Human Factors in Instructional Augmented Reality for Intravehicular Spaceflight Activities and How Gravity Influences the Setup of Interfaces Operated by Direct Object Selection

DISSERTATION
to obtain the academic degree of
Doktor-Ingenieur (Dr.-Ing.)
of the Faculty of Computer Science and Electrical Engineering
at the University of Rostock

submitted by
Daniela Markov-Vetter
M.Sc. Dipl.-Inf (FH)
born on 02.02.1972 in Potsdam, Germany

Rostock, 14.10.2016

urn:nbn:de:gbv:28-diss2018-0189-1

Reviewers:

Prof. Dr. Oliver Staadt
University of Rostock
Faculty of Computer Science and Electrical Engineering

Prof. Dr. Dieter Schmalstieg
Technical University of Graz
Institute of Computer Graphics and Vision

Dr. Uwe Mittag
German Aerospace Center (DLR)
Institute of Aerospace Medicine

Date of Submission: 14.10.2016
Date of Defense: 12.07.2017

Acknowledgement

At this point, I would like to say thank you to all people who have motivated me and supported my research. First and foremost I would like to acknowledge Dr. Uwe Mittag, my supervisor at the German Aerospace Center (DLR), for his innovative thinking, which has enabled the release of this interesting research topic, but also for letting me space to find my way that brought me here. Secondly -but no less important - I would like to thank Dr. Oliver Staadt for hosting my research at his professorial chair at the University of Rostock. In particular for sharing and trust in my ideas and intentions, as well as for directing me when I fell to narrow my thoughts. Further acknowledgments go to Prof. Dr. Dieter Schmalstieg for reviewing this thesis and traveling to my defense of my doctors thesis in Rostock.

Also I would like to thank all staff of the Institute for Aerospace Medicine for the amazing and exciting time during my five years of residence at the DLR in Cologne. Thereby, special thanks goes to Stefan Moestl, Jens Hauslage, Bernd Johannes, Christoph Steger, Vanja Zander, Lars Uffmann, Wolfram Sies and Ruth Hemmersbach. Another big thank you goes to Peter Gauger for his great support during the research under parabolic flight conditions. Without his experiences and capabilities, several studies would not have been possible. Furthermore, I would like to thank Wolfgang Doering for his great support in general and the numerous talks about my work and ideas from a technical point of view, as well as for moderating and monitoring the usability study at the ESA European Astronaut Centre (EAC), which was also greatly assisted by Anja Simon. In this context I would also like to thank the team of the EAC, in particular Uwe Müllerschkowski and Norbert Illmer, who supported and enabled me to perform the study.

Gratitude also goes to the team of Novespace, who excellently organizes, coordinates and performs parabolic flight campaigns in Europe. Special thanks to Yannick Bailhe for his technical support and advices, as well as for preparing our experiments for the flights. Regarding those experiments, I would also like to thank the ESA and the DLR space agency for selecting and financially supporting this research.

A special thanks to Dr. Eckard Moll and Jörg Sellmann from the Julius-Kühn Institute for their strong support and advices for the numerous statistical evaluations.

Finally, I would like to thank my family and friends, who motivated my work and tolerated my one-dimensionality throughout those years. The deepest thanks goes to Tillmann Heier, the best friend, which I can only wish for. I would like to thank him for his endless effort spent on reading and correcting this thesis, but also for numerous interesting discussions about my research and, especially, on the English language and its meaning.

Apart from all others appreciative words of thanks, I would like to thank the most important people in my life. Thanks David, for your never ending patience and for being a constant source of motivation. But mostly for your love to the one I am.

Abstract

In human spaceflight, advanced user interfaces are becoming an interesting mean to facilitate human-machine interaction, enhancing and guaranteeing the sequences of space operations. During their average six month stay aboard the International Space Station, a majority of the astronauts' working time is scheduled for maintenance and mission specific tasks at research payloads. To ensure the reliability of such operations, each payload task is assigned to a standardized procedure that is provided on the Station Support Computer (SSC), a laptop computer laterally fixed to the working area, which constantly calls for switching between perceiving and executing an instruction. Besides failures of alternating attention, such guidance mode is characterized by effortful top-down processing, unnecessarily loading the working memory, which can be caused by the complexity and inappropriate presentation of the instructional information as well as by the visual search to localize the target area in question. As a result, the performance quality of payload operations can suffer from an increased workload level and thus bears the risk of astronaut errors.

The efforts hitherto made to ease astronauts' intravehicular operations have shown strong interests in using novel human-computer interaction in a variety of ways. One promising approach is the utilization of Augmented Reality (AR), which not only enables the integration of the instructional information into astronauts' field of view, it also supports fast bottom-up processing to direct attention towards the area of interest. Several projects have been launched to take advantage of AR, but were predominantly driven by technological innovations, neglecting the aspect of Human Factors, which, however, is required to provide an effective and efficient system, fulfilling the users' needs.

Complementary to such technological-driven approaches, the work presented in this thesis is directed towards a user-driven design for AR-assisted space payload operations, iteratively solving issues risen from the problem space. Compensating the lack of considering Human Factors implies different challenges, whereby the first one needed to be met, demands for controlled field testing in a summative way to verify that AR has the real potential to replace the common guidance interface and, in the best case, is capable to decrease the effort spent on task performance, increasing its quality. A further gap, not closed by competitive works, is the thorough investigation of specific conditions for task localization, which not only implies the visual search of the off-screen target area, but also the subsequent detection of the in-view target objects. While this research is targeted towards solving the trade-off between exocentric and egocentric cue stimuli and mode of perception, other research needs to be conducted to counterbalance the loss of input modalities usually provided by the SSC. Before discussing concrete approaches to design an AR input device, its placement is just as important as fundamental, which requests an appropriate spatial reference system. Assuming that such an interface is operated by direct touch, it can be placed in a world-, head- or hand-reference. Because the underlying visuomotor coordination is influenced by gravity, such research on AR device placement needs to consider being operable under modified gravity conditions, especially under weightlessness.

By closing these gaps by relevant research on human-computer interaction as well as by tackling challenging questions arisen from this research, this thesis presents its justified course of action and rephrases its findings to provide guidelines and principles, not only complementing existing design rules for usability, but also supporting interface designers to deploy a good conceptual model for AR assisted space payload operations from a Human Factors perspective.

Kurzfassung

Mit dem Fortschritt der bemannten Raumfahrt steigt auch das Interesse an dem Einsatz innovativer Benutzerschnittstellen um nicht nur die Mensch-Maschine Interaktion zu erleichtern oder effektiver zu gestalten, sondern auch um einen korrekten Arbeitsablauf sicherzustellen. Während ihrer durchschnittlich sechs monatigen Aufenthalte auf der Internationalen Raumstation, besteht die Tätigkeit der Astronauten hauptsächlich aus Wartungsarbeiten sowie der Durchführung von Forschungsaufgaben an sogenannten Nutzlastmodulen. Zur Durchführung dieser Arbeiten ist jede Aufgabe mit Hilfe einer standardisierten Prozedur beschrieben, welche eine genaue Durchführung garantieren soll. Die Darstellung solch einer Prozedur wird auf dem Station Support Computer (SSC) bereitgestellt, einem Laptop Computer der gewöhnlich seitlich am Rack befestigt ist. Diese Anzeigemodalität zwingt den Operator zu einem ständigen Wechsel zwischen der Informationsaufnahme und dem Arbeiten am Forschungsmodul, sowie zu einer aufwändigen visuellen Neuausrichtung vor jedem Ausführungsvorgang. Neben dem Problem der ständigen Allokation alternierender Aufmerksamkeit, führt diese Form der Anleitung zu aufwändigen konzeptgeleiteten (top-down) Verarbeitungsprozessen, welche durch die Komplexität und der unangemessenen Darstellung der Anleitungsinformationen, aber auch durch die visuelle Suche zur Objektlokalisierung verursacht wird und somit das Arbeitsgedächtnis in unnötigem Maße belastet. In Folge dieser Mehrbelastung steigt das Fehlerrisiko während der Durchführung der Arbeitsaufgaben und birgt somit die Gefahr menschlichen Versagens.

In den vergangenen Jahren wurden wiederholt Anstrengungen unternommen, die Innenbordarbeiten an den Forschungsmodulen zu erleichtern, welches ein vielfältiges Interesse an neuartiger Mensch-Computer Interaktion mit sich brachte. Als ein vielversprechender Ansatz zeigt sich hier die Anwendung der Erweiterten Realität (engl. Augmented Reality, AR), welche nicht nur eine Integration der relevanter Informationen im Blickfeld des Astronauten ermöglicht, sondern auch zu schnellen datengeleiteten (bottom-up) Verarbeitungsprozessen führt, wobei die visuelle Aufmerksamkeit präattentiv gesteuert werden kann. Bis heute wurden verschiedene Projekte initiiert um diese Vorteile von AR für die Arbeiten an Forschungsmodulen zu nutzen. Diese waren jedoch vorwiegend technologiegetrieben und vernachlässigten den Aspekt der menschlicher Einflussgröße, welcher jedoch eine Voraussetzung für wirkungsvolle und leistungsfähige Systeme ist.

Komplementär zu diesen technologieorientierten Ansätzen, konzentriert sich die vorliegende Arbeit auf einen nutzerorientierten Ansatz, welcher zum Ziel hat die vorherrschenden Probleme schrittweise in einem iterativen Designprozess zu lösen. Der Ausgleich des Mangels an Berücksichtigung des Faktors Mensch erfordert zunächst die Durchführung eines kontrollierten Feldtests, um zu prüfen ob AR überhaupt qualifiziert ist das vorhandene Informations- und Assistenzsystem zu ersetzen und im besten Fall sogar den Aufwand bei der Aufgabenausführung senken und so deren Qualität verbessern kann. Ein weiterer Iterationsschritt erfordert eine gründliche Untersuchung spezifischer Methoden zur visuellen Objektlokalisierung an den Forschungsmodulen. Diese besteht nicht nur aus der initialen Suche außerhalb der visuellen Anzeige, sondern auch in der nachfolgenden Zielortung im Sichtbereich. Während sich diese Forschung auf das Lösen des Zielkonfliktes zwischen exozentrischen und egozentrischen Hinweisreizen beziehungsweise Wahrnehmungsweisen konzentriert, haben andere Untersuchungen das Ziel den Verlust der herkömmlichen Eingabegeräte,

wie zum Beispiel der Tastatur, auszugleichen. Bevor aber konkrete Ansätze zum Gestalten einer kompensierenden AR Eingabeschnittstelle diskutiert werden, ist die Frage nach der räumlichen Position grundlegender und wichtiger. Vorausgesetzt, dass solch ein Eingabegerät mittels direkter Berührung bedient wird, können Welt-, kopf- oder handbezogende Referenzsysteme zum Positionieren genutzt werden. Da die dabei zugrundeliegende visiomotorische Koordination wesentlich von der Schwerkraft beeinflusst wird, sollten diesbezügliche Untersuchungen die Auswirkungen verändernder Schwerkraftbedingungen, insbesondere der Schwerelosigkeit, berücksichtigen.

Bei Beantwortung dieser Fragestellung durch relevante Forschung im Bereich Mensch-Computer Interaktion, sowie bei der Verfolgung weiterführender Forschungsfragen, präsentiert die vorliegende Arbeit neben deren Bedeutsamkeit und methodische Vorgehensweise, auch die Formulierung ihrer Erkenntnisse in Form von Richtlinien und Grundsätzen. Diese sollen nicht nur bestehende Normen der Usability im Allgemeinen ergänzen, sondern konkret Interfacedesigner dabei unterstützen, ein benutzerzentriertes Konzept zur Gestaltung AR-basierter Assistenzsysteme zur Unterstützung von Innenbordarbeiten an ISS Forschungsmodulen zu entwerfen.

List of Figures

List of Tables

List of Abbreviations

2D	Two dimensional
3D	Three dimensional
0*g*	Microgravity, Weightlessness, Zero-G
1*g*	Normogravity, Earth condition
ANS	Autonomous Nervous System
API	Application Programming Interface
AR	Augmented Reality
ATV	Automated Transfer Vehicle
AV	Augmented Virtuality
Biolab	Biological Experiment Laboratory
BIOTESC	Biotechnology Space Support Center
BMI	Body Mass Index
B-USOC	Belgian User Support and Operation Center
CATIA	Computer Aided Three-Dimensional Interactive Application
CCS	Crew Commanding Station
CI	Confidence Interval
CNES	Centre National d'études Spatiales
CNS	Central Nervous System
Col-CC	Columbus Control Center
COTS	Commercial Off-The-Shelf
CSA	Canadian Space Agency
CTA	Cognitive Task Analysis
DAC	Danish Aerospace Company
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
DOF, DF	Degrees of freedom
EAC	European Astronaut Center
EADS	European Aeronautic Defense and Space Company
ECG	Electrocardiogram
EDR	European Drawer Rack
EMU	Extravehicular Mobility Unit
EPM	European Physiology Modules
E-USOC	Spanish User Support and Operation Center
EVA	Extravehicular Activity
ESA	European Space Agency
ESTEC	European Space Research and Technology Centre
FPS	Frames per second
FSC	Facility Support Center
FOV	Field of View
FSL	Fluid Science Laboratory
FRC	Facility Responsible Center
FSC	Facility Support Center

GPS	Global Positioning System
HCI	Human-Computer-Interaction
HF	High Frequency (in the context of HRV)
HFE	Human Factor Engineering
HMD	Head-Mounted Display
HMI	Human-Machine Interface
HMM	Hidden Markov Model
HRV	Heart Rate Variability
HTK	Hidden Markov Model Toolkit
HUD	Head-Up Display
iPV	international Procedure Viewer
ISMAR	International Symposium on Mixed and Augmented Reality
ISPR	International Standard Payload Rack
ISS	International Space Station
IVA	Intravehicular Activity
JAXA	Japan Aerospace Exploration Agency
LAHC	Long-Arm Human Centrifuge
LEO	Low Earth Orbit
LF	Low Frequency (in the context of HRV)
LF/HF	Ratio between LF and HF (in the context of HRV)
LOD	Level of Detail
M	Arithmetic Mean
MARS	Microgravity Advanced Research and Support Centre
MIT	Massachusetts Institute of Technology
MR	Mixed Reality
MSG	Microgravity GloveBox
MSL	Material Science Laboratory
MUSC	Microgravity User Support Center
N	Number of Observations
NASA	National Aeronautics and Space Administration
NASA TLX	NASA Task Load Index
NASA RTLX	NASA Raw Task Load Index
NEEMO	NASA Extreme Environment Mission Operations
NTSC	National Television Systems Committee
N-USOC	Norwegian User Support and Operation Center
ODF	Operation Data File
OST	Optical see-through
PAT	Procedure Authoring Tool
PDA	Personal Digital Assistant
PF	Parabolic Flight
PFC	Parabolic Flight Campaign
POC	Proof Of Concept
PODF	Payload Operation Data File
R&D	Research and Development
Roscosmos	Russian Federal Space Agency
SSC	Station Support Computer
SD, Stddev	Standard Deviation

SE, Stderr	Standard Error
SI	Stress or Strain Index
SIFT	Scale-Invariant Feature Transform
SLAM	Simultaneous Localization And Mapping
SODF	Station Operation Data File
SURF	Speeded-Up Robust Features
UI	User Interface
USOC	User Support and Operation Center
VR	Virtual Reality
VGA	Video Graphics Array
VST	Video see-through
WIM	World In Miniature
XML	Extensible Markup Language

Table of Contents

Chapter 1

Introduction

"A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools."

(Douglas Adams, *Mostly Harmless*, 1992)

The introduction chapter presents the motivation for the use of Augmented Reality to assist astronauts during intravehicular space operations. Describing the problem identifies missing issues and promotes the research objective by formulating the thesis statements. An overview of the thesis activities constitutes the interdisciplinary scope and establishes the basis to state the major contribution of this work. Final outlining of the workflow of this thesis illustrates the organization of the chapters that are briefly described.

1.1 Research Motivation

On an orbital space station, like the International Space Station, astronauts have to deal with a huge amount of work under challenging environmental conditions. In the absence of gravity they need to perform complex tasks that require deep concentration and high-precision movements. The adaptation to weightlessness and human-machine interaction are challenges that strongly impact the level of astronaut's performance and is of extreme importance to ensure the success of a space mission. As shown in Figure 1.1, the tasks that astronauts have to accomplish are generally classified into two categories as extravehicular activities (EVA) and intravehicular activities (IVA). Beside of extraterritorial explorations, like the lunar surface exploration by the Apollo program from 1969 to 1972, EVAs are long-duration activities conducted outside of a spacecraft and are always associated with a spacewalk, which requires wearing of an Extravehicular Mobility Unit (EMU), a spacesuit that provides astronauts with life support functions. For example, from 1993 to 2009 several EVA missions were required to conduct repair and maintenance tasks of the Hubble telescope. However, regular EVA tasks are currently conducted to maintain and extend the International Space Station (ISS). In contrast, IVAs are task activities done inside a space station and are including a variety of different functions that have already been conducted from the beginning of manned space stations aboard the Russian Salyut 1 in 1971 until now aboard the ISS.

Such IVA tasks are ranging from operating to maintaining the space station to rendezvous maneuvers, like the docking of the Automated Transfer Vehicle (ATV). A majority of intravehicular tasks for an astronaut consists of the performance of space science experiments. Such experiments are conducted for scientists on ground, whose research is engaged with better understanding of the effect of weightlessness on physical and life science phenomena. While for human life science experiments, astronauts also act as subjects (see Fig. 1.2, left), biological, physical and material science experiments are conducted by means of research payloads that are regularly housed in a standard rack, also known as International Standard Payload Rack (ISPR), that is an uniformly sized rack to accommodate different types and configurations of science equipment. The operation of the ISPRs requires regular maintenance and repair tasks as needed. A single experiment or maintenance task at an ISPR may last between thirty minutes and two hours on average and may amount to significant

workload. To ensure the consistency and reliability of such ISPR tasks the International Partners of the ISS program have designed and developed a standardized procedure format (Pruzin et al., 2003). The Operation Data File (ODF) format itemizes the necessary tools, the prescribed time period and a sequence of text-based instructions complemented by visual resources. There is a specific ODF procedure for every ISS operation that will be conducted step-by-step in cooperation between the astronaut and the ground team of the responsible User Support Operation Center (USOC) to ensure that a task will be performed as precisely as possible. Because astronauts have to carry out numerous ODF procedures, strictly-formatted procedures guidance is crucial for the success of an operational task and is therefore intensively practiced as part of the astronaut's ground training.

