixed design studies involve at least one between-subjects factor and at least one within-subjects factor [96]. This type of study design combines the advantages of within-subjects factors, namely greater statistical power and between-subjects factors, namely less risk of subjects discovering the hypothesis.

It should be noted that after defining the independent variables, people need to be selected based on the principle of representative sampling. It is important that participants represent the group of people one is interested in analysing [96].

To get a better understanding, the study designs adopted by the studies reviewed in section G.2 were investigated.

The study by Braly, Nuernberger, and Young Kim [15] involved between-subject studies. In total 30 trials from three critical areas on the instrument, namely CPU, science module and electrical filter were carried out. 50% of the trials were assigned as mate trials and 50% as demate trials in a random manner. After that, two sets of trials were created to ensure that test subjects did not carry out the same trials in every instruction method (AR and paper) [15]. The random selection for each set of trial was done according to these constraints: the same amount of mate and demate trials as well as trials from each of the three areas and number of connection types had to be fulfilled. Finally, the order of the 15 trials in every set was randomised. Test subjects were randomly allocated to one of two instruction method orders defining which instruction method was used first by the subjects [15].

The study by Markov-Vetter and Staadt [160] on the other hand involved a within-subject study, test subjects were exposed to two different instruction method conditions (PDF, AR), each subject performed the procedure for every independent variable in all levels resulting in a 2x2 factorial design. Fatigue effects were taken into account and mitigated [160]. A within-subject study was carried out to evaluate the stowage application developed in the study of Furuya et al. [79]. Nevertheless, between-subject pilot studies with 25 individuals including experts were carried out as well prior to the final experiments [79].

To the best of my knowledge, other AR projects involving space applications such as EdcAR [124], WEKIT[34], mobiPV [3], WEAR [220] and T2 [230] performed user reviews sessions to evaluate the usability of the system. For ARPASS [198] online surveys and user reviews on desirability and feasibility were performed. While no information on the types of evaluations performed for the projects ARAMIS [72], Sidekick [175], MOON [114] and OnSight [175] could be found.

From the survey performed by Dey et al. [7] it was concluded that most of the AR studies, across all disciplines, were designed as within-subjects studies, namely 73.7% of the studies while 14.1% where designed as between-subjects studies and 4.3% as mixed studies.

G.7.3. Measurement

Many usability metrics including effectiveness, efficiency and satisfaction can be used. As mentioned earlier, effectiveness is defined as the accuracy and completeness with which a goal is achieved. Quality of the user's solution and error rates are effectiveness indicators [173]. Efficiency is defined as the relation between accuracy and completeness with which users achieve a preset goal and the resources used to achieve it. Task completion time and learning time are indicators of efficiency [173]. Satisfaction is defined as the users' comfort with respect to the use of the system as well as the users' positive attitudes toward this. To assess satisfaction user ratings and surveys addressing satisfaction with the interface are used.

The System Usability Scale (SUS) [93], [2], a 5 point Likert scale, is a common tool used to assess user satisfaction. The SUS has been used for WEKIT's [34] evaluation studies for instance, see Figure G.12 and Figure G.13. To calculate the SUS score, one needs to sum the score contributions from every item. Score contribution range from 0 to 4. For item: 1, 3, 5, 7 and 9 the score contribution equals the scale position minus 1. For item: 2, 4, 6, 8 and 10, the score contribution equals 5 minus the scale position [93]. Finally, the SU result is obtained by multiplying the sum of the scores by 2.5. Scores range from 0 to 100 [93]. Validation studies concluded that an SUS score starting from 68-70 stands for an acceptable system usability level. The Acceptability ranges are as follows: 0-50 not acceptable; 50-70 marginal; 70-acceptable [2, 35, 94].

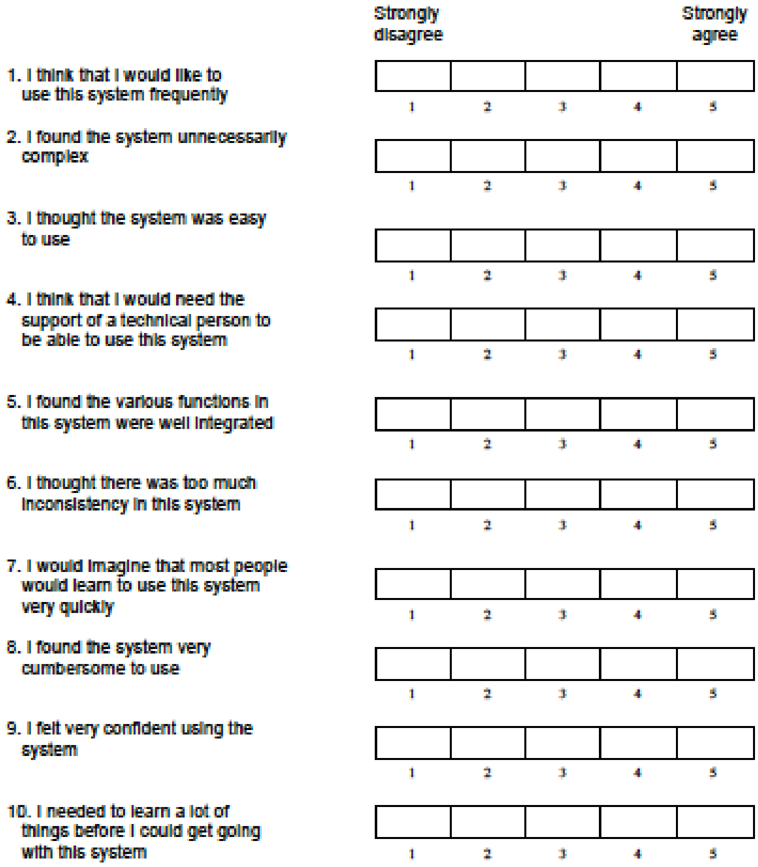


Figure G.12: System Usability Scale [93]

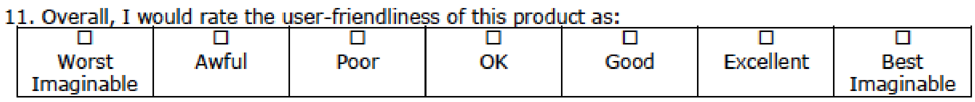


Figure G.13: Adjective Rating Scale added to the System Usability Scale [2]

Other questionnaires relevant to AR studies include the Questionnaire User Interaction Satisfaction (QUIS) and the Smart Glasses User Satisfaction questionnaire (SGUS) (please refer to Figure F.1 in Appendix F to view the SGUS), both were used as evaluation methods for WEKIT [128]. The SGUS was specifically created for the WEKIT project, it consists of 11 items and it is a 7 point Likert scale as can be seen in Figure F.1. To get insights on the results, the overall average of all items is calculated together with the averages of the individual items. Every individual average is then compared to the overall average and based on this comparison and the individual average scores conclusions are drawn.

It should be noted that the three metrics, efficiency, effectiveness and satisfaction have a low correlation among them [113], making all three measurements appropriate.

To the best of my knowledge most space-related AR projects (see subsection G.2.2) mainly used direct observations and user feedback rather than questionnaires, surveys or user rating scales (e.g. MobiPV [3], WEAR [220], T2 [230] and ARPASS [198]). It should be noted however that on project MOON [114], OnSight [175] and Sidekick [175] no information could be found on the evaluation methods.

Subjective and objective measurements can be distinguished. Subjective measurements include workload, situation awareness, fatigue, safety, acceptance, trust and comfort. Objective measurements include task completion times and accuracy/error rates. The differences between the type of measures should be considered, as for instance subjective measures are easier to retrieve and are less expensive. There are certain experimental contexts where subjective measures are better predictors while in other cases the subjective measures are not ideal as performance is not a well represented feature. An example of this situation, could be that following a subjective evaluation people prefer color to monochromatic displays while this choice according to objective measures undermines performance [6]. Nevertheless, the two measures are often complementary [96].

Based on the survey performed by Dey et al. [7], the most common objective measures were error rates or accuracy and task completion time. Error rate or accuracy was the objective measure chosen in the following space-related studies, mentioned in subsection G.2.2, by Braly, Nuernberger, and Young Kim [15], Markov-Vetter and Staadt [160], Tang et al. [18] and Uva et al. [9], while task completion time was selected in the following space-related studies by Braly, Nuernberger, and Young Kim [15], Furuya et al. [79], Boyd et al. [3]. Other studies evaluating the effectiveness of AR related to maintenance and repair used time and error rates as objective measures [210], similarly did a study evaluating the effectiveness of AR for procedural tasks [211]. Other objective measures identified in the survey by Dey et al. include: response rate, distance walked, head movement and physiological measures [7], however these were specifically related to the nature of the experiment.

Subjective measures include workload for instance. Workload is defined as the mental and physical effort required to meet the demands of the assigned task [173]. It appears to be a relevant measure for the proposed research, hence ways to assess this measure will be investigated.

NASA-STD-3001, volume 2 [178] specifies workload requirements. With respect to cognitive workload it is stated in V2 5007 that:

> *"Cognitive workload shall be accommodated in the design of all system elements that interface with the crew for all levels of crew capability and all levels of task demands"*

Workload is usually obtained through user ratings, the user is requested to rate the workload on a subjective scale. Subjective scales can be divided into two categories: unidimensional and multidimensional ratings [216]. Unidimensional rating scales are said to be simple to use as they do not involve complicated analysis techniques. These scales have only one dimension as the name suggests, nevertheless generally unidimensional scales are more sensitive than the multidimensional ones [42]. The multidimensional workload scale represents a more complex and time consuming form of measurement. It has from three to six dimensions. Unidimensional rating scales have not been given much validity in past research since they are often considered too simple to measure the complexity of workload. Despite this, through further analysis unidimensional scales acquired validity and managed to outperform multidimensional scales. Examples of unidimensional scales are the modified Cooper-Harper Scale (MCH) and the Overall Workload Scale (OW) [216].

Multidimensional scales are the most widely accepted form of measuring workload using subjective means. The two main multidimensional measures currently being used in the real-world and simulated environment are the NASA-Task Load Index (NASA-TLX) scale, and the Subjective Workload Assessment Technique (SWAT) [216]. There are also several other scales that are less well known [216].

The NASA Task Load Index (TLX) is a frequently used subjective measurement tool (see Figure G.14), used to assess workload, it consists of six subscales representing independent clusters of variables, three dimensions are related to the demands imposed on the subject (Mental, Physical, and Temporal Demands) and the other three relate to the interaction of the subject with the task (Effort, Frustration, and Performance) [206]. Scale ratings are score based, the user can give a score from 0 to 100 and the weights range from 0 to 5. Weighted averages are computed by multiplying the raw score of each scale by the number of times the associated workload factor was chosen in the paired-choice task, then dividing by the sum of the weights (i.e., 15). Unweighted/Raw TLX scores are also commonly published [206].

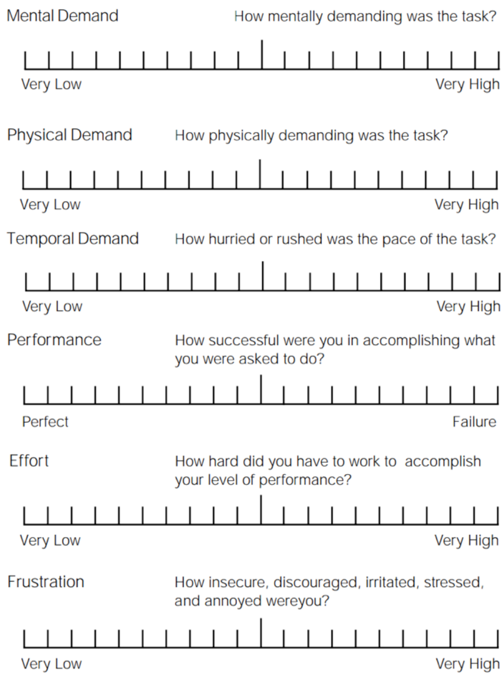


Figure G.14: NASA Task Load Index (TLX) scale [206]

The Subjective Workload Assessment Technique (SWAT) uses three levels, low, medium and high for each of the three dimensions of time load, mental load and physiological stress load to assess workload [207]. The SWAT technique scales the measurement scores to generate one rating scale with interval properties [216]. The technique aims at gaining insights into the way the human uses resources to process information [216]. When the SWAT scale is compared to the NASA-TLX, the latter is generally considered to be the better scale to measure mental workload [207] [185].

As mentioned earlier there are many less well known subjective scales used for the measurement of various types of workload, such as Cooper-Harper rating scale which is unidimensional and mainly used for measurement of psychomotor workload of pilots [216]. The Rating Scale Mental Effort (RSME) is another example of unidimensional scale, it consists of a line with a length of 150 mm marked with nine anchor points, each with a descriptive label indicating a degree of effort [17].

NASA uses the unidimensional Bedford scale [11] to assess workload [173] (see Figure G.15). The scale is characterized by a decision-tree format and was developed for test pilots with their help. The subjects are required to start answering questions from the bottom left corner and each question in order. This scale is feasible to carry out verification processes as it provides anchors for every rating, it is familiar to crew, it also gives a decision gate in which ratings above this gate gives evidence that workload is not satisfactory without a reduction in spare capacity or not tolerable [173].

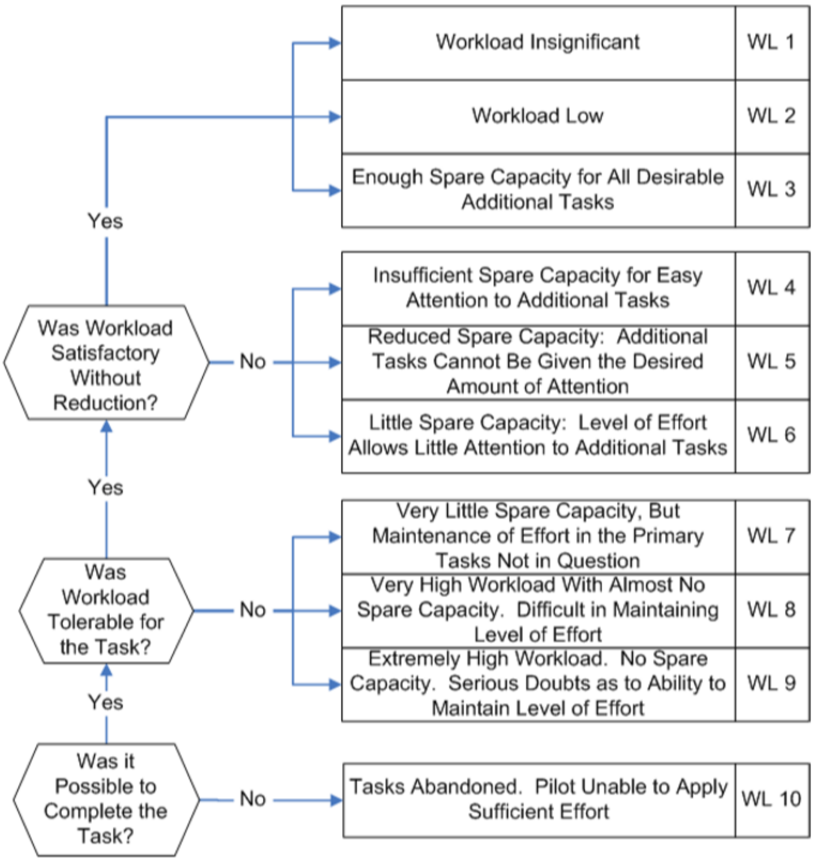


Figure G.15: Bedford scale [11], [173]

Bibliography

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash . Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys and Tutorials*, 17(4): 2347–2376, 2015. doi: 10.1109/COMST.2015.2444095.

[2] A. Bangor, P. T. Kortum and J. T. Miller . An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6):574–594, 2008. doi: 10.1080/10447310802205776.

[3] A. Boyd, A. Fortunato, M. Wolff, and D. Martinez-Oliveira. mobiPV: A new, wearable real-time collaboration software for Astronauts using mobile computing solutions. May 2016. doi: 10.2514/6.2016-2306.

[4] A. Butz, J. Baus, A. Krüger, and M. Lohse. A hybrid indoor navigation system. In *International Conference on Intelligent User Interfaces, Proceedings IUI*, 2001. doi: 10.1145/359784.359832.

[5] A.-C. Haugstvedt and J. Krogstie. Mobile augmented reality for cultural heritage: A technology acceptance study. pages 247–255, 2012. doi: 10.1109/ISMAR.2012.6402563.