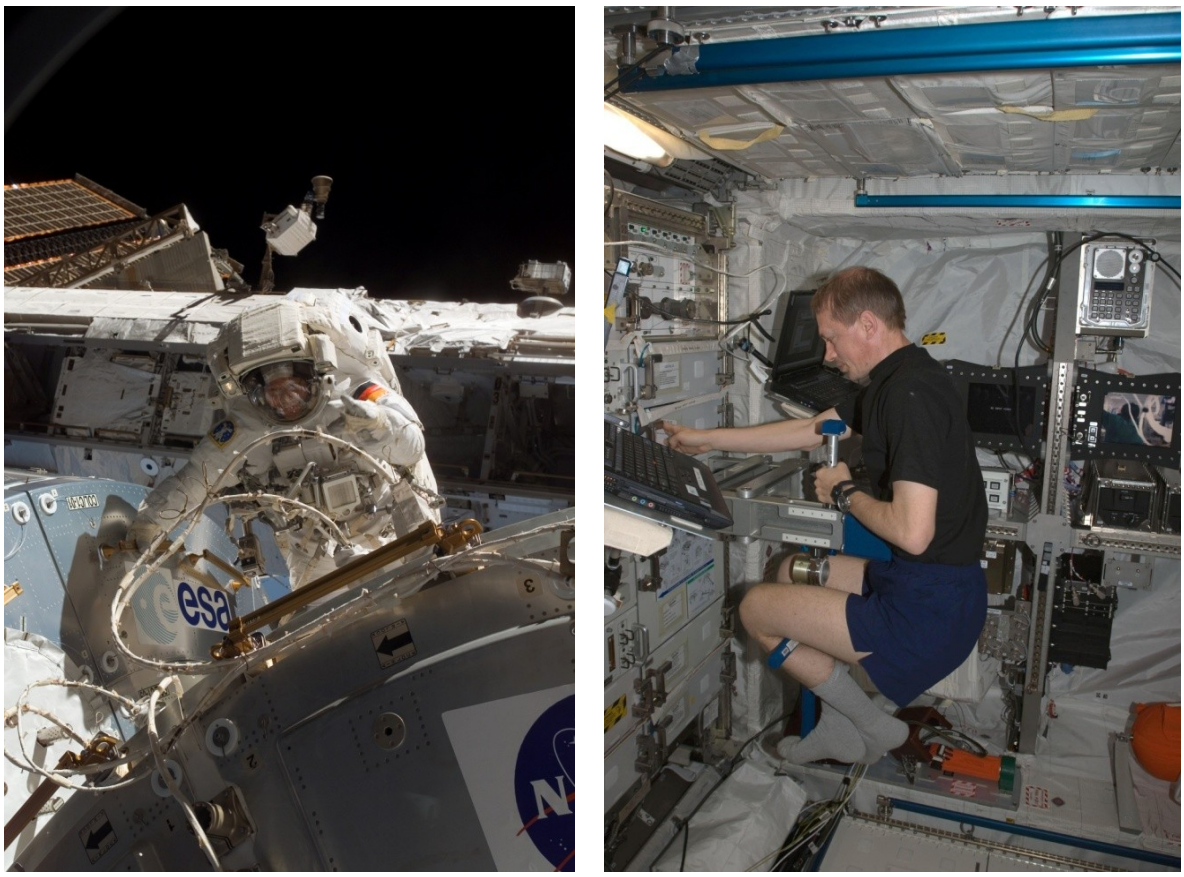


Figure 1.1: Extra- and intravehicular activities of astronauts' work: *(left)* ESA astronaut Hans Schlegel conducting extravehicular activities (EVA) working at Quest Airlock of the ISS [Photo © NASA, 2008], and *(right)* ESA astronaut Frank de Winne conducting intravehicular activities (IVA) working at the Human Research Facility (HRF) Rack in the European Columbus laboratory of the ISS. [Photo © NASA, 2009]

1.2 Problem Statement

Successful space missions are vitally depending on human performance to complete tasks accurately in a reasonable time. While astronauts are carrying out maintenance and mission specific tasks at ISPRs using PODFs (Payload ODFs), they obtain the necessary information and instructions from a laptop computer, called the Station Support Computer[1] (SSC), which is laterally installed at the rack

[1] The "Station Support Computer" (SSC), also referred to as "Crew Commanding Station" (CCS), is a laptop computer to provide crew operations and share ISS resources by the"Operations LAN" (OPS LAN), while the"Crew Support LAN" (CSL) provides personal Web access. Most SSCs are Lenovo ThinkPads, complying with the T61p model.

construction (see Fig. 1.2, right). During such guided operations, astronauts are forced to change their view constantly between the SSC and the region of operation at the ISPR, which can lead to loss of focus, loss of concentration and attention. Besides this risk of failure of attention, it is conceivable that the current guidance mode additionally implies increased costs of human information processing, which is conditioned by a high perceptual and cognitive load during the reception of the instruction information and the subsequent task localization. Such shortcomings may reduce the performance quality, increase the operator workload and, thus, can lead to astronaut errors. This turned out to be even more relevant during long-duration procedures and emergency procedures that were not trained in advance. These issues indicate the need of a proper contemporary human-machine interface (HMI) to assist the astronaut during scheduled and unscheduled ISPR payload operations in space.

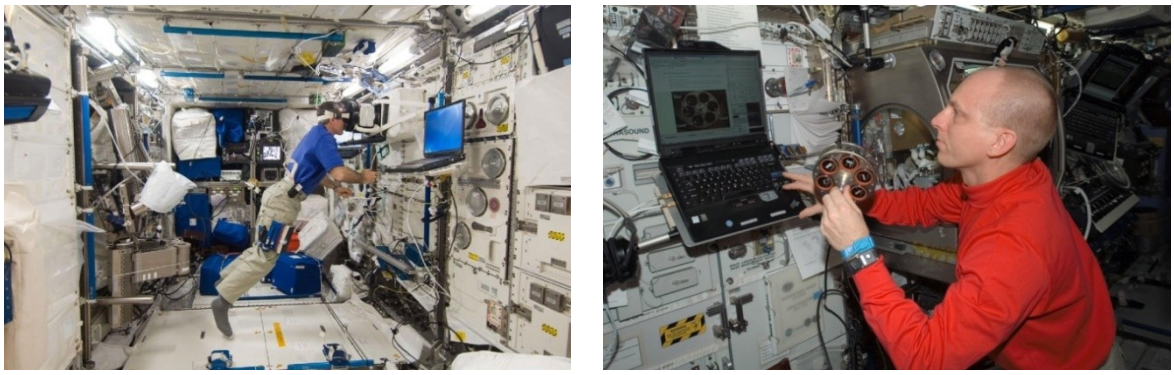


Figure 1.2: Experiment Performance at ISS research payloads: *(left)* NASA astronaut Dan Burbank conducts the Neurospat[2] experiment in the ESA Columbus [Photo © NASA, 2012], and *(right)* NASA astronaut Clay Anderson works at the Microgravity Science Glovebox in the U.S. Destiny module [Photo © NASA, 2007].

Over the last decades the *National Aeronautics and Space Administration* (NASA) and the *European Space Agency* (ESA) have shown strong interest to ease astronaut's work by applying advanced human-computer interfaces, including such as Augmented Reality (AR). Although it has been established that the procedure standard is limited by its supply of visual resources and needs to be advanced, the efforts made to improve astronauts' operational activities by AR not only suffered from a fundamental lack of understanding the problem space, the existing approaches seem to be more technology-driven, not considering directives given by Human Factors engineering. Thus, in order to optimize astronauts' payload activities in an adequate way, it is necessary to identify the drawbacks that need to be encountered as well as to induce an iteratively design process, satisfying the system and user needs. Such course of action would also imply the consideration of environmental factors. Space as probably the most extreme working environment is not conducive to human performance and thus the perpetuation of the usability under such extreme environmental conditions constitutes a critical requirement for the design of advanced payload interfaces. Because the absence of gravity has, amongst others, consequences on the spatial orientation and visuomotor coordination, relevant restrictions may reduce the benefit of advanced payload interfaces, especially those which imply a direct interaction modality. Although most of the existing AR approaches were field tested onboard the ISS, they neither were focused on interactive features nor were covered by controlled studies, investigating the effect of weightlessness on astronaut's physical and cognitive well-being. Consequently, the design process additionally needs to take into account changes of human sensorimotor system caused by modified conditions of the gravity force. This enables, for example, to

[2] The Neurospat experiment examined changes in spatial orientation and perception affected by spaceflight conditions. (ESA Erasmus Archive, URL: http://eea.spaceflight.esa.int/portal/exp/?id=9143, last visit: 28.09.2016)

detect and compensate visuomotor deficits, which was not considered so far. Thus, the research presented in this thesis is based on the concept of appropriate integration of Human Factors while studying the effect of Augmented Reality on payload operations.

This research was initiated by the *Institute of Aerospace Medicine*[3] of the German Aerospace Center (DLR) in 2009. The department of Biomedical Research is, amongst others, responsible for the preparation of scientific experiments in the Biolab ISPR that enables biological research with a wide variety of laboratory functionalities to perform under microgravity. In cooperation with the *Microgravity User Support Center*[4] (MUSC), a German USOC (User Support and Operation Center) responsible for the operation of the Biolab ISPR, the DLR has requested to evaluate the benefit of Augmented Reality for space payload operations, using the Biolab ISPR as case study.

1.3 Advanced User Interfacing using Augmented Reality

In order that humans benefit from computational devices the interaction to the computer system shall be proper designed in an effective, efficient, easy, and enjoyable way (Dix et al., 2003). While until the late 1970s only experts in information technology were interacting with computer systems, today, 45 years later, everyone will be able to use computers, whether professionally at work or privately at home. With the arising of the desk metaphor introduced by Xerox and popularized by Apple since 1984, the foundation for *Human-Computer Interaction* (HCI) as a research field in computer science was laid. As stated by Hewett and colleagues of the Curriculum Development Group of the ACM SIGCHI (Special Interest Group on Computer-Human Interaction) in 1992:

> "Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them." (Hewett et al. 1992)

By considering ergonomic and physiological aspects, HCI has become a tool in developing and investigating user-centred computer interfaces consolidating Human Factors engineering and cognitive science. Beyond desktop computing, soon the technological development revealed novel device contexts ranging from laptop and handheld devices to *Ubiquitous Computing* as pervasive presence of computing in human surroundings, first introduced by Mark Weiser (1993). Whether in cars, accommodation, clothing or articles of daily use in general, ubiquitous computing is intended to be everywhere, regardless of the device, location or format. While ubiquitous computing enriches the physical world with computers, other approaches are mixing the physical world with the digital environment of computers. Thus, *Mixed Reality* is bridging the gap between the real and virtual world that has established new conditions for HCI towards a combined multisensory level of perception. The term "Mixed Reality" (MR) was defined by Paul Milgram and Fumio Kishino in 1994 (Milgram & Kishino, 1994) in their description of the *Reality-Virtuality Continuum* (see Fig. 1.3) that describes a gradation between the real and the virtual world. Thereby Mixed Reality can range from applications in the real world with synthetic objects, referred to as *Augmented Reality* (AR), to a synthetic world with real objects, known as *Augmented Virtuality* (AV). While AV applications are integrating something from the real world into a virtual environment in real-time and are used for

[3] Institute of Aerospace Medicine (DLR-ME). URL: http://www.dlr.de/me/, last visit: 28.09.2016
[4] Microgravity User Support Center (MUSC). URL: http://www.dlr.de/musc/, last visit: 28.09.2016

instance in concepts using a virtual studio, AR applications are using a reverse concept by dynamically setting up the physical reality superimposed with virtual objects.

Figure 1.3: The Reality-Virtuality Continuum (reproduced from Milgram & Kishino, 1994).

In 1968, Ivan Sutherland already built the foundation for AR by his development of the first see-through display (see Fig. 1.3, left) coupled with a rudimentary mechanical tracking system (Sutherland, 1968). But only in 1992, Bajura and colleagues conducted the first attempts of practical AR use by overlaying 3D visualizations of ultrasound data onto patients to support surgical procedures (Bajura et al., 1992). In the same year, Caudell and Mizell (1992) introduced the term "Augmented Reality" and discussed its application for aircraft manufacturing. With the technological development of powerful devices for computing and tracking in the 1990s and beside stationary interface setups, first mobile AR systems used for outdoor applications were appearing. In 1993, Loomis and colleagues constructed an outdoor personal guidance system for visual impaired (Loomis et al., 1993) using a laptop computer, a head worn electronic compass and a GPS (Global Positioning System) receiver. Nevertheless the first visual mobile AR system was presented by Feiner and colleagues in 1997. Their Touring Machine was developed as campus information system and used differential GPS for position sensing and mainly consisted of a see-through head-mounted display (HMD) and a laptop computer worn on the person's back (see Fig. 1.4, middle and right). In 1997, Ronald Azuma reviewed the by then status of AR systems and defined AR more closely as requiring the fulfillment of the following three criteria:

1. **Combine the real and virtual world:** AR merges the physical world with synthetic information, where a user still perceives the usual surrounding.
2. **Interactive in real-time:** AR systems are interactive and respond to user input in real-time.
3. **3D registered in the real world:** AR spatially integrates virtual objects into a real surrounding by tracking the 3D pose.

Thus, the static presentation of the flight altitude in a pilot's head-up display as an example, fulfills the requirements according to Milgram's and Kishino's definition of AR, but lacks the third criterion from Azuma. With the compliances to Azuma, Augmented Reality offers a direct interface to the physical reality enriched with virtual spatially registered information in real-time. This has not only the potential of increasing the perception of our daily surroundings, but also during the performance of complex working tasks. Until now, AR is investigated and used in a variety of applications areas, such as, but not limited to medicine, architecture, education, training, tourism, construction and entertainment. The design and development of AR interfaces ranges across a wide field of researches. Many projects, like ARVIKA (Friedrich, 2002), Studierstube [5] (Schmalstieg et al., 2002) or AVILUSplus [6], are focused on improving of the key technologies including pose tracking, registration, ergonomic displays, visualization and rendering. Alternatively to highly complex AR

[5] Studiestube. URL: http://www.icg.tugraz.at/project/studierstube, last visit: 28.09.2016
[6] AVILUSplus. URL: http://www.avilusplus.de, last visit: 28.09.2016

setups, some researchers are focused on handheld AR (Wagner & Schmalstieg, 2003) enabling anyone to experience AR on mobile phones (Gervautz & Schmalstieg, 2012). Overall AR occupies an important position in the field of HCI research with the potential of being one of the most influential interface technologies for operations with digital content.

Figure 1.4: Augmented Reality in application: *(left)* the first see-through display (from Sutherland, 1968), *(middle and right)* the Touring Machine as first mobile AR system (from Feiner et al., 1997).

Beside of innovative technologies intended to make AR ready-to-use in smart technical environments, the human performance decides finally on the usability. While AR as interface technology claims being in the physical reality, resulted AR systems should solve practical real-world problems and effectively ensure a consistent level of task integration with a proven eligibility[7]. To successfully apply an AR system, it needs to be evaluated related to its task domain. This also includes that adequate usability testing should leave bulky laboratory settings and take place within the corresponding environment to which the AR system is intended to operate, as well as being conducted from a representative range of end-users[8]. Thus, AR interfaces should be adapted to human needs and be optimized for the environmental conditions to which the AR interface will be applied (Kruijff et al., 2010; Huang et al., 2013; Li & Duh, 2013). Besides the lack of adequate controlled field tests, the current research in Human Factors on handling AR interfaces is presuming terrestrial conditions and therefore normogravity. Its adoption to operations under weightlessness is a unique challenge that strongly affects the level of usability and thus it is of particular importance to investigate whether the human performance is adequately preserved under altered gravity conditions.

1.4 Research Objective and Thesis Statements

In order to support payload operations in an efficiently way, this research discusses and investigates the capability of mobile Augmented Reality in consideration of a user-driven design process. Because of the fact that altered gravity has a tremendous effect on the human physiology and psychology, and therefore also on users' performance of handling interactive AR interfaces, the research needs to be timely addressed on environmental restrictions given by the condition of weightlessness. Because payload operations also demand to enter command-based instructions, the loss of traditional input

[7] "The long-term goal for many AR researchers is to build usable applications that people prefer over conventional methods and with which they perform tasks better. This requires identifying applications that can benefit from AR and building good interfaces to support those tasks." (Livingston, 2005)

[8] "Getting useful results with cognitive evaluations usually requires having subjects who are reasonable precise representatives of the intended end-user population." (Huang et al., 2013)

devices needs to be compensated. As a result, this research aims to deepen the understanding of problems faced by payload operations in reference to Human Factors, as well as how AR can contribute to this problem space. Finally, the research findings should provide guidelines to support and optimize the design of an adequate AR assistance from a Human Factors perspective.

Hence, the research objective can be simply stated by the following main goal:

> To improve the work of astronauts during intravehicular payload operations by means of advanced user interfaces using Augmented Reality.

To reach and verify the main goal, the following hypothesis statements have been drafted:

H1: Augmented Reality is capable to fulfill the requirements of payload operations.

H2: Augmented Reality enhances the performance of payload operations.

H3: Augmented Reality decreases the workload during payload operations.

H4: Augmented Reality is operable under altered gravity conditions.

H5: Handheld Augmented Reality input devices enhance the quality of visuomotor coordination under altered gravity conditions.

H6: Tactile sensation enhances the quality of visuomotor coordination while operating Augmented Reality input devices under altered gravity conditions.

1.5 Thesis Activities

The research needed to prove the validity of the hypothesis mentioned above, have implied various underlying activities that I conducted in the scope of this thesis at the Institute of Aerospace Medicine at the German Aerospace Center (DLR) in Cologne. Listed below is an overview of the thesis activities, whose results are incorporated into the content of this thesis:

(1) Studying astronauts' workplace, particularly in the context of intravehicular payload operations.

(2) Studying the effect of microgravity on human sensorimotor system.

(3) Studying theories and concepts of Human Factors and usability engineering.

(4) Selecting methods for usability testing in an AR environment.

(5) Compiling a survey of selected works done in the field of instructional AR.

(6) Compiling a survey of research on AR task localization.

(7) Compiling a survey of research on the placement of interactive AR interfaces.

(8) Compiling a survey of advanced human-machine interfaces for space-related work processes, in particular regarding AR interfaces.

(9) Identifying challenges and their roots, astronauts are facing during payload operations.

(10) Identifying gaps, not closed by competing works on AR supported space operations.

(11) Implementing a high-fidelity prototype of Mobile AR for Space Operations (MARSOP).

(12) Designing and coordinating a proof-of-concept study to test the MARSOP system at the European Astronaut Center (EAC).

(13) Designing and coordinating a ground-based user study investigating the spatial orientation of target cueing during payload operations.

(14) Acquiring, designing and coordinating two experiments under parabolic flight conditions, investigating the placement of interactive AR interface.

(15) Developing and ground-based usability testing of two approaches to enhance direct AR object selection by means of increased sensorimotoric load.

(16) Implementing high-fidelity prototypes used for usability testing related to points (12)-(15).

In addition, one master project and two master theses related to this thesis topic were announced and supervised in cooperation with the Department of Computer Science of the University of Applied Sciences Bonn-Rhein-Sieg:

Shazad, M. (2011). Detection and tracking of pointing hand gestures for AR applications.

Millberg, J. (2011). Generierung von Keyframes als Basis für markerloses Tracking in 3D Augmented Reality Anwendungen. (Generation of keyframes for markerless tracking in Augmented Reality applications)

Millberg, J. (2012). Markerloses, modellbasiertes Echtzeit-Tracking für AR-Applikationen. (Markerless, model-based real-time tracking for AR applications)

1.6 Contribution

This thesis contributes to the area of human-computer interaction (HCI) in computer science and, as a consequence, also to the superordinate area of Human Factors Engineering (HFE) in ergonomics. Regarding the specific aspect of altered gravity conditions, this thesis also contributes to hyper- and microgravity research in space life science and space technology. The overall contribution of this thesis is the improvement of the working process during payload operations that are used for ISPR tasks by the introduction of AR supported information delivery, also taking its consequences into account, such as the loss of traditional input devices. This includes the following contributions that were originated in the scope of this thesis:

(1) Field Testing of Mobile Augmented Reality for Space Operations.

(2) Evaluation of the effort spent on AR task localization by multidimensional workload assessment.

(3) Verification of the relevance of human individual predisposition in spatial processing for AR task localization.

(4) Determination of spatial placement conditions on AR input devices under altered gravity.

(5) Introduction of gravity-adapted interface resizing for direct AR selection to enhance visuomotor coordination.

(6) Enhancement of visuomotor performance of direct selection on nonhaptic AR interfaces.

(7) Providing guidelines and principles to support the design of AR-assisted payload operations.