[6] A. D. Andre and C. D. Wickens. When Users Want What's not Best for Them. *Ergonomics in Design*, 3(4): 10–14, 1995. doi: 10.1177/106480469500300403.

[7] A. Dey, M. Billinghurst, R. W. Lindeman, and J. E. Swan. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Frontiers in Robotics and AI*, 5:37, 2018. ISSN 2296-9144. doi: 10.3389/frobt.2018.00037. URL https://www.frontiersin.org/article/10.3389/frobt.2018.00037.

[8] A. Di Serio, M. B. Ibanez and C.D. Kloos. Impact of an augmented reality system on students' motivation for a visual art course. *Computers and Education*, 68:586–596, 2013. doi: 10.1016/j.compedu.2012.03.002.

[9] A. E. Uva, M. Gattullo, V. M. Manghisi, D. Spagnulo, G. L. Cascella and, M. Fiorentino. Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. *The International Journal of Advanced Manufacturing Technology*, 94(1-4):509, 2018. doi: 10.1007/s00170-017-0846-4.

[10] A. Finkbeiner. Inside the Breakthrough Starshot Mission to Alpha Centauri, March 2017. URL https://www.scientificamerican.com/article/inside-the-breakthrough-starshot-mission-to-alpha-centauri/. [Accessed: 2019-11-10].

[11] A. H. Roscoe. In-flight assessment of workload using pilot ratings and heart rate. 1987.

[12] A. H. Stevens et al. Tactical Scientific Decision-Making during Crewed Astrobiology Mars Missions. *Astrobiology*, 19(3):369–386, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1837.

[13] A. Kak, E. Guven, U. E. Ergin, and I. Akyildiz. Performance Evaluation of SDN-Based Internet of Space Things. pages 1–6, December 2018. doi: 10.1109/GLOCOMW.2018.8644237.

[14] A. L. Brady et al. Strategic Planning Insights for Future Science-Driven Extravehicular Activity on Mars. *Astrobiology*, 19(3):347–368, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1850.

[15] A. M. Braly, B. Nuernberger and S. Young Kim. Augmented Reality Improves Procedural Work on an International Space Station Science Instrument. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 61(6):866–878, September 2019. doi: 10.1177/0018720818824464.

[16] A. Pagani, J. Henriques, and D. Stricker. Sensors for location-based augmented reality the example of galileo and egnos. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2016. doi: 10.5194/isprsarchives-XLI-B1-1173-2016.

[17] A. Sartang, M. Ashnagar, E. Habibi, and S. Sadeghi. Evaluation of Rating Scale Mental Effort (RSME) effectiveness for mental workload assessment in nurses. *Journal of Occupational Health and Epidemiology*, 5:211–217, October 2016. doi: 10.18869/acadpub.johe.5.4.211.

[18] A. Tang, C. Owen, F. Biocca and W. Mou. Comparative effectiveness of augmented reality in object assembly. *Conference on Human Factors in Computing Systems - Proceedings*, pages 73–80, January 2003. doi: 10.1145/642611.642626.

[19] A. W. Johnson, J. A. Hoffman, D. J. Newman, E. M. Mazarico, and M. T. Zuber. An integrated Traverse Planner and Analysis tool for planetary exploration. In *AIAA SPACE Conference and Exposition 2010*, 2010. ISBN 9781600869662. doi: 10.2514/6.2010-8829.

[20] A. Y. C. Nee, S. K. Ong, G. Chryssolouris, and D. Mourtzis. Augmented reality applications in design and manufacturing. *CIRP Annals - Manufacturing Technology*, 61(2):657–679, 2012. doi: 10.1016/j.cirp.2012.05.010.

[21] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1):22–32, 2014. doi: 10.1109/JIOT.2014.2306328.

[22] I. Andrea, C. Chrysostomou, and G. Hadjichristofi. Internet of things: Security vulnerabilities and challenges. In *2015 IEEE Symposium on Computers and Communication (ISCC)*, pages 180–187, July 2015. doi: 10.1109/ISCC.2015.7405513.

[23] ARToolKit. URL http://www.artoolkitx.org/. [Accessed: 2019-11-08].

[24] ATLAS.ti. URL https://atlasti.com. [Accessed: 2020-05-07].

[25] B. Ens, J. D. Hincapie-Ramos, and P. Irani. Ethereal Planes: A Design Framework for 2D Information Spaces in 3D Mixed Reality Environments. 10 2014. doi: 10.1145/2659766.2659769.

[26] B. Shneiderman and C. Plaisant. Designing the user interface 4th ed. January 2005.

[27] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57(4):247–262, 2002. ISSN 1071-5819. doi: 10.1006/ijhc.2002.1017. URL http://www.sciencedirect.com/science/article/pii/S1071581902910177.

[28] C. Carter and R. Graham. Experimental comparison of manual and voice controls for the operation of in-vehicle systems. In *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Association, 'Ergonomics for the New Millennium'*, 2000. doi: 10.1177/154193120004402016.

[29] C. D. Wickens and C. M. Carswell. The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, 37(3):473–494, 1995. doi: 10.1518/001872095779049408.

[30] C. D. Wickens and J. Long. Object versus space-based models of visual attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied*, 1(3):179–193, 1995.

[31] C. E. Stewart. Chapter 22 - Extravehicular Activity Safety. In Gary Eugene Musgrave, Axel (Skip) M Larsen, and Tommaso Sgobba, editors, *Safety Design for Space Systems*, pages 705–723. Butterworth-Heinemann, Burlington, 2009. ISBN 978-0-7506-8580-1. doi: 10.1016/B978-0-7506-8580-1.00022-1. URL http://www.sciencedirect.com/science/article/pii/B9780750685801000221.

[32] C. Garcia-Galan, S. Uckun, W. Gregory, and K. Williams. Advance Technologies for Future Spacecraft Cockpits And Space-based Control Centers.

[33] C. M. de Morais, D. Sadok, and J. Kelner. An IoT sensor and scenario survey for data researchers. *Journal of the Brazilian Computer Society*, 25(1), 2019. doi: 10.1186/s13173-019-0085-7.

[34] C. Vizzi, K. Helin and J. Karjalainen. Exploitation of augmented reality for astronaut training. 2017. This article was not included in the published proceedings;European Association for Virtual Reality and Augmented Reality Conference, EuroVR-2017, EuroVR-2017; Conference date: 12-12-2017 Through 14-12-2017.

[35] C. Vizzi, T. Kuula, P. Sharma, F. Wild, B. Limbu, M. Fominykh. D6.6 Implementation of Evaluation Trials in Space, 2017. URL http://wekit.eu/wp-content/uploads/2017/09/WEKIT_D6.6.pdf.

[36] T. P. Caudell and D. W. Mizell. Augmented reality: an application of heads-up display technology to manual manufacturing processes. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, volume II, pages 659–669, January 1992. doi: 10.1109/HICSS.1992.183317.

[37] W. J. Clancey. Roles for agent assistants in field science: understanding personal projects and collaboration. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2): 125–137, May 2004. doi: 10.1109/TSMCC.2004.826284.

[38] D. B. Eppler. Lighting Constraints on Lunar Surface Operations, May 1991. URL https://www.lpi.usra.edu/lunar/strategies/human_ex/lighting_constraints.pdf. [Accessed: 2020-03-06].

[39] D. C. Foyle, A. D. Andre, and B. L. Hooey. Situation Awareness in an Augmented Reality Cockpit: Design, Viewpoints and Cognitive Glue. *Proceedings of the 11th International Conference on Human Computer Interaction*, 2005.

[40] D. Datcu and S. Lukosch. Free-hands interaction in augmented reality. In *SUI 2013 - Proceedings of the ACM Symposium on Spatial User Interaction*, 2013. ISBN 9781450321419. doi: 10.1145/2491367.2491370.

[41] D. Datcu, S. Lukosch, and F. Brazier. On the Usability and Effectiveness of Different Interaction Types in Augmented Reality. *International Journal of Human-Computer Interaction*, 2015. ISSN 15327590. doi: 10.1080/10447318.2014.994193.

[42] D. De Waard. The measurement of drivers' mental workload. 1996.

[43] D. Fonseca, E. Redondo and, S. Falip. Mixed-methods research: a new approach to evaluating the motivation and satisfaction of university students using advanced visual technologies. *Universal Access in the Information Society*, 14, July 2014. doi: 10.1007/s10209-014-0361-4.

[44] D. Furio, S. Gonzalez-Gancedo, M.-C. Juan, I. Segui, and N. Rando. Evaluation of learning outcomes using an educational iPhone game vs. traditional game. *Computers & Education*, 64:1–23, 2013. ISSN 0360-1315. doi: 10.1016/j.compedu.2012.12.001. URL http://www.sciencedirect.com/science/article/pii/S0360131512002850.

[45] D. J. Mayhew. The usability engineering lifecycle. In *CHI 98 Conference Summary on Human Factors in Computing Systems*, CHI '98, pages 127–128, New York, NY, USA, 1998. ACM. ISBN 1-58113-028-7. doi: 10.1145/286498.286575.

[46] D. Jo and G. J. Kim. IoT + AR: pervasive and augmented environments for "Digi-log" shopping experience. *Human-centric Computing and Information Sciences*, 9(1), 2019. doi: 10.1186/s13673-018-0162-5.

[47] D. Ledo, S. Greenberg, N. Marquardt, and S. Boring. Proxemic-Aware Controls: Designing Remote Controls for Ubiquitous Computing Ecologies. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pages 187–198, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3652-9. doi: 10.1145/2785830.2785871. URL http://doi.acm.org/10.1145/2785830.2785871.

[48] D. Newman, and M. Barratt. LIFE SUPPORT AND PERFORMANCE ISSUES FOR EXTRAVEHICULAR ACTIVITY (EVA). doi: 10.1.1.583.4840. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.4840&rep=rep1&type=pdf.

[49] D. Norman. The design of everyday things. pages 188–189, 1990.

[50] D. R. Williams. The Apollo Lunar Roving Vehicle, May 2016. URL https://nssdc.gsfc.nasa.gov/planetary/\lunar/apollo_lrv.html.

[51] D. S. S. Lim et al. The BASALT Research Program: Designing and Developing Mission Elements in Support of Human Scientific Exploration of Mars. *Astrobiology*, 19(3):245–259, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1869.

[52] D. Tedone et al. Augmented reality for the support of space exploration missions and on-ground/on-orbit operations. In Isabel L. Nunes, editor, *Advances in Human Factors and System Interactions*, pages 141–151, Cham, 2017. Springer International Publishing. ISBN 978-3-319-41956-5.

[53] D. Wagner and D. Schmalstieg. First steps towards handheld augmented reality. In *Proceedings - International Symposium on Wearable Computers, ISWC*, 2003. ISBN 0769520340. doi: 10.1109/iswc.2003.1241402.

[54] D. Woods, E. S. Patterson, J. M. Corban, J. Watt-Englert. Bridging the gap between user-centered intentions and actual design practice. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 40:967–971, October 1996. doi: 10.1177/154193129604001903.

[55] E. Anandapadmanaban, J. Tannady, J. Norheim, D. J. Newman, J. Hoffman. Holo-SEXTANT: an Augmented Reality Planetary EVA Navigation Interface. In *48th International Conference on Environmental Systems*, 2018.

[56] E. Graham-Rowe et al. Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: A qualitative analysis of responses and evaluations. *Transportation Research Part A: Policy and Practice*, 2012. ISSN 09658564. doi: 10.1016/j.tra.2011.09.008.

[57] E. M. Jones. Apollo Lunar Surface Journal, 1999. URL https://www.hq.nasa.gov/alsj/.

[58] E. McLellan, K. M. MaCqueen and J. L. Neidig. Beyond the Qualitative Interview: Data Preparation and Transcription. *Field Methods*, 15(1):63–84, 2003. ISSN 15523969. doi: 10.1177/1525822X02239573.

[59] Eclipse Paho. Eclipse Paho MQTT. URL https://www.eclipse.org/paho/. [Accessed: 2020-04-02].

[60] ESA. Horizons mission - Installing life-support system with astronaut aid mobiPV. URL https://www.esa.int/ESA_Multimedia/Videos/2018/12/Horizons_mission_Installing_life-support_system_with_astronaut_aid_mobiPV. [Accessed: 2019-11-08].

[61] ESA. Saving time in space. URL https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Research/Saving_time_in_space. [Accessed: 2019-11-08].

[62] ESA. Pangaea-X 2018, 2018. URL https://www.flickr.com/photos/europeanastronauttraining/45088839325/. [Accessed: 2020-06-08].

[63] ESA. IN-SITU RESOURCE UTILISATION (ISRU) DEMONSTRATION MISSION. September 2019. URL https://exploration.esa.int/s/8dlKaj8. [Accessed: 2019-04-19].

[64] Espressif. ESP8266. URL https://www.espressif.com/en/products/hardware/esp8266ex/overview. [Accessed: 2020-04-17].

[65] G. Turini et al. A Microsoft HoloLens Mixed Reality Surgical Simulator for Patient-Specific Hip Arthroplasty Training. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018. ISBN 9783319952819. doi: 10.1007/978-3-319-95282-6_15.

[66] F. A. A. S. D. T. Rometsch. MARVIN Project - Internship Report, 2019.

[67] F. H. Mohammed and R. Esmail. Survey on IoT Services: Classifications and Applications. 2015.

[68] F. J. Seagull and D. Gopher. Training head movement in visual scanning: An embedded approach to the development of piloting skills with helmet-mounted displays. *Journal of Experimental Psychology: Applied*, 3(3):163–180, 1997.

[69] F. P. Vidal et al. Principles and applications of computer graphics in medicine. *Computer Graphics Forum*, 25(1):113–137, 2006. doi: 10.1111/j.1467-8659.2006.00822.x.

[70] G. Alce and M. Roszko and H. Edlund and S. Olsson and J. Svedberg and M. Wallergard. [POSTER] AR as a User Interface for The Internet of Things-Comparing Three Interaction Models. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 81–86, October 2017. doi: 10.1109/ISMAR-Adjunct.2017.37.

[71] G. B. Dalrymple. The age of the Earth in the twentieth century: A problem (mostly) solved. *Geological Society Special Publication*, 2001. ISSN 03058719. doi: 10.1144/GSL.SP.2001.190.01.14.

[72] G. Lentini, E. Afelli, A. Piras, R. Sapone, M. Crisconio, and G. Valentini. ARAMIS - Augmented reality application for maintenance, inventory and stowage. volume 2018-October, 2018.

[73] G. M. Sullivan and A. R. Artino. Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, 5(4):541–542, 2013. ISSN 1949-8349. doi: 10.4300/jgme-5-4-18.

[74] G. Q. Huang, X. Chen, T. Qu, J. H. Wang, and H. Fu. System dynamics approach for the performance evaluation of internet-of-things system implementation in a dynamic production logistics system. 2016.

[75] G. Reinhart and C. Patron. Integrating Augmented Reality in the Assembly Domain - Fundamentals, Benefits and Applications. *CIRP Annals*, 52(1):5–8, 2003. ISSN 0007-8506. doi: 10.1016/S0007-8506(07)60517-4.

[76] G. Reitmayr and D. Schmalstieg. Location based applications for mobile augmented reality. In *AUIC '03*, pages 65–73, 2003.

[77] G. White, C. Cabrera, A. Palade, and S. Clarke. Augmented Reality in IoT. September 2018. doi: 10.1007/978-3-030-17642-6_13.

[78] H. B. Booher. RELATIVE COMPREHENSIBILITY OF PICTORIAL INFORMATION AND PRINTED WORDS IN PROCEDURALIZED INSTRUCTIONS. *Human Factors*, 1975. ISSN 00187208. doi: 10.1177/001872087501700306.

[79] H. Furuya, L. Wang, C. Elvezio, and S. Feiner. A comparative ground study of prototype augmented reality task guidance for international space station stowage operations. volume 2018-October, 2018.

[80] H. H. Schmitt. Apollo 17 and the Moon. In *Encyclopedia of Space Science and Technology*. 2003. doi: 10.1002/0471263869.sst066.

[81] H. H. Schmitt, A. W. Snoke, M. A. Helper, J. M. Hurtado, K. V. Hodges and J.W. Rice. Motives, methods, and essential preparation for planetary field geology on the Moon and Mars. *Special Paper of the Geological Society of America*, 483(01):1–15, 2011. ISSN 00721077. doi: 10.1130/2011.2483(01).