The previously listed contributions are briefly described below:

(1) Field Testing of Mobile Augmented Reality for Space Operations (MARSOP Study)

Contrary to common laboratory usability tests, a field test enables to systematically observe the behavior of the intended user group in the target environment in question. To verify whether AR can contribute to payload operations and replace the common guidance interface, a controlled field study is required, comparing both conditions in the course of a summative evaluation. Although previous work on AR-assisted space payload operations was tested by astronauts onboard the ISS (Arguello, 2009; Scheid et al., 2010), there was no consideration of usability engineering in an adequate way, it was rather focused on technological apsects and operational tests. The opposite is covered by most AR research beyond the spaceflight sector. Although AR raises the claim for the physical reality, most usability tests take place under controlled conditions in laboratories and neglect the external validity with respect to the transferability in practice. Only few usability tests considered adequate field conditions to investigate the benefit of AR to instructional real-world tasks (e.g., De Crescenzio et al., 2011; Henderson & Feiner, 2009, 2011a, Webel et al., 2011b). Whether by insufficient usability test or missing field conditions, to prevent such inadequacies, it is necessary to strike the right balance between the levels of control and realism. Hence, to demonstrate its feasibility and get adequate feedback, the concept of AR was proven by field testing at the ESA European Astronaut Center, which not only provides a real payload environment, it also enabled to recruit participants, who were representative of the intended user population. However, this should be preceded, on the one hand, by an analytic evaluation of the problem space, which ideally comprises the analysis of representative tasks, the end-user group and the environmental conditions. This not only has enabled to identify the drawbacks of the current guidance interface (see section 5.1.1), but also to clarify why and how intravehicular payload tasks could benefit from AR (see section 5.1.2). On the other hand, the prior development of a full-fledged AR system is required. This was realized by implementing a high-fidelity prototype (see section 6.1), which enables the authoring and viewing of Mobile AR for Space Operations (MARSOP) in an appropriate manner. Nevertheless, the contribution of this research is on the one hand to evaluate the benefit of AR to payload tasks in a summative way, which was not considered so far, and on the other hand to provide evidence that AR supports a higher level of performance quality than the traditional guidance method. This research contributes to HCI by applying AR to space technology. While the underlying challenge is discussed in section 5.2.1, details on the controlled field test and its results are presented in Chapter 6.2.

(2) Evaluation of the Effort Spent on AR Task Localization by Multidimensional Workload Assessment (ARGuide Study)

One fundamental issue while developing user-centered AR interfaces for ISPR operations is the spatial alignment of the target cues, supporting the localization of the task area and object in question. Thereby the visuospatial perception delivers the necessary information to reach objects in the visual field and to facilitate the human ability to shift the gaze to specific points of interests. This can be influenced by the reference frame, in which the target cues are coded. Consequently, a target cue can be coded in an exocentric or egocentric way, each providing its own coordinate system and process to update the spatial information. Taking this into account, the common method to guide payload operations by the SSC implies exocentric target cueing, forcing the astronaut to an effortful visual search at the payload. In contrast, the use of an AR

display, such as a HMD, offers an egocentric perspective, but can vary in the way the target cues are visualized. While an exocentric display automatically implies the same orientation of the visualization, an egocentric display can provide either an egocentric visualization, optimally 3D-registered, as done by AR, or an exocentric one, visualizing the target cues in a static manner on top of the display, not registered to the visual information attended. As shown by AR research, it is not clear which kind of spatial updating supports the best task localization (e.g., Tönnis et al., 2005; Tönnis & Klinker, 2006; Robertson et al., 2008; Medenica et al., 2011; Khuong et al., 2014). Because it can be reasonably assumed that the process of task localization depends on the task requirements in question, it is needed to be checked for payload operations. Thus, a trade-off study was conducted to find out the most suitable reference frame for payload operations. By doing so, an important criterion was identified by the lack of assessing the effort spent on task localization. Not only in research on task localization, evaluating the efficiency by the resulting workload level in an adequate way is usually neglected in AR research. Thus, the trade-off study claimed to assess the workload in its full dimension, analyzing primary and secondary task indicators, subjective ratings as well as cardiovascular feedback. Therewith, this research makes not only a contribution to HCI by applying AR to space technology, it also contributes to usability testing in AR. To date, there was no work, showing the relevance and importance of assessing the workload level in an adequate way. While the underlying challenge of the research on task localization is further underpinned in section 5.2.2, details on the trade-off study and its results are presented in Chapter 7.

(3) Verification of the Relevance of Human Individual Predisposition in Spatial Processing for AR Task Localization (ARGuide Study)

Spatial tasks, like solving of navigation problems to localize a target object, require human visuospatial ability of orientation, which is associated with the use of a specific reference frame. To localize objects in their environment, humans are in general capable to use both reference frames (egocentric and exocentric), but are often influenced by a preference. Research in biopsychology even showed that human spatial behavior is also affected by an individual predisposition (Gramann, 2013) and has proven the existence of discrete strategies in spatial navigation, which means that humans use distinct reference frames in the overall population for spatial updating (Goeke et al., 2013). This suggests that such distinctions in spatial strategy should be considered in studies investigating spatial orientation behavior, like the evaluation of AR systems for navigation or in general for spatial cuing tasks. Thus, in the course of the research on AR task localization, the relevance of individual human predisposition in spatial processing was investigated additionally. This aspect was considered for the first time in AR usability research and, therefore, this research contributes to AR in HCI. As the research on task localization, this research is presented in Chapter 7.

(4) Determination of Spatial Placement Conditions on AR Input Devices Under Altered Gravity (3DPick Study)

To release PODF related actions, currently an astronaut uses the classical keyboard and touchpad of the SSC laptop as input devices. Besides viewing the content, a full-fledged AR system, to assist payload operations, needs to provide interaction modalities for control or symbolic-input tasks, so that the astronaut is able to communicate with the PODF procedure in an adequate way, complying with the requirements demanded by AR. Before designing such interfaces to optimize, for example, their control elements, as important as fundamental is the problem given by the spatial placement. Presuming direct object selection, an efficient

handling of such an interface is featured by its underlying visuomotor coordination. As shown by related AR research, this is influenced by the reference the interface is placed in (e.g., Feiner et al., 1993; Lindeman et al., 1999; Bowman & Wingrave, 2001; Harrison et al., 2011). Correspondingly, such interfaces can be placed by a world, body or head reference, whereby a body reference here will be restricted to the hand. While such research on placement of AR input devices has been only conducted under normogravity, research on device placement for astronauts' payload operations involves taking into account their environmental conditions, implying the absence of the gravity force. Thus, to evaluate the hand-eye coordination, resulted from a world, hand and head reference, it needs to be considered that each of them implies its own sensorimotor demand, which is generally affected by altered gravity conditions, especially those which are causing deficits in visuomotor coordination (e.g., Berger et al., 1993; Mechtcheriakov et al, 2002; Bock et al., 2003; Fowler et al., 2008). Although microgravity research extensively investigated goal-directed pointing movements, it does not provide relevant findings to answer the question which placement reference is preserving the highest quality of visuomotor coordination in AR direct object selection. Thus, it was required to study the trade-off between the reference types under altered gravity conditions. To answer the question on device placement, two experiments were conducted under parabolic flight conditions, which provides short-term hyper- and microgravity. Up until now, neither microgravity research nor AR research was undertaken to investigate the effect of altered gravity on visually controlled hand-eye coordination resulting from a world, hand and head referenced AR interface. Therewith this research contributes in two ways: first, to space life science and, second, to AR in HCI. While the underlying challenge of this research is thoroughly discussed in section 5.2.3, the description of the parabolic flight studies and their results are presented in Chapter 8.

(5) Introduction of Gravity-Adapted Interface Resizing for Direct AR Selection to Enhance Visuomotor Coordination (LAHC/Weight Study)

Although using the surface of an ISPR would be meaningful to place interactive control AR interfaces, the parabolic flight research has proven that the visuomotor performance on world-referenced interfaces is deteriorated under short-term hyper- and microgravity. To counteract the shortcoming caused by spatial transformation between the body and world coordinate systems, a resizing approach was introduced that uses the present gravity load to influence the size and the distance of the interface's pointing targets. This approach can be applied in two different ways, resulting in either a compressed or elongated interface. While compressed resized interfaces contributes with shorter distances, but implies smaller targets, an elongated resized interface advantages from greater targets, but causes longer distances between the targets. To detect possible effects compared to an unaffected interface and solve the trade-off between the resizing conditions, a usability test was conducted, which initially only considered different hypergravity loads. To provide such environmental condition, hypergravity was induced by two different simulation methods, which resulted in a case study, using a long-arm human centrifuge, and a follow-up experiment using additional arm weighting (Guardiera et al., 2008). Because this research was intended to enhance visuomotor performance on world-referenced AR interfaces, it contributes to AR in HCI, but also to space life science by research under hypergravity. Details on the resizing approach and its usability testing are presented in Chapter 9.1.

(6) Enhancement of Visuomotor Performance of Direct Selection on Nonhaptic AR Interfaces (Haptic Study)

Through the sense of touch, operating direct selection interfaces benefits from haptic feedback that provides short-term regeneration and supports the arm positions sense. In contrast to world- and body-referenced interfaces, head-referenced interfaces never provide haptic cues for tactile sensation. Both, under normogravity and altered gravity conditions, operating such nonhaptic interfaces drastically affects the visuomotor performance and, thus, they are unusable. Nonetheless, using head-referenced interfaces would be meaningful, because of their short reach and permanent presence. In order to compensate the lack of haptic feedback, it was required to search for an appropriated countermeasure. Because an additional stimulation of the isometric force level of the pointing arm amplify the proprioceptive sense of aimed pointing movements (Bringoux et al., 2012), it is conceivable that additional weighting, as was used for simulating hypergravity, as presented in the last paragraph, can enhance the handling of nonhaptic interfaces. Hence, a usability study was conducted, comparing the visuomotor performance revealed under different loads to unloaded performance. To prove that possible effects are exclusively linked to the missing tactile sensation, the study was also performed under a haptic condition. By improving the usability of nonhaptic interfaces, this research not only contributes to AR, but also to VR in HCI. This research on nonhaptic interfaces is reported in Chapter 9.2.

(7) Providing guidelines and principles to support the design of AR-assisted payload operations

As a result of the research conducted in the scope of this thesis, the findings are reworded to provide a set of recommendations and guidelines to support the design of AR-assisted payload operations. These guidelines should complement existing design rules, needed to comply with the concept of usability. Besides guidelines, which are restricted to payload or similar operations, some findings are more general and, thus, represent principles, providing a high level of generality to its application. By means of these guidelines and principles, an interface designer should be able to provide a good conceptual model for AR assisted space payload operations or to continue the iterative design process induced in this thesis. The guidelines and principles are presented as part of the conclusion chapter (see section 10.3).

1.7 Publications and Interdisciplinary Collaboration

In the scope of this thesis several publications were released differing in the type of publication (poster, conference paper, journal article). Conference papers were presented by the first author. All publications are listed below according to their category. Chapters of this thesis document, containing parts of the publications, will be marked as such at their beginning. An interdisciplinary collaboration with the Institute of Cardiology and Sports Medicine of the German Sport University Cologne provided the opportunity for the assessment of biofeedback for evaluating the workload by the heart rate variability during several studies presented in this thesis.

Journal Publication:

(1) Markov-Vetter, D., Moll, E. & Staadt, O. (2013). Verifying Sensorimotoric Coordination of Augmented Reality Selection under Hyper- and Microgravity. *Intern. Journal for Advanced Computer Science (IJACS)*, vol. 3, no 5, pp. 217-226.

Peer-Reviewed Conferences Publications and Book Chapters:

(2) Markov-Vetter, D., Moll, E. & Staadt, O. (2012). Evaluation of 3D Selection Tasks in Parabolic Flight Conditions: Pointing Task in Augmented Reality User Interfaces. *Proc. of the ACM SIGGRAPH Intern. Conference on the 11th Virtual-Reality Continuum and its Applications in Industry (VRCAI '12)*, pp.287-293.
DOI: 10.1145/2407516.2407583

(3) Markov-Vetter, D. & Staadt, O. (2013). A Pilot Study for Augmented Reality Supported Procedure Guidance to Operate Payload Racks On-Board the International Space Station. *Proc. of the 12th Intern. Symposium on Mixed and Augmented Reality (ISMAR '13)*, pp. 1-6.
DOI: 10.1109/ISMAR.2013.6671832

(4) Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2013). The Impact of Altered Gravitation on Performance and Workload of Augmented Reality Hand-Eye-Coordination: Inside vs. Outside of Human Body Frame of Reference. *Proc. of Joint Virtual Reality Conference of EGVE – EuroVR (JVRC '13)*, pp. 65-72.
DOI: 10.2312/EGVE.JVRC13.065-072

(5) Markov-Vetter, D., Zander, V., Latsch, J., & Staadt, O. (2015). The Influence of Gravity-Adapted Target Resizing on Direct Augmented Reality Pointing Under Simulated Hypergravity. *Proc. of the 10th Intern. Conference on Computer Graphics Theory and Applications (GRAPP '15),* pp. 401-411. *(Best Paper Award in the area of Interactive Environments)*
DOI: 10.5220/0005316604010411

(6) Markov-Vetter, D., Zander, V., Latsch, J., & Staadt, O. (2016). Enhancement of Direct Augmented Reality Object Selection by Gravity-Adapted Resizing. In Book: J. Braz, J. Pettré, P. Richard, A. Kerren, L. Linsen, S. Battiato, F. Imai (Eds.), Computer Vision, Imaging and Computer Graphics - Theory and Applications. International Joint Conference, VISIGRAPP 2015, Berlin, Germany, March 11-14, 2015. *Springer International Publishing Switzertland*, pp. 75-96.
DOI: 10.1007/978-3-319-29971-6_5

Conference Publications (abstract submission):

(7) Kuijpers, E.A., Carotenuto, L., Malapert, J.-C., Markov-Vetter, D., Melatti, I., Orlandini, A. & Pinchuk, R. (2012). Collaboration on ISS Experiment Data and Knowledge Representation. *63rd Intern. Astronautical Congress (IAC '12)*, paper-code: IAC-12-D5.2.11.

(8) Markov-Vetter, D., Millberg, J. & Staadt, O. (2013). Mobile Augmented Reality for Space Operation Procedures: A Generic Approach of Authoring and Guiding On-Board Payload Activities. *Proceedings of the 64th International Astronautical Congress (IAC '13)*, 1(14), pp. 4542-4555.

(9) Markov-Vetter, D., Moll, E. & Staadt, O. (2013). The Effect of Hyper- and Microgravity on Visuomotor Coordination of Augmented Reality Selection in Correlation with Spatial Orientation and Haptical Feedback. *Proc. of the 64th Intern. Astronautical Congress (IAC '13)*, 1(14), pp. 35-47.

Poster Publications:

(10) Markov-Vetter, D., Mittag, U. & Staadt, O. (2009). An Augmented Reality supported Rack Guidance System in Space Flight. *Proc. of Joint Virtual Reality Conference of EGVE – EuroVR (JVRC'09)*, pp. 33-34. *(peer-reviewed)*

(11) Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2014). A Proof-Of-Concept Study on the Impact of Artificial Hypergravity on Force-Adapted Target Sizing for Direct Augmented Reality Pointing. *Proc. of IEEE Virtual Reality Conference (VR'14),* pp. 95-96. *(peer-reviewed)*

1.8 Outline of the Thesis

This section gives an overview and a short description of each chapter provided in this thesis. Figure 1.5 demonstrates the outline and the corresponding affiliations of these chapters that should mirror the workflow of the activities.

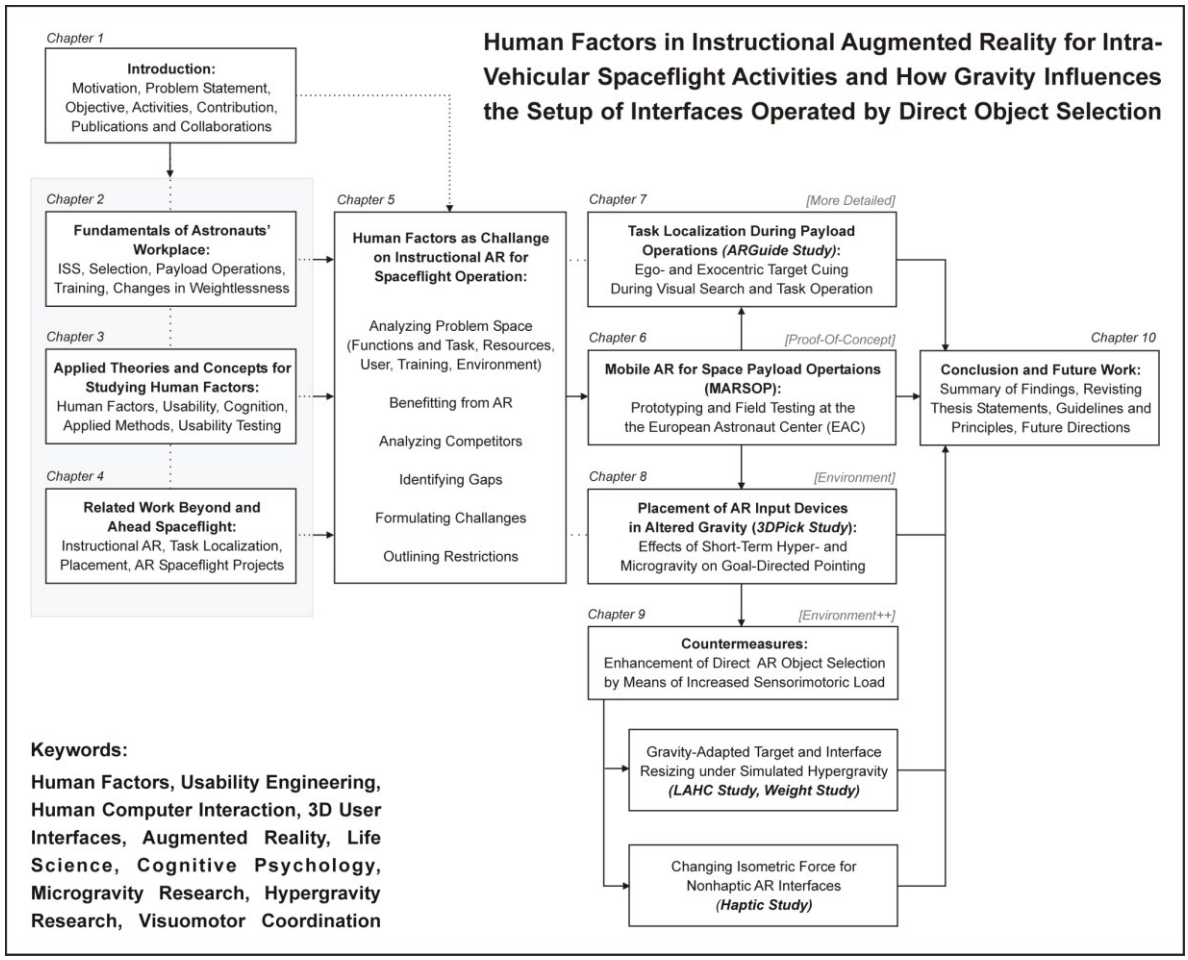


Figure 1.5: Outline and workflow of this thesis.

Chapter 1: Introduction

The introduction chapter initially describes the motivation and problem statement addressed in this thesis. Before clarifying the thesis's meta-objective, which is supported by adequate thesis statements, the concept of Augmented Reality is introduced as advanced user interface technology, established in the field of human-computer interaction. To give an idea about the work that was needed to fulfill the thesis's objective, the thesis activities are listed thereafter, which is followed by pointing out the contributions, reflecting the significance of this research. Besides listing the related publications, this chapter ends with outlining and positioning the thesis chapters, as well as sketching the workflow, which represents the management of the work done.