[82] H. Xue, P. Sharma, and F. Wild. User Satisfaction in Augmented Reality-Based Training Using Microsoft HoloLens. 10 2018. doi: 10.20944/preprints201810.0594.v1.

[83] Hiber. The first truly global IoT network. And everyone's welcome, 2017. URL https://hiber.global. [Accessed: 2019-10-15].

[84] I. Christofilos, M. Haunschild, F. Huber and, E. Messerschmid. TELDASAT - Satellitengestuetzte globale Vernetzung des "Internet of Things" IoT. 2017. Deutscher Luft- und Raumfahrtkongress 2017, Muenchen.

[85] I. F. Akyildiz and A. Kak. The internet of space things/cubesats: A ubiquitous cyber-physical system for the connected world. *Computer Networks*, 150:134–149, 2019. ISSN 1389-1286. doi: 10.1016/j.comnet.2018.12.017. URL http://www.sciencedirect.com/science/article/pii/S1389128618314191.

[86] I. F. del Amo, J. A. Erkoyuncu, R. Roy, R. Palmarini, and D. Onoufriou. A systematic review of augmented reality content-related techniques for knowledge transfer in maintenance applications. *Computers in Industry*, 103:47–71, 2018. ISSN 0166-3615. doi: 10.1016/j.compind.2018.08.007.

[87] IBM SPSS Software. URL https://www.ibm.com/nl-en/analytics/spss-statistics-software. [Accessed: 2020-05-07].

[88] IconFinder. URL https://www.iconfinder.com. [Accessed: 2020-04-02].

[89] Icons For Free. URL https://icons-for-free.com. [Accessed: 2020-04-02].

[90] Interaction Design Foundation. How to Use Voice Interaction in Augmented Reality. URL https://www.interaction-design.org/literature/article/how-to-use-voice-interaction-in-augmented-reality. [Accessed: 2020-07-02].

[91] International Space Exploration Coordination Group (ISECG). The Global Exploration Roadmap. pages 1–36, January 2018.

[92] ISO. ISO 9241-210:2019 Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems, 2019. URL https://www.iso.org/standard/77520.html.

[93] J. Brooke. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 1996.

[94] J. Brooke. SUS: a retrospective. *Journal of Usability Studies*, 2013. ISSN 1931-3357.

[95] J. C. Jakobsen, C. Gluud, J. Wetterslev, and P. Winkel. When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Medical Research Methodology*, 17(1):1–10, 2017. ISSN 14712288. doi: 10.1186/s12874-017-0442-1.

[96] J. D. Lee, C. D. Wickens, Y. Liu, and L. N. Boyle. *Designing for People: An introduction to human factors engineering*. August 2017. ISBN 1539808009.

[97] J. E. Swan and J. L. Gabbard. Survey of user-based experimentation in augmented reality. 2005.

[98] J. J. Marquez et al. Future Needs for Science-Driven Geospatial and Temporal Extravehicular Activity Planning and Execution. *Astrobiology*, 19(3):440–461, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1838.

[99] J. Karasinski, R. Joyce, C. Carroll, J. Gale, and S. Hillenius. An Augmented Reality/Internet of Things Prototype for Just-in-time Astronaut Training. pages 248–260, May 2017. ISBN 978-3-319-57986-3. doi: 10.1007/978-3-319-57987-0_20.

[100] J. L. Gabbard and J. E. Swan II. Usability engineering for augmented reality: Employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):513–525, May 2008. doi: 10.1109/TVCG.2008.24.

[101] J. M. Owens, S. B. McLaughlin, and J. Sudweeks. On-road comparison of driving performance measures when using handheld and voice-control interfaces for mobile phones and portable music players. In *SAE Technical Papers*, 2010. doi: 10.4271/2010-01-1036.

[102] J. Nielsen. Iterative user-interface design. *Computer*, 26(11):32–41, November 1993. doi: 10.1109/2.241424.

[103] J. Nielsen. Usability engineering. Academic Press, San Francisco, CA, 1993.

[104] J. Nielsen. How to conduct a heuristic evaluation. 1994. URL https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/. [Accessed: 2020-01-19].

[105] J. Nielsen. 10 Heuristics for User Interface Design, 1995.

[106] J. Nielsen. How many test users in a usability study? 2012. URL https://www.nngroup.com/articles/how-many-test-users/. [Accessed: 2020-01-10].

[107] J. Nielsen and R. Molich. Heuristic evaluation of user interfaces. In *Conference on Human Factors in Computing Systems - Proceedings*, 1990. ISBN 0201509326. doi: 10.1145/97243.97281.

[108] J. Norheim et al. Architecture of a surface exploration traverse analysis and navigational tool. In *IEEE Aerospace Conference Proceedings*, 2018. ISBN 9781538620144. doi: 10.1109/AERO.2018.8396510.

[109] J. Norris. Mixed Reality on the Space Station, August 2015. URL http://drjeffnorris.com/project/sidekick-holograms-in-orbit.

[110] J. Okamoto and A. Nishihara. Assembly assisted by augmented reality (A3R). *In Y. Bi, S. Kapoor, R. Bhatia (Eds.), Intelligent systems and applications. Studies in computational intelligence. Cham: Springer*, 650: 281–300, 2016. doi: 10.1007/978-3-319-33386-1_14.

[111] J. Rambach, A. Pagani and D. Stricker. Augmented things: Enhancing ar applications leveraging the internet of things and universal 3d object tracking. pages 103–108, 2017. doi: 10.1109/ISMAR-Adjunct.2017.42.

[112] J. Rekimoto, and Y. Ayatsuka. CyberCode: Designing Augmented Reality Environments with Visual Tags. In *Proceedings of DARE 2000 on Designing Augmented Reality Environments*, DARE '00, pages 1–10, New York, NY, USA, 2000. ACM. doi: 10.1145/354666.354667. URL http://doi.acm.org/10.1145/354666.354667.

[113] J. Sauro and J. R. Lewis. Correlations Among Prototypical Usability Metrics: Evidence for the Construct of Usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1609–1618, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518947. URL http://doi.acm.org/10.1145/1518701.1518947.

[114] J. Servan, F. Mas, J. Menendez, and J. Rios. Using augmented reality in AIRBUS A400M shop floor assembly work instructions. volume 1431, pages 633–640, September 2011. doi: 10.1063/1.4707618.

[115] J. Zhu, S. K. Ong, and A. Y. C. Nee. An authorable context-aware augmented reality system to assist the maintenance technicians. *International Journal of Advanced Manufacturing Technology*, 2013. ISSN 02683768. doi: 10.1007/s00170-012-4451-2.

[116] J.D. Gould and C. Lewis. Designing for usability: Key principles and what designers think. *Commun. ACM*, 28(3):300–311, March 1985. ISSN 0001-0782. doi: 10.1145/3166.3170. URL http://doi.acm.org/10.1145/3166.3170.

[117] D. Jo and G. J. Kim. ARIoT: scalable augmented reality framework for interacting with Internet of Things appliances everywhere. *IEEE Transactions on Consumer Electronics*, 62(3):334–340, August 2016. doi: 10.1109/TCE.2016.7613201.

[118] Johnson Space Centre - NASA. Apollo Program Summary Report, 1975.

[119] Johnson Space Centre - NASA. Human Integration Design Handbook, 2014. URL https://www.nasa.gov/sites/default/files/atoms/files/human_integration_design_handbook_revision_1.pdf.

[120] E. M. Jones and K. Glover. Apollo Lunar Surface Journal, 1995. URL https://www.hq.nasa.gov/alsj/. [Accessed: 2019-10-06].

[121] K. H. Beaton et al. Using Science-Driven Analog Research to Investigate Extravehicular Activity Science Operations Concepts and Capabilities for Human Planetary Exploration. *Astrobiology*, 19(3):300–320, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1861.

[122] K. H. Beaton et al. Assessing the Acceptability of Science Operations Concepts and the Level of Mission Enhancement of Capabilities for Human Mars Exploration Extravehicular Activity. *Astrobiology*, 19(3): 321–346, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1912.

[123] K. H. Beaton et al. IAC-19-B3.5.7x50867 Immersive Mixed Reality Capabilities for Planning and Executing Exploration Extravehicular Activity Kara H. Beaton. pages 21–25, October 2019.

[124] K. Helin. Augmented reality for AIT, AIV and operations. 2017. URL http://old.esaconferencebureau.com/2017-events/17m38. SDA: SHP: ForIndustry ; Space Engineering and Technology Final Presentation Days, SET-FPDs ; Conference date: 23-05-2017 Through 24-05-2017.