Chapter 2: Fundamentals of Astronaut's Workplace

To lay the foundations for a better understanding of the problem space of payload operations, this chapter presents the astronauts' working environment, specifically those aspects needed for such operations, which includes the resources provided by the current interface mode, such as the structure of payloads and they arrangement aboard the ISS, the defined standard of procedures to guide payload operations and how they are viewed. It is also briefly described how astronaut are prepared and trained for their mission, especially for operational tasks at payloads. To get an idea, what demands an astronaut candidate has to meet and thus supplying the skills of the intended user group, a typical process of astronaut selection is outlined. Because of the fact that working in space is influenced by environmental factors, fundamentals and related research on changes of human sensorimotor system under weightlessness are mentioned and described.

Chapter 3: Applied Theories and Concepts for Studying Human Factors

More generally, this chapter provides basics about Human Factors engineering. Besides describing what Human Factors means, the concept of usability is defined as well as what design rules are needed to be considered. Because Human Factors claims to meet a user-driven design, and, thus, it is needed to consider the capabilities and capacities of the user, this chapter also provides fundamentals and theories of human information processing. This should support the later analysis of the problem space, presented in Chapter 5. Finally, this chapter is completed by describing the methods applied in this thesis to analyze the users' needs.

Chapter 4: Related Work on Instructional AR Beyond Spaceflight and Advanced User Interfacing for Spaceflight Operations

This chapter is intended to give an overview of selected works done in the field of instructional Augmented Reality for maintenance, assembling and servicing tasks. Thereby it is taking a closer look at works complying with the claim of covering real-world problems and verifying that they are benefitting from AR. Additionally, this chapter presents a set of reviews of research on task localization and the placement of AR input interfaces, both playing an important part when AR is adopted for payload operations, and not considered so far. While the prior reviews considered only works conducted beyond the space flight sector, the last section is directed towards works and projects of advanced user interfacing to ease astronauts' work, especially by AR. Because the number of such works is limited, five projects will be presented in detail, which supports the competitor analysis presented in Chapter 5.

Chapter 5: Human Factors as Challenge on Instructional AR for Spaceflight Operation

To identify challenges tackled in the scope of this thesis, this chapter discusses in detail the problem space in question, which is supported by the knowledge yielded from the previous chapter. Following common inspection methods to analyze user's needs, it presents the results of various analyses of the current mode of payload operations. Taking this into account, it is clarified why payload operations could benefit from AR. By discussing the works done by competitors, the gaps needed to be filled are specified, which is followed by detailed descriptions of the identified challenges to corroborate their needs. On order that the reader of this thesis knows which related issues are not touched, the restriction and delimitations of this work are outlined finally.

Chapter 6: Proof-Of-Concept: Prototyping and Field Testing of Mobile AR for Space Operations (MARSOP Study)

This chapter outlines and describes the components of the initial prototype, providing Mobile AR for Space Operations (MARSOP), while overcoming the drawbacks of the current guidance mode presented in Chapter 5. Background information required to compile an full-fledged AR prototype and not explained up to here, will be clarified. The core part of this chapter is formed by the performance of a proof-of-concept study that meets the requirement for controlled field testing. To comply with this request, the MARSOP system was field tested in the European Astronaut Center (EAC) in Cologne, Germany. Thus, this chapter thoroughly describes the conditions and results of this field test, which considered a real payload environment and application domain experts as representative users.

Chapter 7: Task Localization During Payload Operations: Ego- and Exocentric Target Cuing During Visual Search and Task Operation (ARGuide Study)

Besides understanding an instruction, the astronaut has to spatially process the given target cues to localize the task area at the payload. To examine more closely the effect of spatial processing during visual navigation, this chapter presents research on task localization, which can be affected by the way the target cues are spatially oriented. By considering an ego- and exocentric orientation, not only of the display, but also of the visualization scheme, this chapter describes a study investigating the trade-off between appropriate correlations. In addition to results yielded from analyzing the visuomotor performance, this chapter presents results from analyzing the workload level, which was assessed in its full range, using a dual-task model, biofeedback and subjective experiences.

Chapter 8: Placement of AR Input Devices in Altered Gravity: Effects of Short- Term Hyper- and Microgravity on Goal-Directed Pointing (3DPick Study)

This chapter reports on experimentation on spatial placement of AR input devices that are operated by direct pointing. Such interfaces can be placed by a world, body or head reference. Unlike the studies presented in Chapter 6 and Chapter 7, this research takes the environmental conditions, in which an astronaut has to work, into account. This implies modified conditions of the gravitational force. To study the trade-off between the placement references in compliance with this demand, two experiments under parabolic flight conditions were conducted. Thus, this chapter contains experimental results from studying the effect of altered gravity on visuomotor coordination, resulted from pointing towards world-, body- and head-referenced interfaces. The results are supplemented by taking a closer look at the underlying pointing movement. By analyzing the target's mislocalization and the trajectory of the aimed movement, this motion analysis aims to make a contribution to the understanding of the findings.

Chapter 9: Countermeasures: Enhancement of Direct AR Target Selection by Means of Sensorimotoric Load

As revealed by research conducted in Chapter 8, the visuomotor coordination resulted from pointing towards interfaces that are referenced to world's objects and to user's head is not optimal. This chapter reports on potential countermeasures to improve their handling. Therefore two approaches using increased sensorimotoric load are introduced. While world-referenced interfaces are enhanced by applying gravitational information to characteristics of the interface, the non-usability of head-referenced interfaces is counteracted by exaggerating the isometric force level of the pointing arm. Both approaches and related experimentation are described and discussed within this chapter.

Chapter 10: Conclusion

The design process to apply Augmented Reality to space payload operation resulted in six user studies. In this chapter their findings are summarized, which should support the following review of the initial thesis statements. By verifying their fulfillments, it is discussed whether the meta-objective of this thesis was achieved. Afterwards, the findings are reworded to provide a set of guidelines and principles to support the process of designing AR interfaces for space payload operations from a Human Factors perspective. Finally, the future work section is intended to clarify the next steps to continue the work presented in this thesis.

Appendix A

This appendix presents the NASA Raw Task Load Index (NASA RTLX) to measure subjective experience of mental workload. This rating scale was used during several user studies presented in this thesis.

Appendices (B-I)

These appendices are arranged by the chapters and named after their related study acronym. They provide additional contextual materials, such as file formats, study tasks, resources, subjective feedback, sample parabolas and schematics of experiment installations, but mostly descriptive and inferential statistics not presented in the chapters so far.

Chapter 2

Fundamentals of Astronaut's Workplace

"Initially, when a problem is first presented, it must be recognized and understood."

(Alan Newell & Herbert A. Simon, *Human problem solving*, 1972)

Payload operations are conducted by astronauts aboard the International Space Station (ISS). This broadly defines the problem space covered in this thesis. To increase its understanding, this chapter provides information about fundamentals of astronaut's workplace, paying particular attention to International Standard Payload Racks and their operating mode supported by standardized procedures, which are developed by User Support Operation Centers and provided by the international Procedure Viewer, running on a common laptop computer. Because payload operations are exclusively operated by astronauts the criteria for their selection are outlined, and therewith constitute the profile of the intended user group targeted by this thesis. Before astronauts embark to their missions in space, they will be trained, in particular for payload operations. Describing these trainings should enable to understand the astronauts' level of knowledge and experiences, on which their later task performance is based. The last important aspect considers that astronauts need to carry out their work in the absence of gravity. Hence, the effect of weightlessness on the human sensorimotor system is explained and research on it is outlined.

2.1 Working in Space aboard the International Space Station

Traveling in space aboard the International Space Station (ISS), while orbiting Earth every ninety minutes, astronauts' surrounding provides a unique and most unusual place for a human to work. The astronauts' workplace is located in a nearly circular Low Earth Orbit (LEO) with an altitude ranging from 350 km to 460 km and an inclination of 51.6 degrees to Earth's equator. Due to the low altitude and the ISS's dimensions a periodic reboost every 10 to 45 days is necessary to keep the ISS on its orbit (NASA, 2000). The ISS (see Figure 2.1) is dimensioned with a length of 73 meter, a width of 108 meter and a height of 20 meter that results in a pressurized volume of 916 m^3, whereby 388 m^3 volume can be used for habitable purposes, and thus, the ISS is the largest artificial manned space station that has ever been in orbit (NASA, 2015). Travelling to the workplace on the ISS, an astronaut has to accept a journey of around six hours per space flight, today with the Russian Soyuz TMA-M spacecraft, where his or her body experiences altered gravity[9] conditions from 1*g* up to 4*g* on launch, close to 0*g* in orbit and up to 10*g* during reentering the Earth's atmosphere. Today, an astronaut or cosmonaut will stay up to six months on the ISS that is operated by a full time crew of six astronauts of different nationalities. To ensure a permanent occupation, the six crew members are alternating with two separate shuttle missions with three astronauts per shuttle and an overlap in their stays.

[9] The Earth gravitational force, or shortly gravity, is almost always the same in low orbits. Thus, the term „altered gravity“ is not correct in its physical meaning. In free fall or with an increased acceleration only the weight force affecting the body is changing. Because of its common use, the term "gravity" is defined as the weight force in the scope of this thesis.

Figure 2.1: The International Space Station as astronauts' workplace. The photo shows the ISS from space shuttle Endeavour. [Photo © NASA, 2010]

Inside this orbiting space station, the physical circumstances provide a microgravity laboratory that enables astronauts to conduct research on behalf of the scientists on the ground that is focused on gaining a better understanding of the effect of microgravity[10] to enhance knowledge in space, physical, and biological sciences. To provide habitable place to live and work (see Figure 2.2), the current ISS configuration is, beside others, featured with three nodes and five modules, which are mainly pressurized cylinders. Accommodations are located in the Russian Zvezda module (Service Module, SM), the U.S. Unity node (Node 1), the U.S. Destiny Laboratory, and the European nodes Harmony (Node 2) and Tranquility (Node 3). These include compartments for waste management, the toilet, the galley, individual crew sleeping compartments, and some of the exercise equipment, like a bicycle ergometer and diverse treadmills (NASA, 2015). Beside of external research accommodation, working inside the ISS is possible almost everywhere in the modules and nodes. While flight control and propulsion systems are mainly placed in the U.S. Zarya (Functional Cargo Block) and the Russian Zvezda module, research accommodations to support microgravity science investigations are mainly located in the research modules (U.S. Destiny, European Columbus, Japanese KIBO). While the Russian Zvezda module was originally intended only for crew habitation, it is meanwhile featured as multipurpose research laboratory on the ISS. Until now, the five space agencies, NASA, Roscosmos, ESA, JAXA and CSA, have hosted investigators from 83 nations to conduct over 1700 experiments under long-term microgravity conditions aboard the ISS (NASA, 2015). Thus, major research activities are conducted by the astronauts using research racks that house the experiment hardware and materials. Currently the ISS provides 34 scientific racks with 13 racks in the U.S. Destiny laboratory, 10 racks in the European Columbus laboratory and 11 racks in Japanese KIBO laboratory. If astronaut's mission includes the performance of experiments using a

[10] Contrary to Earth condition or normogravity, where the gravity is defined as $1g = 9.81$ m/s^2, microgravity condition creates an environment where objects and humans are accelerated relative to their surroundings with a reduced gravity level of 9.81×10^{-6} m/s^2. In this thesis the term *microgravity* is used as synonym for weightlessness and 0g.

research rack, he or she needs to be extensively familiarized and trained with the specific rack and the experimental procedure.

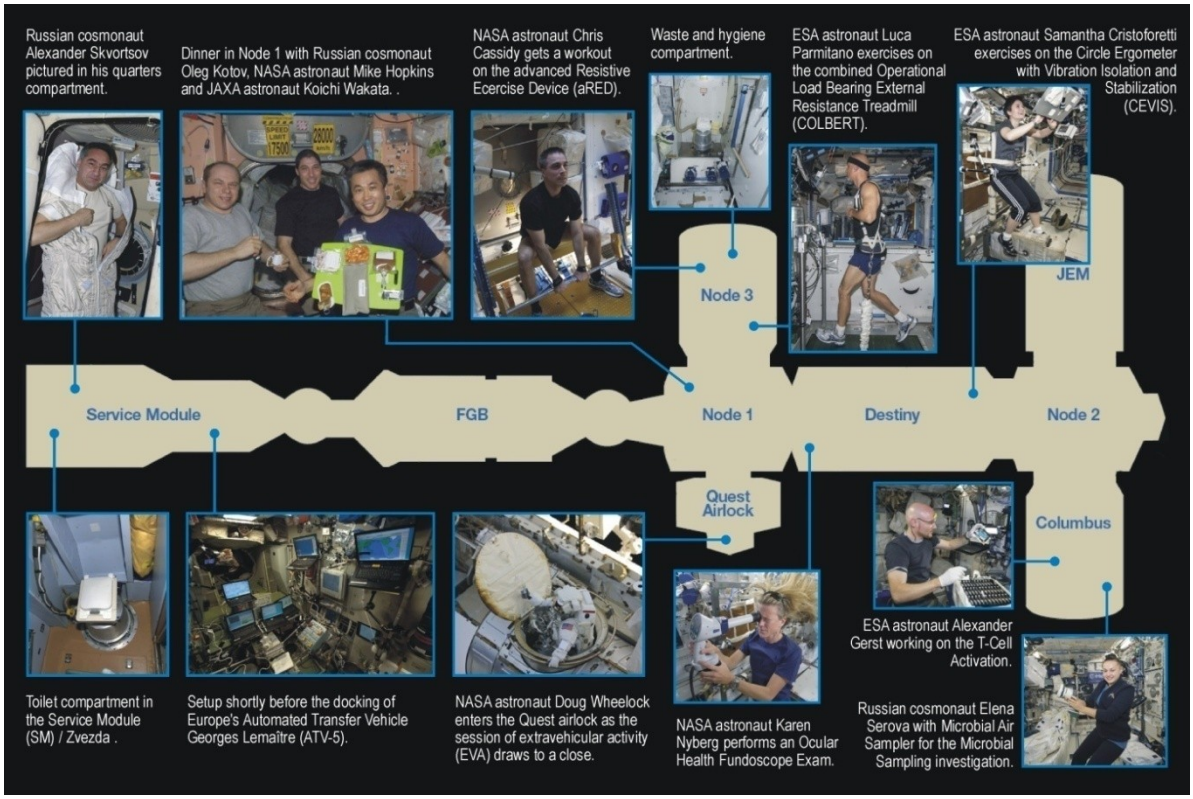


Figure 2.2: The habitable modules and nodes of the ISS. They are providing working and living space for the flight crewmembers (adapted from NASA, 2015).

Because a space mission requires huge efforts at high costs, crew time is rare and precious. Thus it is important to efficiently benefit from the astronauts' abilities and operations. Therefore the daily routine of an astronaut is scheduled in detail by the ground mission control and there is not much room for unplanned outages. Besides command, control, maintenance and housekeeping duties, within a weekly work schedule about 35 hours the astronauts are conducting experiments for research in biology, human physiology, physical and space sciences. Thereby astronauts are not only responsible for preparing and conducting scientific experiments, sometimes they also need to act as subject during some of those experiments.

2.2 Astronaut Selection

Working in space is an outstanding job that requires a high level of perceptual, cognitive and psychomotor skills. A careful astronaut selection ensures a certain degree of aptitude to meet its high demands, but is no warranty for the best performance at work. In space, far away from the Earth, human errors may have serious consequence causing extremely high costs and efforts, or even life-threatening situations. Thus, prior exclusion of psychological and physiological dysfunctions provides the best prerequisites to master these extreme working conditions. To be chosen as an astronaut, applicants have to meet a range of strict criteria and pass a wide range of tests. This selection process will be outlined based upon the last ESA selection campaign, which was initiated in May 2008 searching for a six person crew, whereby the final decision was announced one year later,

in May 2009. As shown in Figure 2.3, before being invited to a specific job interview, each applicant needed to pass four stages of pre-selection: online application, psychology test 1, psychology test 2 and medical examination. Generally everyone who met the following formal qualifications could apply for the astronaut job: (ESA, 2009)

- university degree in natural science (e.g., physics, biology, chemistry), engineering or medicine;
- three years postgraduate professional experience or experience as a pilot;
- ESA Member State nationality;
- not older than 55 years (preferred age ranged from 27 to 37);
- pilot medical certificate (*JAR-FCL 3*[11] Class 2 or equivalent).

In 2008, 8413 Europeans applied to become an astronauts, completed the online application form and provided the medical certificate. Only about eleven percent of them complied with the formal requirements and were accepted for further tests. These candidates were between 24 and 46 years old, with a mean of 33.2 (Maschke et al., 2011). In the following psychological selection the cognitive capabilities and personality traits of the applicants were sequentially tested. These tests were conducted in Hamburg, Germany, at the Psychology Department of the Institute of Aerospace Medicine of the German Aerospace Centre (DLR) in cooperation with the French Institute for Space Medicine and Physiology (MEDES). Beside skills in englisch language, engineering, technology, physics and mathematics, the first psychology selection included various cognitive tests to detect impairments in: (Maschke et al., 2011)

- mental arithmetic, memory function, attention, perception, spatial comprehension, reasoning, psychomotor coordination, multi-tasking and work capacity
- relevant basic personality skills (emotional stability, conscientiousness, agreeableness, openness, self-confidence, flexibility and resilience)

Only about 21 percent of the 912 candidates mastered these tests successful and were invited to the second phase of the psychological evaluation. This was more focused on applicants' personality skills diagnosing the behavior-oriented personality and investigating skills related to problem-solving and cooperativeness, which was assessed by the *Dyadic Cooperation Test*[12] involving decision making, reliability, working style and stress resistance. This psychological phase concluded with an individual structured interview in front of a selection committee, whereby the results, obtained from personality questionnaires acquired in the psych I phase, formed the basis of it. From the 192 candidates tested, only 45 complied with the high requirements applied to the psychological profile and competed in the medical test that was conducted by the Institute of Aerospace Medicine (DLR) in Cologne again in cooperation with MEDES. In the scope of the medical examination the candidates were requested to provide information about their medical history, current health state, their family medical history and diseases, as well as about their social habits and lifestyle. It also implied a set of clinical

[11] JAR-FCL 3 (in German). URL: http://www.luftrecht-online.de/banz/JAR-FCL3-deutsch_vE.pdf, last visit: 28.09.2016

[12] The Dyadic Cooperation Test is part of the German pilot test, also referred to as "DLR test". It is a computer aided test to examine skills in cooperative problem solving. Under time pressure the participants has to solve a cognitive task (loading of bridges with different weights), first alone than in a two-person team communicating via a headset and alternating the roles as operator and coordinator. Thereby the task is interfered by disturbances and further sub-tasks. This enables to detect applicants' behavioral traits in taking ones role and responsibilities for others. (Schuler, 2007, p. 318)

examinations, like anthropometry and diagnostic imaging (Magnetic Resonance Imagery, Computerized Tomography, Carotid Doppler, Echocardiography), and was finished with a review by the ESA Medical Selection Board (Pruett et al., 2013). Typically, vision is the most common reason for disqualification of applicants in aviation and spaceflight, but surprisingly, this time most candidates were disqualified by cardiovascular causes. The 22 candidates that passed the medical tests successfully, were invited to a job interview. In May 2009, one year after the announcement, the ESA finally decided for the following six astronauts: (see photo in Figure 2.3, from left to right) Andreas Mogensen (DK), Alexander Gerst, (DE), Samantha Cristoforetti (IT), Thomas Pesquet (FR), Luca Parmitano, (IT) and Timothy Peake (UK). Until now, the ISS has been continuously staffed with an international crew resulting in more than two hundred astronauts from over fifteen countries.