[125] K. Helin. MobiPV4Hololens 4000125238/18/NL/AF/ as Prototype a Media Helmet for MobiPV Implemented Using MS Hololens - Executive Summary Report, April 2019.

[126] K. Helin, J. Karjalainen, P. Kiernan, M. Wolff, and D. M. Oliveira. Mixed reality user interface for astronauts procedure viewer. *Proceedings - 2019 International Conference on Cyberworlds, CW 2019*, pages 17–20, 2019. doi: 10.1109/CW.2019.00011.

[127] K. Helin, J. Perret and V. Kuts, editor. *The application track, posters and demos of EuroVR: Proceedings of the 16th Annual EuroVR Conference - 2019*. Number 357 in VTT Technology. VTT Technical Research Centre of Finland, Finland, 2019. doi: 10.32040/2242-122X.2019.T357. 16th EuroVR International Conference-EuroVR 2019 ; Conference date: 23-10-2019 Through 25-10-2019.

[128] K. Helin, T. Kuula, C. Vizzi, J. Karjalainen, and A. Vovk. User experience of augmented reality system for astronaut's manual work support. *Frontiers in Robotics and AI*, 5:106, 2018. ISSN 2296-9144. doi: 10.3389/frobt.2018.00106. URL https://www.frontiersin.org/article/10.3389/frobt.2018.00106.

[129] K. Matthews. Stanford and NASA Launch Tiny IoT Satellites Into Earth's Orbit, July 2019. URL https://theiotmagazine.com/stanford-and-nasa-launch-tiny-iot-satellites-into-earths-orbit-9e5f92487500. [Accessed: 2019-10-23].

[130] K. V. Hodges and H. H. Schmitt. A new paradigm for advanced planetary field geology developed through analog experiments on Earth. *Special Paper of the Geological Society of America*, 483(02):17–31, 2011. ISSN 00721077. doi: 10.1130/2011.2483(02).

[131] Keyhole, Inc., Google. Keyhole Markup Language. URL https://developers.google.com/kml/. [Accessed: 2019-11-08].

[132] Kindpng. HoloLens (1st gen) clicker. URL https://www.kindpng.com/imgv/iwTwhbo_hololens-clicker-hd-png-download/. [Accessed: 2020-06-08].

[133] L. Atzori, A. Iera, and G. Morabito. The internet of things: A survey. *Computer Networks*, 54(15):2787–2805, 2010. doi: 10.1016/j.comnet.2010.05.010.

[134] L. Bessone, F. Sauro, M. Maurer and M. Piens. Technologies and operational concepts for field geology and exploration on the Moon: the ESA PANGAEA-eXstension campaign in Lanzarote (Canary Archipelago, Spain). 2018.

[135] L. Bessone, L. Turchi, F. Sauro, H. Stevenin, R. Pozzobon and J. Baravykas. An Electronic Fieldbook supporting data collection and situational awareness during astronauts EVA geologic traverses on the lunar surface. 2018.

[136] L. Billings. Reaching for the Stars, Breakthrough Sends Smallest-Ever Satellites into Orbit, July 2017. URL https://www.scientificamerican.com/article/reaching-for-the-stars-breakthrough-sends-smallest-ever-satellites-into-orbit/.

[137] L. Hou, X. Wang, L. Berhold, M. ASCE and P. E. D. Love. Using animated augmented reality to cognitively guide assembly. *Journal of Computing in Civil Engineering*, 27:439–451, 2013. doi: 10.1061/(ASCE)CP.1943-5487.0000184.

[138] L. Mueller, I. Aslan, and L. Kruessen. GuideMe: A Mobile Augmented Reality System to Display User Manuals for Home Appliances. 8253:152–167, January 2013. doi: 10.1007/978-3-319-03161-3_11.

[139] L. Ravagnolo et al. . Evaluation Process and Results for Space case, Deliverable D6.12, WEKIT Project (2019). URL http://wekit.eu/wpcontent/uploads/2019/04/WEKIT_D6.12.pdf.

[140] LoRa Alliance. LoRaWAN. URL https://lora-alliance.org/about-lorawan. [Accessed: 2020-04-17].

[141] Lunar and Planetary Institute. Apollo 17 mission. URL https://www.lpi.usra.edu/lunar/missions/apollo/apollo_17/surface_opp/. [Accessed: 2019-10-06].

[142] Lunar and Planetary Institute. Apollo 17 Mission - Photography Overview. URL https://www.lpi.usra.edu/lunar/missions/apollo/apollo_17/photography/. [Accessed: 2019-11-20].

[143] M. Billinghurst, H. Kato, and S. Myojin. Advanced interaction techniques for augmented reality applications. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009. ISBN 3642027709. doi: 10.1007/978-3-642-02771-0_2.

[144] M. Billinghurst, J. Bowskill, N. Dyer, and J. Morphett. Evaluation of wearable information spaces. In *Proceedings - Virtual Reality Annual International Symposium*, 1998. doi: 10.1109/vrais.1998.658418.

[145] M. Blackmon, M. Kitajima, and P. Polson. Repairing usability problems identified by the Cognitive Walkthrough for the Web. pages 497–504, January 2003. doi: 10.1145/642611.642698.

[146] M. Cardano, M. Ferrino, M. Costa, and P. Giorgi. VR/AR tools to support on orbit crew operations and P/Ls maintenance in the ISS pressurized columbus module. *60th International Astronautical Congress 2009, IAC 2009*, 5:4188–4195, 2009.

[147] M. F. Boucher, E. Hodgson, S. M. Murray, P. Lee, and S. Braham. Investigation of eva information interface technology in a mars analog arctic field science setting. URL https://spacecraft.ssl.umd.edu/design_lib/ICES02-2312.EVA_HUD.pdf.

[148] M. Gläser-Zikuda, G. Hagenauer and M. Stephan. The potential of qualitative content analysis for empirical educational research. *Forum Qualitative Sozialforschung*, 2020. ISSN 14385627. doi: 10.17169/fqs-21.1.3443.

[149] M. Haklay and A. Nivala. Understanding geospatial technologies. *John Wiley & Sons, Ltd*, pages 91–106, 2010. doi: 10.1002/9780470689813.

[150] M. Handosa. Hololens, Unity Plugins, and MQTT, September 2016. URL http://handosa.blogspot.com/2016/09/hololens-unity-plugins-and-mqtt.html.

[151] M. Kerin and D. T. Pham. A review of emerging industry 4.0 technologies in remanufacturing. *Journal of Cleaner Production*, 237, 2019. ISSN 0959-6526. doi: 10.1016/j.jclepro.2019.117805.

[152] M. Lee, M. Billinghurst, W. Baek, R. Green, and W. Woo. A usability study of multimodal input in an augmented reality environment. *Virtual Reality*, 2013. ISSN 13594338. doi: 10.1007/s10055-013-0230-0.

[153] M. M. Conners, D. T. Eppler, and D. G. Morrow. Interviews with the Apollo lunar surface astronauts in support of planning for EVA systems, NASA Technical Memorandum 108846. page 20, September 1994.

[154] M. M. Connors, D. B. Eppler and, D. G. Morrow. Interviews with the Apollo Lunar Surface Astronauts in Support of Planning for EVA Systems Design, 1994. URL https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19950005492.pdf. [Accessed: 2019-10-06].

[155] M. Richardson, G. Jones, and M. Torrance. Identifying the task variables that influence perceived object assembly complexity. *Ergonomics*, 2004. ISSN 00140139. doi: 10.1080/00140130410001686339.

[156] M. S. Hossain, and G. Muhammad. Cloud-assisted Industrial Internet of Things (IIoT) - Enabled framework for health monitoring. *Computer Networks*, 101:192–202, 2016. doi: 10.1016/j.comnet.2016.01.009.

[157] M. Vaismoradi, H. Turunen and T. Bondas. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study, 2013. ISSN 14410745.

[158] M. Wolff, D. Martinez-Oliveira, A. Fortunato, A. Boyd, and A. Cowley . Enhancement of the ESA mobile Procedure Viewer (mobiPV) beyond Low Earth Orbit. 68th International Astronautical Congress (IAC), September 2017.

[159] M. Yeh, C. Wickens, and D. Brandenburg. Head-up vs. head-down: Effects of precision on cue effectiveness and display signaling. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45, October 2001. doi: 10.1177/154193120104502707.