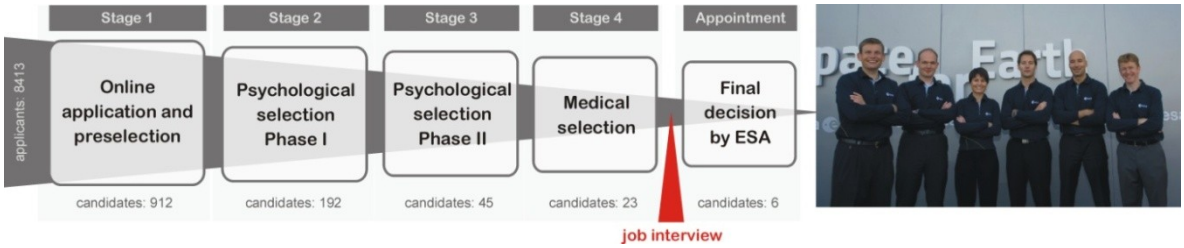


Figure 2.3: Stages and resulted candidates of the 3rd ESA astronaut selection campaign from 2008 until 2009 (adapted from Maschke et al., 2011; numbers from Pruett et al., 2013; Photo © ESA–S. Corvaja, 2009).

2.3 Intravehicular ISS Payload Operations

The working schedule of astronauts is tight and a majority of their daily working time is intended for maintenance and experimental tasks at research payloads, whereby the duration of one session can range from half an hour up to several. This section will consider fundamentals of such payloads in general and the Biolab in particular, as well as how they are operated by standardized procedures and how astronauts are trained for such operations.

2.3.1 International Standard Payload Racks

As mentioned above, currently there are 34 internal ISPRs installed in the ISS's research modules, except in the Russian Zvezda module that does not allow the accommodation of ISPRs. Besides internal (or pressurized) payloads that are installed inside the ISS, external (or unpressurized) payloads are attached at the outside of the ISS modules, like the Matroshka facility, the first external ESA payload. Matroshka is attached to the Zvezda module and consists of a human phantom upper torso that enables the housing of radiation measurement devices to study the effect of the orbital radiation field on astronauts' organs during an extravehicular activity (EVA) (Dettmann et al., 2007). The Columbus laboratory, for example, with a volume of 75 m³ provides the capacity of sixteen internal racks (see Figure 2.5, middle and right) that are arranged in four segments with four racks respectively. While six racks located at the deck and overhead are used to host the subsystems and for storage purposes, the remaining ten racks are active racks designed for scientific and technological experiments in the fields of material, fluid, biological and physiological sciences (see Figure 2.4). The available versatility and resources also provide a suitable environment for other researches in applied sciences, process engineering and prototyping of automatic experiments (ESA, 2015b).

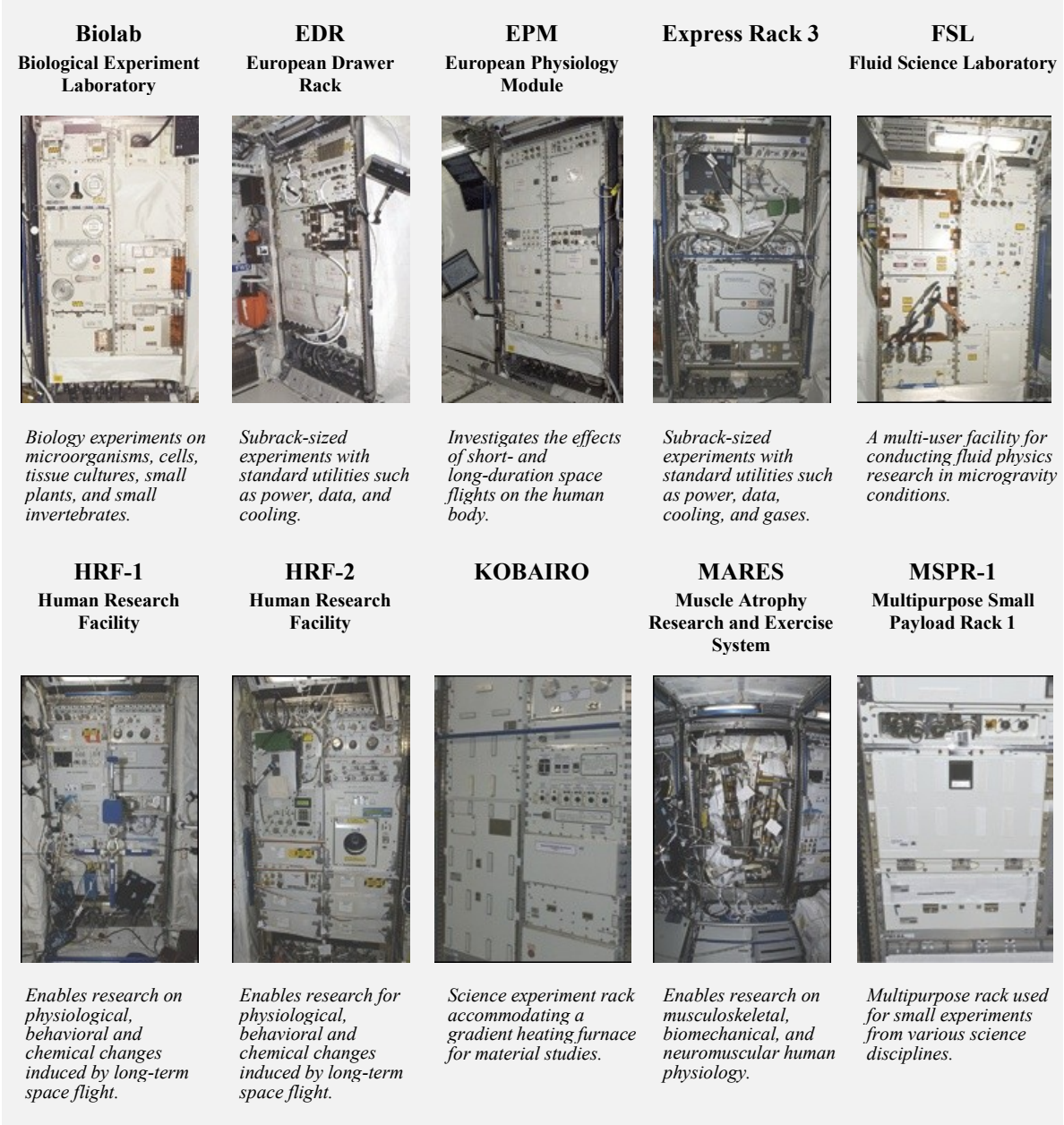


Figure 2.4: The research payloads housed in International Standard Payload Racks accommodated inside the ESA Columbus laboratory (adapted from NASA, 2015).

To carry out scientific and technological experiments under microgravity conditions aboard the ISS, scientists can either bring their own research facilities to the ISS or can use firmly installed payload facilities that enable research in biology, biotechnology, material science, physical sciences, technology development and life science. A research payload within the Destiny, Columbus and KIBO laboratories is usually housed in a standard rack, referred to as International Standard Payload Rack (ISPR). Individual small or medium sized payloads can be fit in ISS lockers carried in a rack framework, which allows a fast and easy installation (NASA, 2015). To house research hardware, an ISPR (see Figure 2.5, left) provides 1.35 m^3 of internal volume with outer dimensions of 1.046 m width, 2.014 m height and 0.858 m depth (ESA, 2015b).

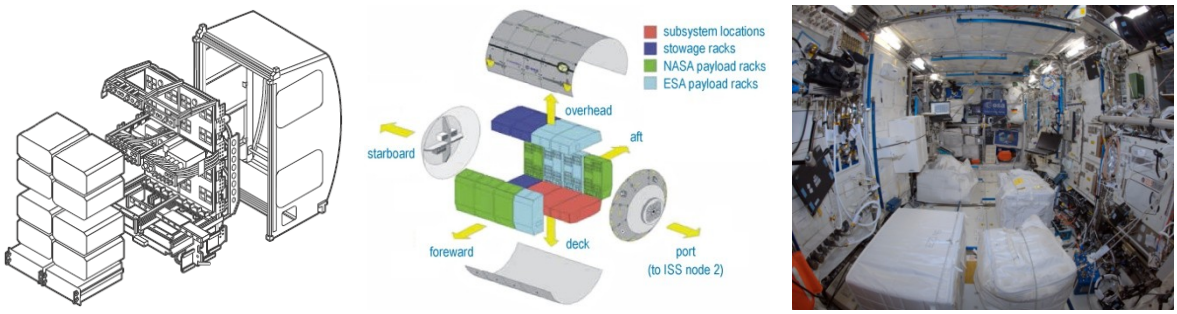


Figure 2.5: International Standard Payload Rack for research aboard the ISS. *(left)* Exploded assembly drawing of an ISPR (from NASA, 2015). *(middle)* The layout of the ISPRs inside the ESA Columbus module (adapted from ESA, 2015b) *(right)* An interior view of the Columbus module [Photo © ESA/NASA, 2014]

2.3.2 The Biolab Payload

The Biolab payload is an ESA multi-user facility and part of the European Columbus laboratory. Preinstalled within the Columbus, the Biolab was launched on board the Space Shuttle Atlantis in February 2008. The Biolab is a modular system housed in an ISPR and is intended for the performance of biological experiments to study the effect of microgravity and space radiation on cell cultures, microorganisms, small plants and small invertebrates. The Biolab's subsystems (see Figure 2.6, left) enable the incubation and preparation of experiment samples, storing them at cool or cold temperatures. Biological samples can also be microscopically or photometrically analysed in orbit, while telescience capability allows scientists on ground to control and monitor their samples and to implement procedural changes just-in-time, if necessary (Brinckmann, 1999; Schuber et al., 2013).

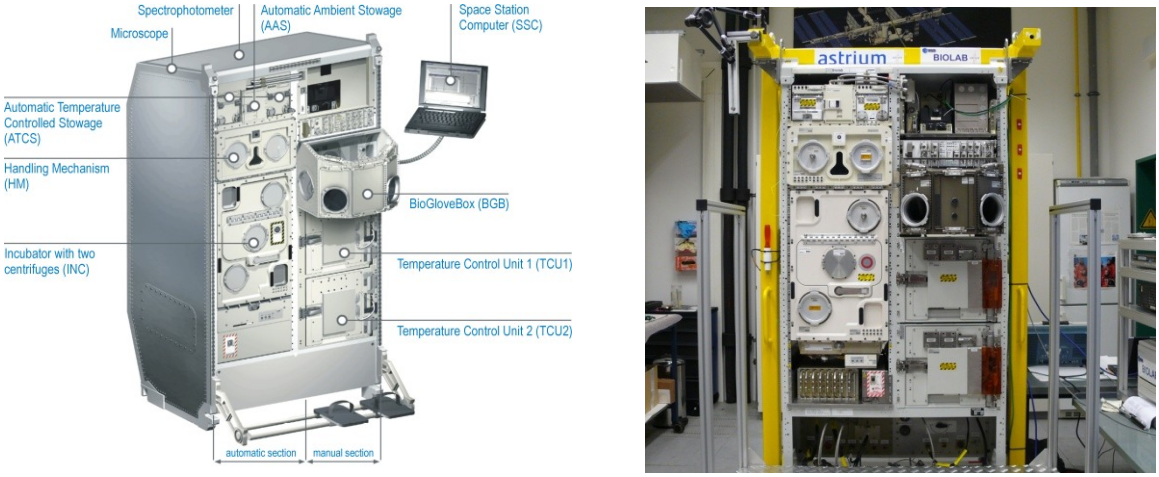


Figure 2.6: The ESA Biolab payload. *(left)* Sections of the Biolab with its components (adapted from ESA, 2015c, and Brinckmann, 1999). *(right)* The engineering ground model hosted by the DLR in Cologne.

The Biolab payload is divided into an automatic and a manual section. Thereby the automatic left part can be controlled from the ground and allows the automated performance of experiments following manual preparation and loading by the astronaut. To enable these features the automatic part is equipped with an Incubator (INC), a Handling Mechanism (HM), two Automatic Temperature Controlled Stowages (ATCS), an Automatic Ambient Stowage (AAS), and diverse analysis instruments (microscope, spectrophotometer). As core part of the Biolab payload, the incubator provides a Life Support System (LSS) and contains two identically centrifuges on two rotor platforms (see Figure 2.7, left). To house the biological samples, one centrifuge can carry up to four Experiment

Containers (EC; see Figure 2.7, right) and two advanced ECs. The LSS enables an environment where the temperature, illumination, gas composition and humidity can be actively controlled. The ECs can be individually isolated from the LSS air flow, which is required in case of contamination. To perform activities on the ECs, a robotic arm, referred to as Handling Mechanism (HM), is, for example, capable to operate the analysis instruments or to transfer fluids between the ECs and the Automated Temperature Control Stowages (ATCS). While the ATCSs provide the storage of samples at low temperatures (from -20°C to +10°C), the AAS facilitates the storage with no active temperature control for the fluids, syringes, tools or other items needed for the experiments, and both facilities are accessible by the Handling Mechanism (Brinckmann, 1999). The manual right part of the Biolab payload is intended to store and prepare the biological samples. It consists of the BioGloveBox (BGB), two Temperature Control Units (TCU) and the Station Support Computer (SSC). Using the BGB enables manual activities on the samples in an enclosed and clean area with an adjustable filtered air stream. To prepare an experiment, the astronaut loads an EC with experiment hardware or new specimen within the BGB. The working volume of the BGB is equipped with interfaces for power and data and contains a video observation system. The two TCUs are available for storing samples and chemicals within an active temperature controlled environment, where the temperature can be set between -20°C and +10°C. The SSC is a laptop computer to control and manage the experiment schedule. While displaying the instructional procedure, it provides also the monitoring of the Biolab's telemetry to enable the mission ground team to remotely control the environment (Brinckmann, 1999).

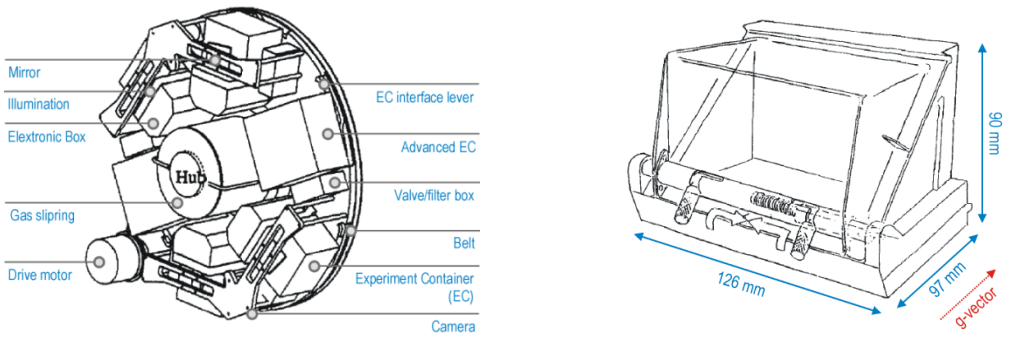


Figure 2.7: The Biolab centrifuge (adapted from Brinckmann, 1999). *(left)* One Biolab centrifuge rotor (Ø 600 mm) holding four Experiment Container (EC) and two advanced ECs. *(right)* A Biolab Experiment Container (EC) with its interface to the centrifuge on the bottom plate.

2.3.3 European User Support and Operation Center

User Support and Operations Centers (USOCs) were established in 1998 by ESA as decentralized infrastructure for the operation of European payloads aboard the ISS. The USOCs are ground laboratories and control centers distributed across Europe (see Figure 2.8, left). They are responsible for operating and implementing European payloads on the International Space Station in cooperation with the scientists and engineers on the ground. Thereby the European USOCs have different roles depending on the technical and operational responsibilities, and either act as Facility Responsible Center (FRC) or as Facility Support Center (FSC) on behalf of the ESA. Thus, each USOC has the authority or expertise for specific payloads. The associated FRC for the Columbus Fluid Science Laboratory, for example, is the Italian MARS in Naples which is supported by the technical expertise of the Spanish E-USOC in Madrid. If a new task or experiment is planned to launch to the ISS, the corresponding FRC assisted by their FSC is responsible to prepare, control and monitor the

experiment in ongoing coordination with the scientists, space hardware developers and the Columbus Control Center (Col-CC) (see Figure 2.8, right). Therefore the USOCs design and develop procedures to optimize and calibrate the payloads and experiments, and finally to support the astronauts' training (ESA, 2015d). After an experiment is successfully launched, the USOCs together with the scientists are remotely interfacing with the astronauts in orbit via telescience to advise and support them during their task operations.

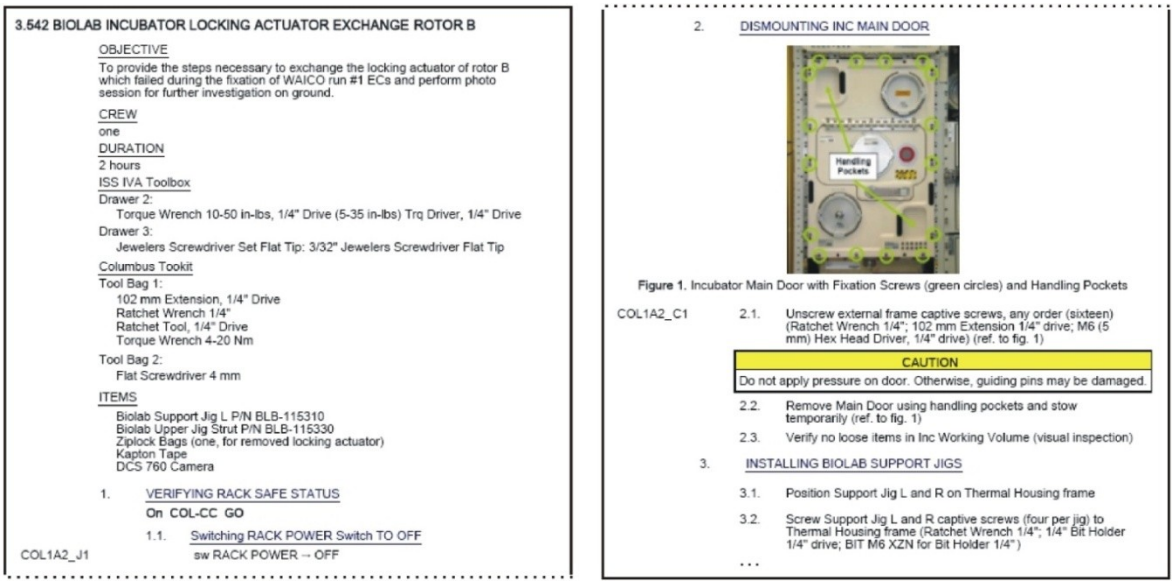
Figure 2.8: The User Support Operation Center. *(left)* The eight European USOCs [Photo © ESA, 2014]. *(right)* The Columbus Control Center (Col-CC) at DLR in Oberpfaffenhofen, Germany [Photo © ESA, 2015].

The Biolab payload, for example, is controlled and monitored by the Microgravity User Support Center (MUSC) as Facility Responsible Center (FRC) at DLR (Cologne, Germany). It is supported by the Biotechnology Space Support Center (BIOTESC, Lucerne, Switzerland), the associated Facility Support Center (FSC), holding the expert knowledge for biomedical research applied to the Biolab payload (Schuber et al., 2013). While an engineering ground model of the Biolab payload, hosted by the DLR in Cologne (see Figure 2.6, right), is used for experimental pre-tests, the astronauts are becoming familiar with the experiment procedures using a training model (see Figure 2.13) that is located at the European Astronaut Center (EAC) in Cologne, Germany.

2.3.4 Standardized Procedures for Operating the ISS

Operating the ISS is a complex ongoing process and astronauts have to conduct a tremendous amount of work to overcome this challenge. While the members of the flight crew are conducting tasks where their physical presence is required, the mission control and support teams can autonomously trigger tasks from the ground that do not need any interaction with the crew, for example for monitoring or commanding purposes. However, crew tasks are regularly accompanied by the ground team to supervise and verify the performance or to conduct a task in cooperation with the astronaut by mandatory callouts (Uhlig et al., 2015). During the task performance the astronaut directly interacts with the mission control center, which in turn is in direct contact with the Facility Responsible Center (FRC) assisted by their Facility Support Center (FSC). To ensure a clear and fast communication during this multi-facility cooperative task performance, it is important to use a unique standardized terminology for ISS operations. Therefore most of the ISS tasks are based on a collection of procedures that are defined by standardized Operations Data Files (ODF), whereby SODFs are assigned to Station tasks and PODF to Payload tasks. This ODF standard was developed by the International Partners of the ISS program and covers most crew and ground activities. As shown in

Figure 2.9, an ODF procedure consists of a leading and an instructional part. Thereby the leading part of the ODF format always itemizes the objective, the number of participating crew members, the task duration, as well as the necessary tools, items and materials, while the instructional part contains the task steps, which are step-by-step instructions describing a series of actions that must be performed in a sequential manner. Instructions can be complemented by visual aids, which "[...] are used for illustration, clarification, or indication of location, but in most cases as complement to a textual instruction [...]" (Aguzzi & Lamborelle, 2012, p. 2).



3.542 BIOLAB INCUBATOR LOCKING ACTUATOR EXCHANGE ROTOR B

OBJECTIVE

To provide the steps necessary to exchange the locking actuator of rotor B which failed during the fixation of WAICO run #1 ECs and perform photo session for further investigation on ground.

CREW

one

DURATION

2 hours

ISS IVA Toolbox

Drawer 2:
Torque Wrench 10-50 in-lbs, 1/4" Drive (5-35 in-lbs) Trq Driver, 1/4" Drive

Drawer 3:
Jewelers Screwdriver Set Flat Tip: 3/32" Jewelers Screwdriver Flat Tip

Columbus Tookit

Tool Bag 1:
102 mm Extension, 1/4" Drive
Ratchet Wrench 1/4"
Ratchet Tool, 1/4" Drive
Torque Wrench 4-20 Nm

Tool Bag 2:
Flat Screwdriver 4 mm

ITEMS

Biolab Support Jig L P/N BLB-115310
Biolab Upper Jig Strut P/N BLB-115330
Ziplock Bags (one, for removed locking actuator)
Kapton Tape
DCS 760 Camera

1. VERIFYING RACK SAFE STATUS

On COL-CC GO

1.1. Switching RACK POWER Switch TO OFF

COL1A2_J1 sw RACK POWER → OFF

2. DISMOUNTING INC MAIN DOOR

Handling Pockets

Figure 1. Incubator Main Door with Fixation Screws (green circles) and Handling Pockets

COL1A2_C1 2.1. Unscrew external frame captive screws, any order (sixteen) (Ratchet Wrench 1/4"; 102 mm Extension 1/4" drive; M6 (5 mm) Hex Head Driver, 1/4" drive) (ref. to fig. 1)

CAUTION

Do not apply pressure on door. Otherwise, guiding pins may be damaged.

2.2. Remove Main Door using handling pockets and stow temporarily (ref. to fig. 1)

2.3. Verify no loose items in Inc Working Volume (visual inspection)

3. INSTALLING BIOLAB SUPPORT JIGS

3.1. Position Support Jig L and R on Thermal Housing frame

3.2. Screw Support Jig L and R captive screws (four per jig) to Thermal Housing frame (Ratchet Wrench 1/4"; 1/4" Bit Holder 1/4" drive; BIT M6 XZN for Bit Holder 1/4")

...

Figure 2.9: Extract of an ODF procedure for the Biolab Payload working at the incubator.

Currently the ODF standard permits only the utilization of static two dimensional visual aids, such as pictures, drawings or schematics. As stated by Aguzzi & Lamborelle (2012), visual aids "[...] are already today an important part of the procedures used on board the ISS" (p. 8), and will become more complex and dynamic in the future, whereby the application of videos, 3D visuals and animations is expected. While most instructions are intended to support payload tasks in a direct way, some instructions demand to enter data or commands via the keyboard under directives, also given by the ODF standard. Considering the compilation process, each ODF procedure is designed and reviewed by multiple experts, and is subjected to conditions by specific rules and syntax (Pels et al., 2004).

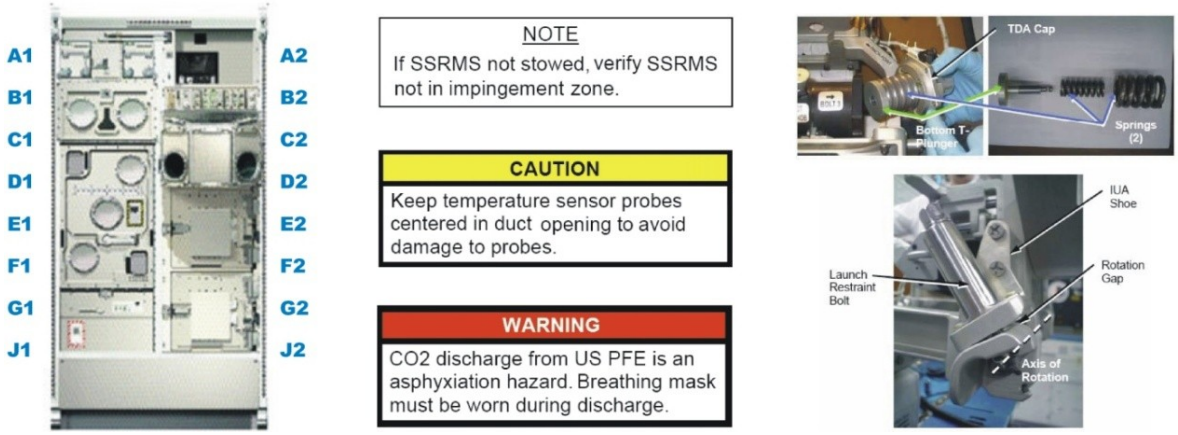


Figure 2.10: Examples of standardized ODF rules. *(left)* The location codes of the Biolab *(middle)* Pre-defined visual appearance of notes, caution and warning. *(right)* Positioning rules for labeling of embedded figures.

This ODF directive consists of more than eighty format settings and standards. Beside rules that are specifying how an instruction must be compiled, there are also rules, for example, indicate how an instruction must be linked to its location code or visual aids, how an embedded figure has to be labeled with annotations, or to specify the appearance of notes, cautions and warnings that have to be placed directly in front of the description of the operation they are clarifying (see Figure 2.10).

ODF procedures are generated by using the *Procedure Authoring Tool* (PAT) and stored as XML files that can be accessed and executed by the *international Procedure Viewer* (iPV), a web-based client server application, which is also replicated on the ground for control centers use. PAT combines XMetaL[13] for generating XML files, and a collection of ODF-specific customizations to create and modify ODF procedures in an easy way. The deployment of PAT[14] and iPV[15] was the result of a cooperative work between NASA and ESA. While the NASA Johnson Space Center was responsible for developing the PAT authoring tool, the ESA developed together with their industrial partners EADS (meanwhile Airbus Group SE[16]) and Skytek[17] the viewer component (Bauer et al., 2003). Besides the execution of procedures, the iPV also provides public and personal annotations, as well as a virtual control panel, which can be used to send and receive actual telemetry data and telecommands. Besides electronic versions using the iPV on astronauts' SSC laptop computer (see Figure 1.2, right), ODF procedures are additionally available on a paper medium (see Figure 2.11) that is also defined by standards and guidelines of the ODF format (Pels et al., 2004).

Figure 2.11: Crewmembers using paper-based ODF procedures. *(left)* ESA astronaut Thomas Reiter looks over a procedure in the Destiny lab. [Photo © NASA, 2006] *(middle)* ESA astronaut Hans Schlegel works at the Biolab in the Columbus while reading a paper-based ODF procedure. [Photo © NASA, 2008] *(right)* NASA astronaut Mike Hopkins reads procedure manuals in the Zvezda Service Module. [Photo © NASA, 2014]

2.3.5 Astronaut Mission and Payload Training

Working aboard the ISS implies to master various complex tasks and to cope with unforeseen situations, thereby posing an enormous challenge for every astronaut, whose skills and performance are decisive factors and contribute significantly to the success of a space mission. Thus, before astronauts embark to their missions in space, they will be trained for long periods in different countries. In contrast to the past, where an astronaut was individually trained to be an expert for a limited number of tasks, today's astronauts should have a common level of knowledge and shall be able to accomplish all basic ISS tasks besides their mission specific duties. Depending on the astronauts' assignment, each ISS partner with its space agency is responsible for training on

[13] XMetaL, URL: http://xmetal.com, last visit: : 28.09.2016
[14] PAT, URL: http://spaceflight.esa.int/eo/EOI/esa-iPV-site/PAT-Training-v1-7-FINAL.pdf, last visit: : 28.09.2016
[15] iPV, URL: http://spaceflight.esa.int/eo/EOI/esa-iPV-site/esa_iPV_index.htm, last visit: : 28.09.2016
[16] Airbus Group SE, URL: http://www.airbusgroup.com, last visit: : 28.09.2016
[17] Skytek - iPV; URL: http://www.skytek.com/wp-content/themes/skytek/assets/pdfs/ipv_a4.pdf, last visit: : 28.09.2016

individual levels to handle their contributed hardware. The Johnson Space Center (JSC) in Houston, the Gagarin Cosmonaut Training Center (GCTC) in Star City and the European Astronaut Center (EAC) in Cologne (see Figure 2.12, left) are some of the world's biggest training centers. They provide physical mockups of ISS components, partial-gravity (see Figure 2.12, middle) and virtual-reality simulators for preparation and conduction of crew training, as well as neutral buoyancy laboratories (see Figure 2.12, right), huge water tanks, in which astronauts will be trained for space walks to perform extravehicular activities (EVA). The NASA also provides NEEMO (NASA Extreme Environment Mission Operations) missions to the U.S. underwater research laboratory Aquarius located at the coast of Florida. It is used by the NASA to simulate everyday life and work on a space station, and to train EVAs. In order to familiarize astronauts with hypergravity to which they are exposed during launch and reentry into Earth's atmosphere, they will be trained on human centrifuges providing high levels of acceleration. Currently NASA facilitates the highest *g*-level centrifuge at the Ames Research Center in California providing up to 20 *g*, while Roscosmos's human centrifuge, hosted at the GCTC, offers *g*-levels up to 18 *g*.

Figure 2.12: Facilities to train astronauts. *(left)* The training hall for Columbus, ATV and payload training at the ESA European Astronaut Center (EAC) in Cologne (Germany). [Photo © ESA–S. Corvaja, 2015] *(middle)* The EAC's neutral buoyancy hall for EVA spacewalk training, equipped with a huge water tank with a depth of 10 meters. [Photo © ESA–S. Corvaja, 2015] *(right)* NASA astronaut Reid Wiseman participates in a spacewalk training session in the *Partial Gravity Simulator*[18] at NASA's Johnson Space Center. [Photo © NASA, 2013]

Candidates intended to be astronauts are commonly selected by the ISS partners' space agencies that will be also astronauts' home base during their trainings and flight missions. Once candidates are selected, they will complete approximately 3 1/2 years of education before becoming eligible for a flight mission. This education covers miscellaneous trainings and practices ranging over three sequential training phases: the Basic Training, the Mission Assignment Training and the Mission Specific Training (Salmen et al., 2011). During the Basic Training, which normally takes place at astronauts' home base, the candidates are schooled in ISS on-board systems, shuttle systems, ground and launch systems, which will be complemented by various trainings, including aircraft operations, robotics, rendezvous and docking maneuvers, scuba diving, as well as training for land and water survival. Candidates will also be schooled in basics related to Earth and space sciences, engineering and meteorology. The following training phase for Mission Assignment will be locally conducted by each ISS partner to familiarize the candidates with the specific hardware under its responsibility. The ESA, for example, provides trainings focused on Columbus' systems and until 2015 also on ATV. During this time candidates will also be familiarized with the scientific payload racks, like the EDR, EPM, FSL, MARES and the Biolab payload. In the third training phase, the Mission Specific Training, candidates are individually trained in different levels of payload operations.

[18] "The Partial Gravity Simulator, also known as POGO, consists of servos, air bearings and gimbals to provide accurate simulations of reduced gravity. It is used for astronaut training and for evaluating their ability to perform tasks in simulated partial and microgravity." (NASA, URL: http://dx14.jsc.nasa.gov/htmls/pogo.htm, last visit: 28.09.2016)

The payload training is organized in a complex process and will be accompanied by experts and payload instructors. To train the candidates efficiently, candidates are assigned to their levels of qualification that can range from a user level, an operator level, to the level of a mission specialist (Aguzzi et al., 2008; Salmen et al., 2011). Thereby a qualification level defines the scope of the payload training and depends on candidates' mission specific onboard tasks and experiments that have been assigned to them. Apart from the individual training content, the three qualification levels share the goal to achieve the mission objectives and to maintain the safety of crew and space station. The user level training is focused on the candidates' capabilities to be aware of payloads' purposes, components and scopes of planned associated experiments. Candidates should also be able to list potential hazards and the relevant payload components to ensure its safety. This training takes place in front of the corresponding training model (see Figure 2.13, left) while the payload instructor provides this lesson by means of slideshow presentations. The training for the operator level intends to extend candidates' payload know-how to perform nominal and maintenance task that should be applied during their mission stay. This time the candidates are trained to understand and handle ODF procedures (see section 2.3.4). They also get knowledge about the specific hardware that is used for experimental operations. The training workflow is similar to the user level training by showing slideshow presentations in front of the training model. Finally a candidate can be trained on the mission specialist level becoming an expert for operating a certain payload. Such candidates are able to perform all nominal, maintenance, repair and published off-nominal tasks and will be trained for mission specific experimental tasks (see Figure 2.13, right). In case of off-nominal events that are not defined by procedures, the payload specialist should have the expertise to handle such situations in a way that the payload system is transferred to a safe state. For recovery purposes the ground team has to develop countermeasures that require the specification of new tasks described by ODF procedures that were not trained on the ground (Salmen et al., 2011). Last but not least it should be mentioned that all practice lessons during the current payload training take place on Earth under normogravity conditions that seriously limits the candidates' experiences (Aguzzi et al, 2008).

Figure 2.13: Payload training for the Biolab facility in the training hall of the European Astronaut Center in Cologne, Germany. *(left)* The Biolab training model. *(middle)* ESA astronaut Alexander Gerst is trained for the "Triplelux-B"[19] experiment by an ESA payload instructor [Photo © DLR, 2014]. *(right)* Training for two JAXA astronauts (Koichi Wakata, Soichi Nogushi) are receiving experiment training for the Biolab glovebox at EAC Cologne. [Photo © ESA/DLR–S. Schneider, 2009]

[19] The Triplelux-B experiment aimed to understand mechanisms of vertebrate and invertebrate cells at cellular level, which are influenced by various phenomena under spaceflight conditions, like the impairment of immune functions or the enhancement of responses to radiation. (ESA Erasmus Experiment Archive, URL: http://eea.spaceflight.esa.int /portal/exp/?id=9175, last visit: 28.09.2016)

2.4 Changes of Human Sensorimotor System in Weightlessness

Besides deep underwater or underground habitats, habitats in high mountain elevations, desert sites, or polar stations, the space environment is defined by physical stressors specifying the most extreme conditions in which a human have had to live and work (Bonin, 2005; Kanas & Manzey, 2008). In space, astronauts are faced with the most unfavorable conditions for a human that are, amongst others, characterized by the exposure to extreme temperature changes leading to thermal stress (Kenefick et al., 2008), the exposure to solar and galactic cosmic radiations (Atwell, 1990; Bagshaw & Cucinotta, 2008; Shiga, 2009), the absence of oxygen and atmospheric pressure (Katuntsev, 1998; Karemaker & Berecki-Gisolf, 2009) leading to decompressive stress (Stepanek & Webb, 2008), as well as an altered level of gravitational force affecting human's physiological system implying changes in cognitive and psychomotor functions (Morphew et al., 2001; Jones, 2010). Without gravity as reference frame for spatial orientation and motor coordination, astronauts experience a condition of extreme sensory discordance that they need to adapt to. Hence, working in space is hampered by microgravity degrading humans' psychomotor skills that restrict astronauts' capability to perform scientific and operational tasks. Kanas and Manzey (2008) have pointed to the role of microgravity on human performance and stated that "[..] degradations of human performance in space can become a serious safety issue during long-duration space missions" and "[...] will be even more important in the future when astronaut tasks and work devices will become more complex" (p. 6). As the MIR-Progress collision in 1997 has already shown, the impairment of human performance in space can have considerable consequences. Thereby the unmanned cargo spacecraft Progress 234 collided with the MIR during a manual docking test causing damages to a solar array and one of the space station modules. As reported by Ellis (2000) the analysis of the accident cause revealed that Human Factors by decreased performance during piloting have significantly contributed to this accident.

To reduce such risks it is important to understand the effect of microgravity on cognitive and psychomotor processes that are involved in human performance abilities (Kanas & Manzey, 2008). Beside of cardiovascular changes (Levine et al., 2003), like *orthostatic intolerance*[20] (Tuday et al., 2007) or the risk of cardiac dysrhythmias (Grenon et al., 2005), and alterations of the musculoskeletal system causing muscle atrophy (Fitts et al., 2001) and bone loss (Holick, 2000), the absence of gravity causes the most pronounced effect on human signal processing in the vestibular and sensorimotor system. Thus working in space is challenged by spatial disorientation and disturbed visuomotor coordination, both resulting from the integration of multiple sensory inputs perceived by the vestibular, visual, and the proprioceptive system, whose sensations are simultaneously interpreted by the central nervous system (CNS) and processed by the brain as position and motion data to coordinate and organize orientation and motor functions.

2.4.1 The Human Vestibular System and Reflexes

For all living beings on Earth, the gravity force constitutes the central point of reference for spatial perception and sensorimotor coordination. Therefore an eligible sensor system is required to be able to move within this gravitational field. Humans' ability to maintain the balance or postural

[20] Orthostatic intolerance is the inability to maintain the upright posture upon astronauts' return to Earth's gravity caused by a fluid shift to regulate the blood pressure. "Removal of all hydrostatic gradients when entering microgravity of spaceflight produces a large head-ward fluid shift; this has been commonly observed and is reflected in measurements of decreased leg volume [...]" (Jennings et al., 2008, p. 518)

equilibrium under Earth condition is based on the vestibular system that is optimized to sense the upright orientation of the body and to determine its position and direction in three-dimensional space (Peters, 1969; Kanas & Manzey, 2008).

From an anatomical point of view, the vestibular system (see Figure 2.14, left) is situated within the human inner ear as non-auditory part next to the cochlea as auditory part. The vestibular apparatus, as measuring device for angular and linear acceleration (see Figure 2.16, left), consists of three semicircular canals and two otolith organs – the utricle and the saccule. The semicircular canals (horizontal, superior, posterior) are sensitive to angular accelerations of the head that are characterized by simultaneous changes of velocity and direction. These canals are filled with a liquid (endolymph) that deflects the cupula, the gelatinous part of the ampullae, which is extending each canal and equipped with hair cells. As soon as the head is angular accelerated the liquid starts to flow through the canals and shifts the cupula that in turn controls the orientation of the ampullae's hair cells stimulating the neurons of the vestibular nerve transmitting this information to the CNS. The semicircular canals are arranged in an orthogonal manner and oriented along human's rotation planes (see Figure 2.14, right). In contrast, the otolith organs sense head's tilts and translations by linear accelerations describing changes in velocity without changes of direction, and are affected by the Earth's gravitational force. Thereby the utricle is more sensitive to acceleration changes in the horizontal plane, as experienced by driving, while the saccule detects changes in the vertical plane, as sensed by an elevator. Their functional principle is similar to those of the semicircular canals. The otolith organs consist of a macula and a gelatinous membrane, which contains calcium carbonate crystals (otoliths) increasing the membrane's density resulting in more inertia that enables the assessment of linear accelerations. The flow direction of the otoliths is downwards oriented along the gravity gradient. The macula provides the sensory cells and is connected with the vestibular part of vestibulocochlear nerve. The extensions of the macula's sensory cells project into the otolithic membrane that bends the hair cells by tilting or translating the head (see Figure 2.15, left). When a human stands in normal posture under normogravity conditions, without translating or tilting the head, then the hair cells of the otolith organs are not bent, whereas during translating or tilting of the head, the otoliths are shifted with reference to gravity displacing the otolithic membrane that bends the hair cells in the corresponding direction (Clément, 2011). The receptors of the macula pass the appropriate sensory information to the vestibulocochlear nerve which transmits to the CNS.