[160] D. Markov-Vetter and O. Staadt. A pilot study for augmented reality supported procedure guidance to operate payload racks on-board the international space station. pages 1–6. IEEE International Symposium on Mixed and Augmented Reality (ISMAR), October 2013. doi: 10.1109/ISMAR.2013.6671832.

[161] Microsoft. Getting Started with MRTK. URL https://microsoft.github.io/MixedRealityToolkit-Unity/Documentation/GettingStartedWithTheMRTK.html. [Accessed: 2019-11-08].

[162] Microsoft. HoloLens (1st gen) clicker, . URL https://docs.microsoft.com/en-us/hololens/hololens1-clicker. [Accessed: 2020-06-09].

[163] Microsoft. HoloLens (1st gen) hardware, . URL https://docs.microsoft.com/en-us/hololens/hololens1-hardware. [Accessed: 2020-06-08].

[164] Microsoft. Microsoft Hololens 2. URL https://www.microsoft.com/en-us/hololens/. [Accessed: 2019-09-24].

[165] Microsoft. HoloLens 1 gestures for authoring and navigating in Dynamics 365 Guides. URL https://docs.microsoft.com/en-us/dynamics365/mixed-reality/guides/authoring-gestures. [Accessed: 2020-06-08].

[166] Microsoft. Voice commanding. URL https://docs.microsoft.com/en-us/windows/mixed-reality/voice-design. [Accessed: 2020-06-08].

[167] N. G. Bailey and G. E. Ulrich. APOLLO 17 VOICE TRANSCRIPT - Pertaining to the geology of the landing site, 1975. URL https://www.hq.nasa.gov/alsj/a17/Apollo17VoiceTranscript-Geology.pdf.

[168] N. Mays and C. Pope. Rigour and qualitative research. 1995. ISSN 09598146. doi: 10.1136/bmj.311.6997.109.

[169] NASA. IN-SITU RESOURCE UTILISATION. URL https://www.nasa.gov/isru. [Accessed: 2020-04-19].

[170] NASA. NEEMO. URL https://www.nasa.gov/mission_pages/NEEMO/index.html. [Accessed: 2019-11-08].

[171] NASA. Apollo 15 - Lunar surface procedures. 1971.

[172] NASA. Lunar L 1 Gateway Conceptual Design Report. 2001.

[173] NASA. Human Integration Design Processes (HIDP), 2014. URL https://www.nasa.gov/sites/default/files/atoms/files/human_integration_design_processes.pdf.

[174] NASA. NASA Space Flight Human-System Standard Volume 2 : Human Factors , Habitability , and Environmental Health. *NASA Technical Standard 3001*, 2:1–10, 2015.

[175] NASA. NASA, Microsoft Collaborate to Bring Science Fiction to Science Fact, June 2015. URL https://www.nasa.gov/press-release/nasa-microsoft-collaborate-to-bring-science-fiction-to-science-fact. [Accessed: 2019-10-09].

[176] NASA. Augmented Reality Application for Maintenance, Inventory and Stowage, 2017. URL https://www.nasa.gov/mission_pages/station/research/experiments/explorer/Investigation.html?#id=2083. [Accessed: 2019-10-19].

[177] NASA. Cube Quest Challenge, 2017. URL https://www.nasa.gov/cubequest/details. [Accessed: 2019-10-15].

[178] NASA. EXTRAVEHICULAR ACTIVITY (EVA) - NASA-STD-3000, 2018. URL https://msis.jsc.nasa.gov/sections/section14.htm. [Accessed: 2019-11-04].

[179] NASA. What is Artemis?, July 2019. URL https://www.nasa.gov/what-is-artemis. [Accessed: 2019-11-12].

[180] NASA. Forward to the Moon: NASA's Strategic Plan for Human Exploration, 2019. URL https://www.nasa.gov/sites/default/files/atoms/files/america_to_the_moon_2024_artemis_20190523.pdf. [Accessed: 2020-06-02].

[181] National Aeronautic and Space Administration Space Station Program. Display and Graphics Commonality Standard - International Space Program (Revision F). 10, December 2001.

[182] OGC. Augmented Reality Markup Language 2.0. URL http://www.opengeospatial.org/standards/arml. [Accessed: 2019-11-08].

[183] OneWeb, 2017. URL http://oneweb.world/. [Accessed: 2019-10-15].

[184] P. Burnard, P. Gill, K. Stewart, E. Treasure, and B. Chadwick. Analysing and presenting qualitative data. *British Dental Journal*, 204(8):429–432, 2008. ISSN 00070610. doi: 10.1038/sj.bdj.2008.292.

[185] P. C. Park. Comparison of Subjective Mental Workload Assessment Techniques for the Evaluation of In-Vehicle Navigation System Usability. 1998.

[186] P. Mayring. *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz, 1983. doi: 10.17169/fqs-21.1.3443.

[187] P. Mayring. Qualitative Inhaltsanalyse Philipp Mayring. *Forum Qualitative Sozialforschung*, 2000.

[188] P. Patierno. M2Mqtt. URL https://m2mqtt.wordpress.com. [Accessed: 2019-11-01].

[189] P. Phupattanasilp and S. Tong. Augmented Reality in the Integrative Internet of Things (AR-IoT): Application for Precision Farming. *Sustainability*, 11(9):2658, May 2019. doi: 10.3390/su11092658.

[190] P. R. Cohen and S. L. Oviatt. The role of voice input for human-machine communication. *Proceedings of the National Academy of Sciences of the United States of America*, 1995. ISSN 00278424. doi: 10.1073/pnas.92.22.9921.

[191] P. Simoens, M. Dragone, and A. Saffiotti. The Internet of Robotic Things: A review of concept, added value and applications. *International Journal of Advanced Robotic Systems*, pages 1–11, February 2019. doi: 10.1177/1729881418759424.

[192] P. W. Jordan. An Introduction to Usability. 1998. doi: 10.1016/B978-0-240-81203-8.00002-7.

[193] PNG Image. URL https://pngimage.net. [Accessed: 2020-04-02].

[194] Pokemon Go. URL https://pokemongolive.com/en/. [Accessed: 2019-11-17].

[195] Q. Zhu, R. Wang, Q. Chen, Y. Liu, and W. Qin. IOT gateway: Bridging wireless sensor networks into Internet of Things. pages 347–352, 2010. doi: 10.1109/EUC.2010.58.

[196] R. A. Light. Mosquitto: server and client implementation of the MQTT protocol. volume 2, May 2017. doi: 10.21105/joss.00265.

[197] R. A. Scheuring et al. The Apollo Medical Operations Project: Recommendations to Improve Crew Health and Performance for Future Exploration Missions and Lunar Surface Operations, August 2007. URL https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20070030109.pdf.

[198] R. Alarcon. Augmented reality for the enhancement of space product assurance and safety. In *Proceedings of the International Astronautical Congress, IAC*, 2018.

[199] R. Gad. Bowerick. URL https://github.com/ruedigergad/bowerick. [Accessed: 2019-09-25].

[200] R. L. Balthazor, M. G. McHarg, C. S. Godbold, D. J. Barnhart, and T. Vladimirova. Distributed space-based Ionospheric Multiple Plasma Sensor networks. pages 1–10, 2009.

[201] R. T. Azuma. A survey of augmented reality. *Presence: Teleoper. Virtual Environ.*, 6(4):355–385, August 1997. ISSN 1054-7460. doi: 10.1162/pres.1997.6.4.355.

[202] Raspberry Pi. URL https://www.raspberrypi.org. [Accessed: 2020-04-02].

[203] S. Cope. Beginners Guide to the MQTT Protocol. URL http://www.steves-internet-guide.com/mqtt/. [Accessed: 20-10-03].

[204] S. Fadden, P. M. Ververs and C. D. Wickens. Pathway huds: Are they viable? *Human Factors*, 43(2):173–193, 2001. doi: 10.1518/001872001775900841.

[205] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the World: 2D Windows for 3D Augmented Reality. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, UIST '93, pages 145–155, New York, NY, USA, 1993. ACM. ISBN 0-89791-628-X. doi: 10.1145/168642.168657. URL http://doi.acm.org/10.1145/168642.168657.

[206] S. G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later, journal = Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50(9):904–908, 2006. doi: 10.1177/154193120605000909.

[207] S. G. Hill, H. P. Iavecchia, J. C. Byers, A. C. Bittner, A. L. Zaklad, and R. E. Christ. Comparison of Four Subjective Workload Rating Scales. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34:429–439, 1992. doi: 10.1177/001872089203400405.

[208] S. Gamm and R. Haeb-Umbach. User interface design of voice controlled consumer electronics. *Philips Journal of Research*, 1995. ISSN 01655817. doi: 10.1016/0165-5817(96)81590-7.

[209] S. J. Henderson and S. Feiner. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. pages 135–144, New York, NY: IEEE, 2008.

[210] S. J. Henderson and S. Feiner. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE Transactions on Visualization and Computer Graphics*, 17(10):1355–1368, October 2011. doi: 10.1109/TVCG.2010.245.

[211] S. J. Henderson and S. K. Feiner. Augmented reality in the psychomotor phase of a procedural task. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pages 191–200, Oct 2011. doi: 10.1109/ISMAR.2011.6092386.

[212] S. J. Payler et al. Developing Intra-EVA Science Support Team Practices for a Human Mission to Mars. *Astrobiology*, 19(3):387–400, 2019. ISSN 15311074. doi: 10.1089/ast.2018.1846.

[213] S. Jamieson. Likert scales: How to (ab)use them. *Medical Education*, 38(12):1217–1218, 2004. ISSN 03080110. doi: 10.1111/j.1365-2929.2004.02012.x.

[214] S. Jamieson. Analyse qualitative data. *Education for Primary Care*, 27(5):398–402, 2016. ISSN 14739879. doi: 10.1080/14739879.2016.1217430. URL http://dx.doi.org/10.1080/14739879.2016.1217430.

[215] S. K. Ong, M. L. Yuan and A. Y. C. Nee. Augmented reality applications in manufacturing: A survey. International Journal of Production Research. *In Proceedings of the IEEE 1998 Virtual Reality Annual International Symposium*, 46:2707–2742, 2008.

[216] S. Miller. LITERATURE REVIEW - Workload Measures, 2001. URL http://www.nads-sc.uiowa.edu/publicationstorage/200501251347060.n01-006.pdf.

[217] S. N. Roscoe. Airborne displays for flight and navigation. *Human Factors*, 10(4):321–332, 1968. doi: 10.1177/001872086801000402.

[218] S. Narayana, R. V. Prasad, V. S. Rao, and C. Verhoeven. SWANS: Sensor Wireless Actuator Network in Space. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, SenSys '17, pages 23:1–23:6, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5459-2. doi: 10.1145/3131672.3131701. URL http://doi.acm.org/10.1145/3131672.3131701.

[219] Smithsonian National Air and Space Museum. Cuff Checklist, EVA 2 & 3, Apollo 16 (Young). URL https://ids.si.edu/ids/deliveryService?id=https://airandspace.si.edu/webimages/collections/full/A20050109000cp01.jpg&max=3000. [Accessed: 2020-06-08].

[220] Space Applications Services and ESA . WEAR - Final Presentation, 2011. (confidential).

[221] Space Applications Services and ESA . WEAR, 2011. URL https://www.esa.int/gsp/ACT/doc/EVENTS/bmiworkshop/ACT-PRE-BNG-WEAR(BMI_Workshop).pdf. [Accessed: 2020-05-19].

[222] T. Richardson et al. Fusing Self-Reported and Sensor Data from Mixed-Reality Training. In *Interservice / Industry Training, Simulation, and Education Conference (I/ITSEC)*, 2014.

[223] TAS and VTT teams. EdcAR AO/1-8100/14/NL/MH Augmented Reality for AIT, AIV and Operations, December 2016.

[224] Tespack. Solar Backpack for Mars, July 2016. URL https://www.tespack.com/solar-backpack-mars/. [Accessed: 2019-10-25].

[225] U. H. Graneheim, and B. Lundman. Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 2004. ISSN 02606917. doi: 10.1016/j.nedt.2003.10.001.

[226] U. Neumann and A. Majoros. Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. *In Proceedings of the IEEE 1998 Virtual Reality Annual International Symposium*, pages 4–11, 1998.

[227] Unity. URL https://unity.com. [Accessed: 2019-09-27].

[228] V. Almonacid and L. Franck. Extending the coverage of the internet of things with low-cost nanosatellite networks. *Acta Astronautica*, 138:95–101, 2017. ISSN 0094-5765. doi: 10.1016/j.actaastro.2017.05.030. The Fifth International Conference on Tethers in Space.

[229] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 2006. ISSN 14780887. doi: 10.1191/1478088706qp063oa.

[230] V. Byrne, J. Mauldin and B. Munson. Treadmill 2 Augmented Reality (T2 AR) ISS Flight Demonstration. URL https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20190028716.pdf. [Accessed: 2019-11-14].

[231] V. Heun, J. Hobin, and P. Maes. Reality Editor: Programming Smarter Objects. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, pages 307–310, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2215-7. doi: 10.1145/2494091.2494185. URL http://doi.acm.org/10.1145/2494091.2494185.

[232] V. R. Preedy and R. R. Watson, editor. *5-Point Likert Scale*, page 4288. Springer New York, New York, NY, 2010. ISBN 978-0-387-78665-0. doi: 10.1007/978-0-387-78665-0_6363. URL https://doi.org/10.1007/978-0-387-78665-0{_}6363.

[233] VisionLib. URL https://visionlib.com. [Accessed: 2019-11-08].

[234] L. Vojtech, M. Nerada, J. Hrad, and R. Bortel. Outdoor localization technique using active RFID technology aimed for security and disaster management applications. In *Proceedings of the 2015 16th International Carpathian Control Conference (ICCC)*, pages 586–589, May 2015. doi: 10.1109/CarpathianCC.2015.7145148.

[235] VTT, April 2017. URL https://www.vttresearch.com/en. [Accessed: 2019-10-25].

[236] Vuforia. URL https://developer.vuforia.com. [Accessed: 2019-11-08].

[237] Vuforia. Getting started with advanced model targets 360. URL https://library.vuforia.com/content/vuforia-library/en/articles/Solution/Getting-started-with-advanced-model-targets-360.html. [Accessed: 2019-11-08].

[238] Vuforia. Object Recognition, . URL https://library.vuforia.com/content/vuforia-library/en/features/objects/object-reco.html. [Accessed: 2019-11-08].

[239] Vuforia. VuMarks, . URL https://library.vuforia.com/content/vuforia-library/en/features/objects/vumark.html. [Accessed: 2019-11-08].

[240] W. Piekarski and B. Thomas. ARQuake: The Outdoor Augmented Reality Gaming System. *Commun. ACM*, 45(1):36–38, January 2002. doi: 10.1145/502269.502291.

[241] W. Wei, A. T. Yang, and W. Shi. Security in internet of things: Opportunities and challenges. In *2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI)*, pages 512–518, October 2016. doi: 10.1109/IIKI.2016.35.

[242] Wikitude. Wikitude AR SDK for HoloLens. URL https://www.wikitude.com/wikitude-ar-sdk-for-hololens/. [Accessed: 2019-11-08].

[243] X. Wang, S. K. Ong and A. Y. C. Nee. A comprehensive survey of augmented reality assembly research. *Advances in Manufacturing*, 4(1):1, 2016.

[244] Y. Boger. Overview of positional tracking technologies for virtual reality. June 2014.

[245] Y.-H. Chen, B. Zhang, C. Tuna, Y. Li, E. A. Lee, and B. Hartmann. A context menu for the real world: Controlling physical appliances through head-worn infrared targeting. 2013.

[246] Y. Park, S. Yun, and K. Kim. When IoT Met Augmented Reality: Visualizing the Source of the Wireless Signal in AR View. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '19, pages 117–129, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-6661-8. doi: 10.1145/3307334.3326079.

[247] Z. Chen, J. Li, Y. Hua, R. Shen and A. Basu. Multimodal interaction in augmented reality. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 206–209, 2017.

[248] Z. Huang, W. Li, P. Hui, and C. Peylo. Cloudridar: A cloud-based architecture for mobile augmented reality. pages 29–34, 2014. doi: 10.1145/2609829.2609832.

[249] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1):22–32, February 2014. doi: 10.1109/JIOT.2014.2306328.

[250] Zigbee Alliance. Zigbee. URL https://zigbeealliance.org. [Accessed: 2020-04-17].