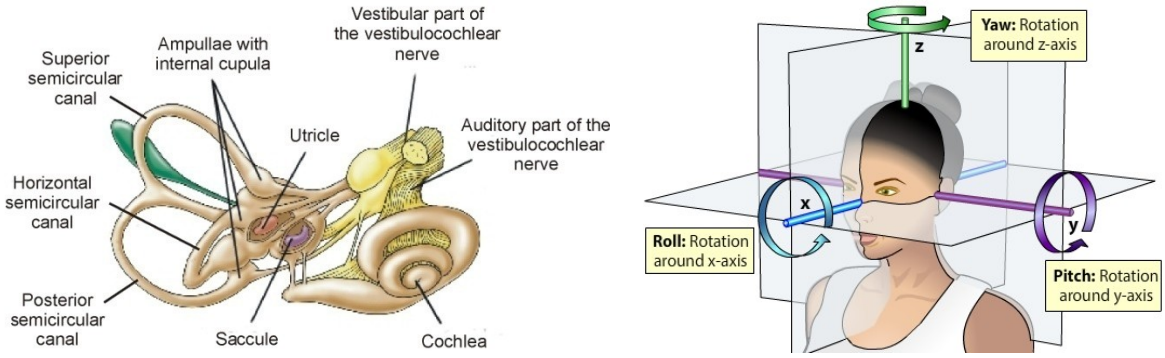


Figure 2.14: The Human vestibular system (from Watson & Breedlove, 2012). *(left)* The vestibular apparatus located within human inner ear. *(right)* Human body coordinates and the resulted rotation planes that correspond with the orientation of the semicircular canals.

To control sensorimotor actions during a task performance, the vestibular system fulfils two important functions, namely the stabilization of the eye position and therefore also the field of vision, as well as the stabilization of the head and body to maintain the posture, both during rest and motion. Thereby the gravity is used as relevant reference under Earth conditions, where we are usually not aware of these functions, because they appear unconsciously at the reflex level. The vestibular system stimulates two major reflexes (vestibulo-spinal, vestibulo-ocular) that are differently responded to by the human musculoskeletal system to compensate movements. Vestibulo-ocular reflexes are related to our visual system, acting on the eye muscles and are responsible to stabilize our gaze on stationary and moving visual targets by compensating head movements. Therefore such reflexes induce highly frequented rotations of the eyes that equally and oppositely map the direction of the preceded head movements (van der Steen, 2009). While during slow motions of the head or of surroundings relative to the head, the visual system stabilizes the image by optokinetic reflexes, like the *optokinetic nystagmus*[21], an increase of the motion's speed or during tilting the head about the horizontal, vertical, or torsional[22] axis, causes the vestibular-ocular reflex to be involved in maintaining clear vision (Parmet & Ercoline, 2008). For example, the *ocular countertorsion* reflex (see Figure 2.15, *right*) is stimulated by otolith-mediated linear accelerations during lateral tilting of the head, whereby the eyes are repositioned along the torsional axis. It is supposed that this reflex responses to direction changes of the gravity force (Parmet & Ercoline, 2008).

Vestibulo-spinal reflexes on the other hand are triggered in the spinal cord and contribute to stabilizing our upright posture during rest and motion (Manzoni, 2009). Because the body's center of mass is above our hip, causing an unstable balance, the vestibulo-spinal reflexes continuously stimulate fast and corrective movements so that we do not fall (Behrends et al., 2009). These compensating movements are conducted by the musculature of our neck, trunk and limbs that is regulated by the spinal reflexes of the propriozeptive[23] system. More specifically, such reflexes are responsible for the regulation of the body tonus responsible for passive and active muscle tension, whereby the passive stretch is the essential basis for resting in upright posture. Via the spinal cord the vestibular organs are directly connected to motor neurons, which trigger the necessary muscular stretch reflexes. Such reflexes enable us not to fall forward because of stimulated vestibular organs, while nodding our head, or trigger reflexively a step forward when stumbling, or moving the arms in the direction of a forward fall.

---

[21] The optokinetic nystagmus is a natural, visuo-mechanical reflex causing rapid involuntary eye movements that can be experienced, for example, while viewing the passing landscape through a window in the train.

[22] Rotation about the torsonial axis means a rotation about the axis of the line of sight, also referred to as visual axis.

[23] Proprioception is the self-perception of our own body and contributes to the sensation of the position and movement of our trunk and limbs relative to each other. It is also referred to as the kinesthetic sense (Proske & Gandevia, 2009) and contrary to our other senses we are usually not aware of it during our daily activities that are predominantly concerned with the generation of movements that in turn causes changes in muscles, joints and tendons. Thereby, the proprioceptive information is conveying from a variety of mechanoreceptors, which are specialized sensory receptors responding to mechanical events that are transduced into neural signals linked in the spinal cord and transmitted to the motor cortex (Riemann & Scott, 2003). Such mechanoreceptors provide information on the actual body position, motion and applied force, and are primarily found in muscles (*muscle spindles*), tendons (*Golgi tendon organs*) and joints. Thereby the Golgi tendon organ provides feedback on changes in the muscle tension to sense the force that a muscle exerts, whereas the muscle spindles detect changes related to the muscle length. The two primary spinal reflexes contributing to proprioception are the *stretch reflex* (regulates muscle length) and the *Golgi tendon reflex* (regulates tension), also called the inverse stretch reflex. Proprioception is complemented with the tactile sensation provided by mechanoreceptors in the skin, whereby their sensory neurons are directly passed to the cortex.

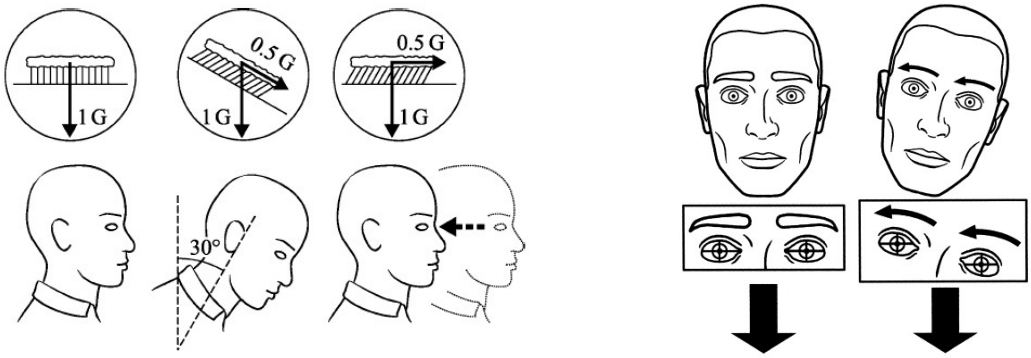


Figure 2.15: Vestibular-mediated mechanisms. *(left)* Tilt and linear acceleration of the head in normogravity are sensed by the otolith organs causing a bend of their hair cells that triggers a neuronal activity of the brain (from Clément, 2011, p. 100). *(right)* Ocular countertorsion, an otolith-mediated vestibulo-ocular reflex stimulated by lateral tilting the head, whereby the eyes rotate to the opposite direction. (from Parmet & Ercoline, 2008)

2.4.2 Vestibular Disorder

In orbital spaceflight, where the gravity is lost as a reference for orientation, the human's vestibular system is in a disorder that drastically impacts human sensorimotor performance by disturbed control of eye and body motions. While the functioning of the semicircular canals is not affected by spaceflight, the vestibular disorder is caused by a misinterpretation of the otolith signals that no longer stimulates the vestibular sense of the vertical, which is required for linear accelerations, like during tilting the head (see Figure 2.16, right). In neuroscience it is hypothesized that under microgravity the CNS reinterprets a tilt of the head as linear motion or translation, which is mentioned as the *Otolith-Tilt-Translation Reinterpretation* (Parker et al., 1985; Young et al., 1984; Clément, 2007). This reinterpretation of vestibular signals has substantial implications that astronauts need to compensate in weightlessness and re-adapt to normal gravity conditions after they return to Earth (Young et al., 1984; Young et al., 1986; Paloski et al., 1998). As the vestibular system is affected by the absence of gravity, the stimulated vestibular reflexes are also subjected to altered information processing responsible for reduced functionality. This leads to vestibulo-ocular adaptations causing alterations of eye movements (Clément, 1998, Clarke et al., 2000) and disturbances of gaze stability by reduced capability to orient towards fix targets and to track moving targets (André-Deshays et al. 1993). The visual system can also be affected by short-term alterations of the optokinetic nystagmus (Clément et al., 1986), or by visual illusions, like the *elevator illusion* (Cohen 1973, Lackner & DiZio, 1996), where the position of visual targets are perceived to low, causing a continuous undershooting, which in turn can lead to faulty task performance. The vestibulo-spinal reflexes and they collaborating proprioceptive spinal reflexes for postural control are also affected by the lack of otolith-mediated information. Thereby the activity of the muscle spindles can be affected by otolith-based gravity changes, causing a sensitivity change of the muscle spindles via the vestibulo-spinal tract in CNS, which can lead to an overestimation or underestimation of the muscle length. Studies in microgravity investigating the *H-reflex*[24] or vibration reflexes, like the *tonic vibration reflex*[25] (TVR), indicate a decreased activity of the muscle spindles by a low H-reflex

[24] "Originally described by Paul Hoffmann in 1910, [...], the Hoffmann reflex (H-reflex) is an electrically induced reflex analogous to the mechanically induced spinal stretch reflex. The primary difference between the H-reflex and the spinal stretch reflex is that the H-reflex bypasses the muscle spindle and, therefore, is a valuable tool in assessing modulation of monosynaptic reflex activity in the spinal cord." (Palmieri et al., 2004)

[25] The tonic vibration reflex is a muscle contraction caused by a high-frequency (30 Hz - 100 Hz) vibration stimulating the muscle spindles besides of the receptors of the skin and tendons (Latash, 2007).

(Reschke et al., 1984) and a reduced TVR (Lackner & DiZio, 1992; Lackner et al., 1992). This can lead to proprioceptive changes (Money & Cheung, 1991) resulting in decreased postural responses (Roll et al., 1992), decreased awareness of the direction and magnitude of motion (Reschke et al., 1984) or in illusion of self-motion, for example, during deep knee bends (Lackner & Graybiel, 1981). The reduction of the vestibulo-spinal reflexes in weightlessness also causes problems with the maintenance of the posture and coordination of walk, upon astronauts' return to the Earth (Clément & Reschke, 1996). The importance of the vestibular functions and their extent on vision and especially on proprioception is typically gaining the astronauts' awareness once they reach weightlessness, in which the gravity is lost as orientation reference. Besides the initial risk of space motion sickness (Oman et al., 1990; Reschke et al., 1994; Lackner and DiZio, 2006), the vestibular adaption to weightlessness affects the astronauts' well-being and degrades their sensorimotor performance induced by spatial disorientation and disturbances in motor coordination as a result of conflicting sensations of vestibular, visual and proprioceptive signals. Thereby the consideration of visuomotor coordination is meaningful to solve operational tasks at complex facilities, like the performance of demanding ISPR tasks.

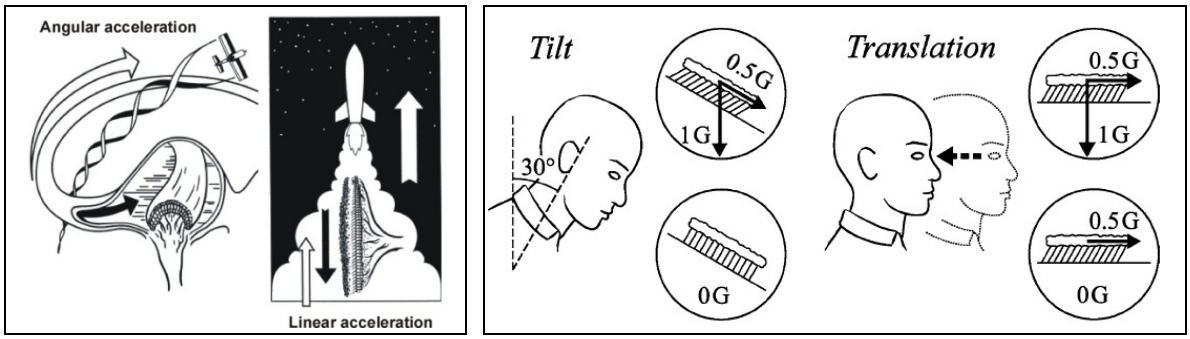


Figure 2.16: Tilt and linear acceleration of the head sensed by the otolith organs: *(left)* under normogravity (from Clément, 2011, p. 100), and *(right)* under microgravity (from Clément, 2011, p. 101).

2.4.3 Spatial Disorientation

Spatial orientation generally specifies humans' ability to continuously reorient themselves at rest (sitting, standing) and motion (walking, running) relating to the external physical space. Thereby the body's position and orientation results from spatial coordinate transformations between body-related intrinsic coordinates (see Figure 2.14, right) and extrinsic coordinates of objects coded in the external frame of reference (Harm et al., 2014). Under normogravity conditions on Earth, this spatial mapping is restricted by gravity that serves as a reference for the contributed multisensory interactions between the vestibular, visual and proprioceptive system. While the vestibular and proprioceptive sensor information are mapped to internal coordinates with respect to the head and the body, the sensor information of the eyes are based on visual cues in the physical surrounding (e.g., wall, ceiling, floor, trees) whereby distal objects (objects not in contact with the body) are coded in extrinsic world-coordinates that needs to be converted into stable body-centered intrinsic coordinates (Lipshits et al., 2005; Thompson & Henriques, 2011). Thereby visual cues affect the perception of spatial orientation in a substantial manner that can interfere the vestibular and proprioceptive inputs increasing the liability to static and dynamic reorientation illusions (Kanas & Manzey, 2008; Harm et al., 2014). Static illusions are related to the stationary position of the human and objects within the visual surrounding that, for example, can be perceived in an inversely furnished room, while dynamic illusions are related to the sense of self-motion or motions in the surrounding that, for example, can be

experienced by sitting in a standing train while observing a slower train on the adjacent track. Howard and colleagues (2000) have proposed "[...] that the dynamic illusion depends on a primitive mechanism of visual-vestibular interaction but that static reorientation illusions depend on learned visual cues to the vertical arising from the perceived tops and bottoms of familiar objects and spatial relationships between objects" (p. A87). By visual exclusion (e.g., eyes closed) the sensory information of the vestibular system and the proprioceptors should be sufficient to preserve a spatial map as mental picture of the environment and its spatial orientation.

However, the microgravity condition during spaceflight causes the absence of the gravitational reference frame by the lack of the otolith-mediated vertical that induces conflicted multisensory interactions between vision, proprioception and the vestibular apparatus needed for subjective sensation of head and body orientation coded in the intrinsic frame of reference. Thus visual and tactile cues from the surrounding will be more relevant for the sensation of spatial orientation, especially the influence of visual information that can partially or even totally overwrite the otolith or participating proprioceptive signals (Howard & Hu, 2001). Hence the maintenance of spatial orientation during work in spaceflight must rely more on visual and tactile sensations (Young et al. 1992, Lackner et al., 2000) that in turn significantly stimulate the occurrence of spatial disorientation (Young et al., 1984, 1993; Lackner, 1993; Glasauer & Mittelstaedt, 1997, 1998; Lackner & DiZio, 2000; Oman; 2007; Harm et al., 2014). Microgravity studies have shown that by visual exclusion the sense of spatial orientation during free floating is detached from the physical surrounding causing a lack in generating the associated spatial map, but can be restored, for example, by the sensation of tactile feedback from familiar parts of the surrounding (Lackner & Graybiel, 1983). Taking into account the visual sense, various static and dynamic orientation illusions (Young & Shelhamer, 1990; Kornilova, 1997; Oman et al., 2003; Harm et al., 2014) have proven a visual dominance, which can trigger the perception of spatial disorientation. The mostly reported static illusion of self-orientation in microgravity is the *inversion illusion* (Graybiel & Kellogg, 1967) that is caused by responses of the otoliths to the absence of gravity, whereby the gravitational force vector may be decomposed into head-fixed components resulting in an internal force-independent orientation vector (Mittelstaedt, 1989). The inversion illusion induces the sense of a sudden turnaround, comparable with hanging upside down, that is reasoned with our natural Earth-based tendency to sense the related "down" direction, whereby we unconsciously assume that the surface seeing under the own feet is the ground, or that the body orientation of somebody else floating upside down indicates the correct direction. The same applies to a view from space towards the Earth that provokes an inverse reorientation illusion if the Earth is located above the head. In general, it might be reasonably assumed that static orientation illusions, especially the inversion illusion, are responsible for the development of space motion sickness (Oman, 1988). Besides static orientation illusions, dynamic illusions including the sensation of moving surrounding objects and false perceptions of self-motion are related to head movements. The perception of motions in the surrounding can be caused by the vestibulo-occular reflex resulting in malfunctioning of the gaze stabilization (André-Deshays et al. 1993; Kanas & Manzey, 2008). This, for example, can provoke horizontal or vertical shifts of stationary objects that are observed while moving the head (e.g., displacement of instruments of an observed control panel). In contrast, self-motion illusions are related to the apparent linear and circular movement of the body (Kornilova, 1997; Oman et al., 2003). Such illusions are most likely caused by the more pronounced sense of vision on spatial orientation that cannot be vestibular corrected through the gravitational reference. The experimental assessment of self-motion illusions is mostly been carried out by means of virtual environment projection systems, like the hemispherical display used for the "rotating dome" experiment (see Figure 2.17) conducted by Young and colleagues (1992). Thereby the subject's head

is placed inside the display and views a moving visual dot pattern that rotates around the head's roll axis with a constant angular velocity of 30, 40 or 60 deg/s. While viewing the visual stimuli the feet are either free floating or fixed to the floor by bungee cords that additionally enable the investigation of multisensory interactions between vestibular, visual and tactile signals. In case of free floating the subject experienced a self-displacement in the opposite direction of the dome rotation (mostly less or equal than 90°). This and multiple other studies (Young et al., 1986, 1996; Young and Shelhamer 1990) have shown that the impact of such illusion effects were more pronounced in microgravity than on Earth and that tactile stimulations on the feet can reduce the illusion of self-motion indicating that tactile sensations can replace the missing otolith information.

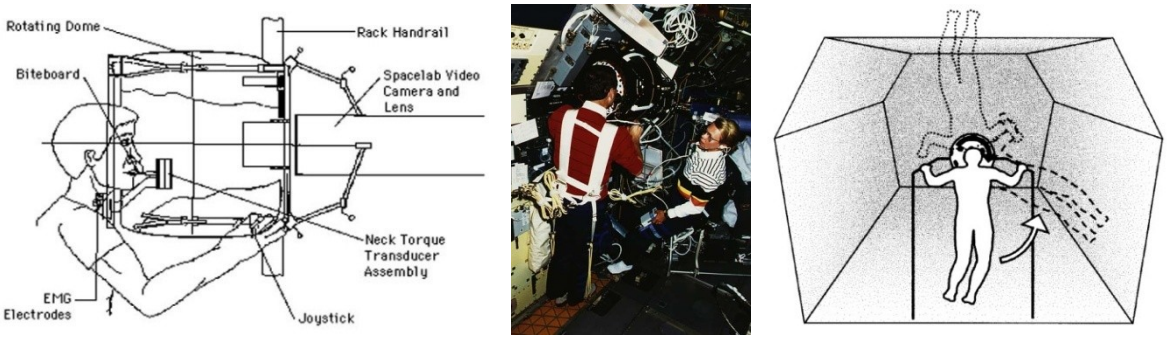


Figure 2.17: The "rotating dome" experiment. *(left)* Schematic of the "rotating dome" assembly (from NASA JSC, 2015). *(middle)* NASA astronaut Drew Gaffney performs the "rotating dome" test on a Spacelab Life Science payload (from NASA JSC, 2015). *(right)* Illustration of visually induced illusory self-rotation in roll, similar to the "rotating dome" experiment (from Lackner & DiZio, 2000).

The various effects of weightlessness on spatial orientation are individually experienced by astronauts with respect to temporal dimensions and natural dispositions (Kanas & Manzey, 2008; Harm et al., 2014). While most of the orientation illusions tend to appear directly after the transition to microgravity and only require a few minutes or hours of physiological and cognitive adaptation, some of them can last for several days or return regularly again over the entire mission period, which could be explained by instable adaptation (Kornilova, 1997). The occurrence of orientation illusions is also depending on astronauts' natural dispositions or tendencies in their default spatial coordinate system that causes different sensations related to the absence of the vertical, in particular under the exclusion of visual and tactile stimuli. Some astronauts can maintain their spatial orientation by a robust egocentric reference frame and thus are still capable to sense the "up" direction in respect of their head (Gurfinkel et al., 1993; Glasauer & Mittelstaedt, 1998; Harm et al., 1998).

2.4.4 Visuomotor Limitations

Whether simple daily activities or the performance of complex task operations: almost any kind of interaction requires goal-directed movements that rely on our capability to control motor functions of our limbs coordinated by the visual system providing knowledge about the dynamic surroundings. Although thereby the afferent information from all sensory receptors are relevant, this section is focused on the visual system and its meaning for the motor commands, both primarily needed for visuomotor coordination, such as for upper limb functions, like hand-eye coordination needed to reach, grasp or point towards objects. The success of visuomotor coordination is dependent on intact pathways of our vision and limb proprioception or kinesthesia, which is actively calibrated to the gravitational frame of reference on Earth. Thereby the proportion of the proprioceptive information

takes on a vital role in the control of goal-directed movements (Park et al., 1999). So, for example, we are able to accurately point towards a remembered target while closing the eyes, even after moving the head. This is possible because the arm is controlled by the visual memory and monitored by proprioceptive information, both stabilized by vestibular reflexes (Whiteside, 1961). Many studies have used this effect of visual exclusion to investigate the effect of microgravity on limb proprioception assessed by goal-directed pointing, usually assessed in upright posture along the gravitational vertical (see Figure 2.18). Thereby the proprioceptive controlled pointing movements was hindered either by not seeing the arm (Whiteside, 1961; Bock et al., 1992; Mechtcheriakov et al., 2002) or the hand (Bock et al., 2001), or being completely without visual feedback just while pointing towards remembered targets (Berger, et al., 1993; Smetanin & Popov, 1997; Watt, 1997). All these studies provided evidences that the absence of gravity causes a significant slowdown of the arm movement. Some of these studies have also investigated the pointing accuracy but yielded inconsistent results showing either that the accuracy remained unaffected (Mechtcheriakov et al., 2002) or, and that mostly so, showing an inaccurate pointing in the vertical under microgravity revealing either in overshoots (Bock et al., 1992), or in lower pointing (undershoots) inducing the illusions of targets moving downwards (Whiteside, 1961; Smetanin & Popov, 1997). Smetanin & Popov (1997) additionally investigated goal-directed pointing towards remembered target locations in lying positions (supine and prone) and showed that arm movements conducted in the lying postures cause less pointing error than in the upright posture and that, relative to the subject's body, the targets were more undershot in the upright and overshot in the lying postures.

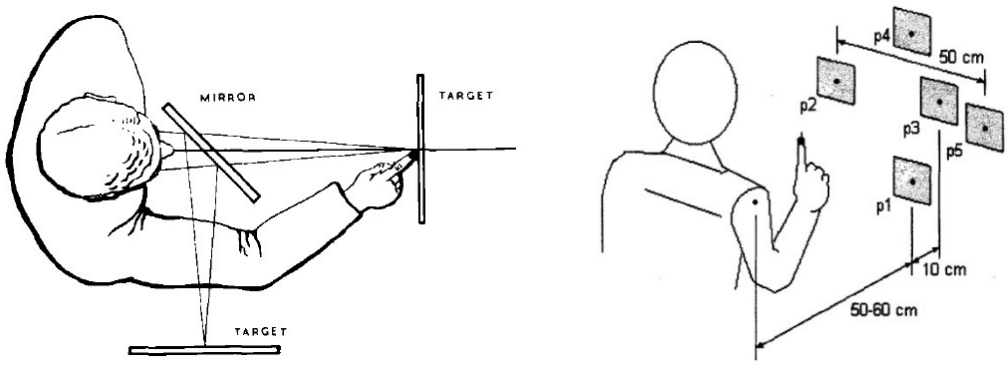


Figure 2.18: Setups of microgravity experimentations to study the motor control of arm movements during goal-directed pointing. *(left)* Pointing without seeing the arm causes the illusion of target moving downwards (from Whiteside, 1961). *(right)* The arm motor program during pointing in the lying posture is not modified compared to pointing in upright posture (from Smetanin & Popov, 1997).

Besides investigating the arm motor control by pointing tasks toward external targets, Watt (1997) also compared pointing at remembered targets with touching at remembered body parts to differentiate the errors differing in state of not knowing either where the arm was pointing or where the target is located. He has suggested that problems during pointing at memorized targets can be explained by a lacking knowledge of the target position and not of the limb position. Here the question arises whether this conclusion is permissible or sufficient. While pointing to memorized targets constitutes a task conducted in extrinsic world coordinates that needed to be adapted by the proprioceptive system to control limb positions, the task of touching body parts however is detached from the surrounding and conducted in the body-related frame of reference. In my opinion these results are showing not more than that the proprioceptive arm motor control is less affected in the body-related frame of reference as in the external world frame. Similar to touching of body parts, Fisk and colleagues (1993) investigated the arm movement control by moving the forearm between

remembered elbow joint angles and compared the performance between horizontal and vertical movements and between rapid and slow movements to evaluate the role of proprioceptive feedback. While horizontal and vertical movements did not differ significantly in performance, this study revealed that the amplitude of movements was depending on the speed. Thereby the amplitude of rapid movements was not affected, while the performance during slow movements showed an effect of microgravity in their amplitudes.

While all above mentioned studies were performed by visual exclusions to reinforce the sense of proprioception, the consideration of the visual guidance is more closely related to real world tasks, even if it causes a visual dominance overriding the otilith-mediated misinformation of the vestibular and proprioceptive system. There are only few studies investigating hand-eye coordination by visual control, which differed in the studied problem statement that includes studies investigating the degree of influence of visual control by comparing different visual conditions (Berger et al., 1993; Mechtcheriakov et al., 2002), the investigation of the associated head movements (Fogt et al., 2002), or just simply the evaluation of a visual guided tracking task conducted by the finger (Bock et al., 2003) and additionally by a stylus or joystick (Fowler et al., 2008). Just like for visual exclusion, these studies have also shown that microgravity causes a deteriorated pointing behavior by slowing down movements or increased pointing errors. The pointing performance was not affected by altering the visual control conditions (Mechtcheriakov et al., 2002), although visual controlled movements showed better adaptations (Berger, et al., 1993). Furthermore, Fogt and colleagues (2002) have evidenced that head movements contribute to the pointing accuracy during rapid, repetitive visual guided pointing tasks. Other studies have dealt with challenges related to the proprioceptive restoration of movement disorder by force and tactile induced stimuli. Therefore Bringoux and colleagues (2012) showed that gravity-like torque generated by increased sensorimotoric load (see Figure 2.19, left) enhances the arm position sense during goal-directed arm movements with eyes closed under microgravity conditions and suggested that the tuned motor planning and control of arm movements causes an optimal activation of the internal models with respect to the gravitational information. In contrast, Lackner and colleagues (2000) conducted a parabolic flight study investigating the effect of fingertip contacts with a stable surface (see Figure 2.19, right) on proprioceptive responses to decrease postural destabilization. This study revealed that vestibular disturbed proprioceptive and motor signals in leg muscles experienced under microgravity can be suppressed by tactile sensations via touching an external surface.

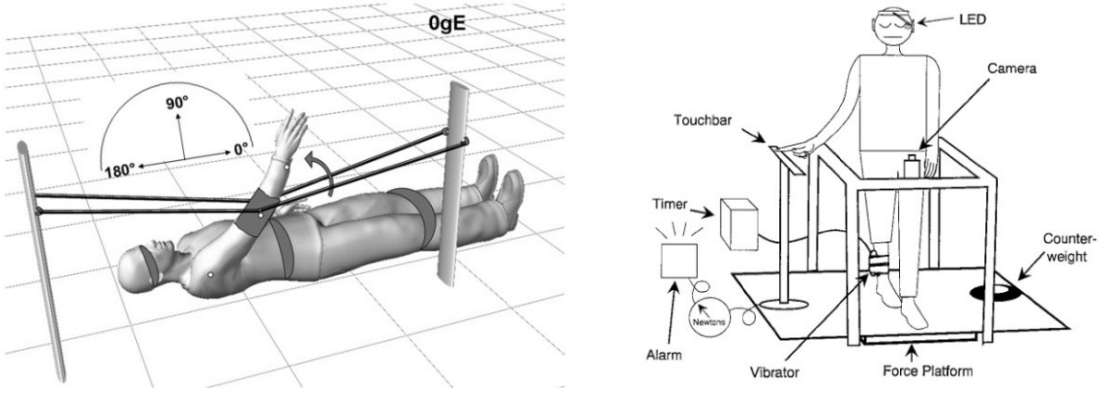


Figure 2.19: Setups of microgravity experimentations to study the effect of tactile and force sensations on limb proprioception. *(left)* Enhancement of arm position sense by gravity-like torque (from Bringoux et al., 2012). *(right)* Tactile cues by fingertip contact suppress proprioceptive destabilizing in leg muscles (from Lackner et al., 2000).

In summary, it can be stated that the visuomotor coordination, discussed by hand-eye coordination, is affected by the absence of gravity resulting in a degraded performance by slowing down movements and inaccurate pointing that can be attributed to otolith-mediated proprioceptive changes or an improper kinesthetic motor control. None of these studies have investigated the effect of altered gravity on visual guided pointing movements towards targets localized in the body-related frame of reference that could benefit from the absence of the coordinate transformation, but could also be disadvantaged by an affected limb proprioception. However, hand-eye coordination conducted in the intrinsic frame can imply a missing of tactile feedback by one-handed operation, or the usage of both hands, whereby one hand serves as pointing surface while the other hand is pointing towards it, resulting in a bimanual handling strategy. There are also no studies investigating the effect of altered gravity on goal-directed pointing without tactile sensation or studying hand-eye coordination during bimanual handling. Research on operation of interactive control interfaces in microgravity should consider these findings and open issues on hand-eye coordination.

Chapter 3

Applied Theories and Concepts for Studying Human Factors

"Proper design can make a difference in our quality of life."

(Donald A. Norman, *The Design of Everyday Things*,1988)

Because the research presented in this thesis claims to meet the user and the environmental needs, it implies a user-centered design, which requires the consideration of the human processing system by nature and the provided tools and concepts to study Human Factors. Therefore this subchapter explains associated aspects in a subdivided manner. On the one hand, this chapter defines the concept of usability and its design rules ranging from a general perspective, to more specific guidelines applicable to mobile AR. On the other hand, it also conveys knowledge about human perception and information processing, as well as related issues on instructional design and visual search, which are needed to explain the mentioned valuables in the context of this thesis. Finally it outlines and describes the applied methods and concepts used to analyze the problem space and the user's needs, as well as to evaluate usability, in the scope of this thesis, with a special focus on the assessment of workload. But first of all, it will be clarified what Human Factors means and what it is focused on.

3.1 What does Human Factors Mean?

When the design of a system is primarily focused on the technology and its features, and lacks of adequate usability testing, it will often not be accepted by people intended to use it. In contrast to such a technology-driven system, a user-driven design considers the human point of view from the very beginning and optimally results in an efficient system. Without adequate consideration of Human Factors, the quality of the interaction between a human and the systems is left to chance. Human factors, which is closely related to ergonomics, engineering psychology or usability engineering in HCI, is concerned with the relationship between human and the working environment considering the limits of human capabilities related to physical, perceptual and cognitive aspects that are impacting the design of technological systems. Human Factors was defined by Sanders and McCormick (1982):

"Human Factors focuses on human beings and their interaction with products, equipment, facilities, procedures, and environments used in work and everyday living. The emphasis is on human beings [...] and how the design of things influences people. Human Factors, then, seeks to change the things people use and the environments in which they use these things to better match the capabilities, limitations, and needs of people." (p. 4).

As an empirical science, Human Factors is predominantly based on collecting quantitative and qualitative information to gain knowledge about human responses and limitations while interacting with the system, and to finally bring this knowledge to bear on its design (Proctor & Van Zandt, 2008).

Within the context of computer systems, Human Factors claims to enhance their usability, whereby the meta-objectives can be simple state as: (Wickens et al., 2003, p. 2)

- enhance performance,
- increase safety, and
- increase user satisfaction.

Thereby, possible tradeoffs between these objectives need to be considered. For example, an increase in performance does not automatically imply an increase in safety, which can be affected by errors of the human operator. In order to comply with all objectives it is required to identify problems or deficits in the interaction between the user and the system. To efficiently expose such deficits and to clearly attribute the cause of possible disturbances, human physical and cognitive restrictions in the context of system requirements must be understood. While physical restrictions are linked to characteristics of the human body, cognition is concerned with our ability to process the information perceived by the environment in consideration of our knowledge. Thus, Human Factors in principle aims "[...] to find a system design that supports the user´s needs rather than making a system to which the users must adapt" (Wickens et al., 2003 p. 15). In the field of usability engineering, an user-centered design process requests: (Nielsen, 1993)

- an early focus on the user and tasks,
- empirical measurements, and
- iterative development and testing of prototypes.

This means that the users' needs, preferences and limitations are influencing the design process from the very beginning, continuously updating the specification requirements, whereby each update requests its own implementation and evaluation, resulting in a design circle that iterates until the users' needs are met. As shown in Figure 3.1 (left) the design lifecycle considering Human Factors is a closed-loop, whereby identified problems can be solved with different design approaches, which in turn can focus on the *equipment* (physical devices), the *task* (work flow), the *environment* (e.g., light, temperature, noise), the *selection* (human individual differences) and the *training* (preparing). Regardless of the used strategy, this design process enables an early and continuous consideration of Human Factors, which may contribute to fulfill the objectives of Human Factors mentioned above.

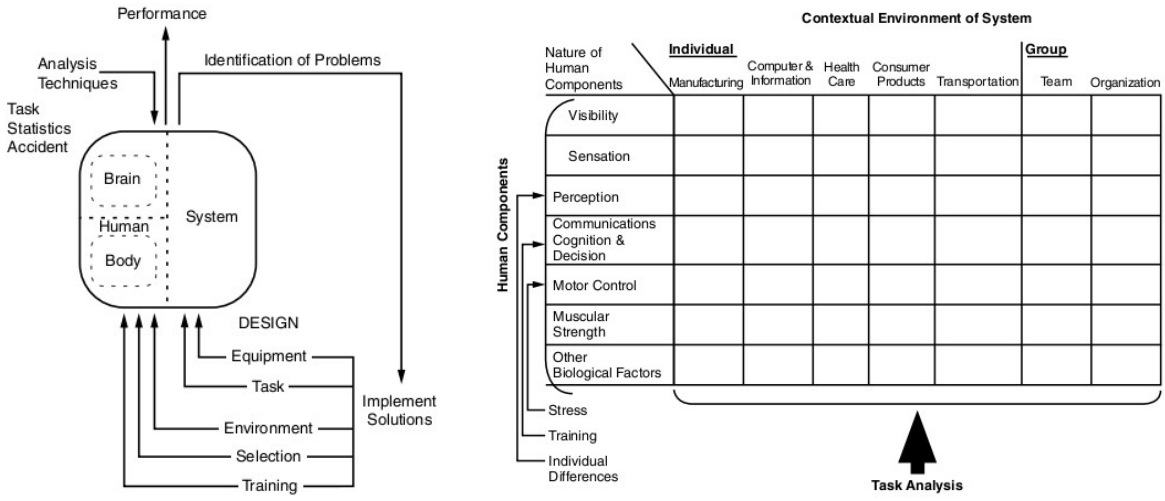


Figure 3.1: Characteristics of Human Factors engineering (from Wickens et al., 2003, p. 3, 5). *(left)* The design lifecycle shows how the human and the system are involved in the design process. *(right)* Matrix of Human Factors outlining the potential task environments suitable to study human components.

Whether for transportation, manufacturing or specifically for computer interfaces, Human Factors can be studied for every conceivable environment in which a human is needed to operate. Depending on the system intended to design or enhance, a subset of human components (see Figure 3.1, right) enables the specification of adequate questions, which need to be answered. Such questions could be, for example: *"Can it be sensed and adequately perceived"*, *"What communications and cognitive processes are involved in understanding the information and deciding to do with it?"*, or *"How are actions to be carry out, and what are the physical and muscular demands of those actions?"* (Wickens et al., 2003, p. 6). Besides selecting the most suitable design approach and asking the right questions, the kind of studying these human components is equally important and varies in the realism of the task environment, the level of control over the experiment and the level of the user's familiarity with the task domain. As deeper the understanding of the human components, as more generalizable and predictable the findings will be across classes of problems that may have common elements.

3.2 The Concept of Usability and Design Rules

The general goal of designing interactive products or systems is to improve and maximize their ease of use and learnability, also referred to as *usability*, which "[..] applies to all aspects of a system with which a human might interact, including installation and maintenance procedures" (Nielsen, 1993, p. 25). In Part 11 of the international standard ISO 9241, titled *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)*, *usability* is defined as "the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments". Thereby the terms effectiveness, efficiency and satisfaction are defined as follows: (Dix et al., 2003, p. 277)

- **Effectiveness:** The accuracy and completeness with which specified users can achieve specified goals in particular environments.
- **Efficiency:** The resources expended in relation to the accuracy and completeness of goals achieved.
- **Satisfaction:** The comfort and acceptability of the work system to its users and other people affected by its use.

In his book *"Usability Engineering"*, Jacob Nielsen (1993) has stated that "[...] usability is a narrow concern compared to the larger issue of system acceptability, which basically is the question of whether the system is good enough to satisfy all the needs and requirements of the users [...]" (p. 24). In this context, he introduced the category *usefulness* to analyze the system's practical acceptability. Thereby a system's usefulness considers two aspects: *utility* and *usability*. While utility is concerned with the intended functional capability of a system, usability determines "[...] how well users can use that functionality" (p. 25). Nielsen has defined usability of user interfaces (UIs) as multi-dimensional, which is associated with the following five *usability attributes*: (Nielsen, 1993, p. 26)

- **Learnability:** The system should be easy to learn so that the user can rapidly start getting some work done.
- **Efficiency:** The system should be efficient to use so that once the user has learned the system, a high level of productivity is possible.
- **Memorability:** The system should be easy to remember so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again.
- **Errors:** The system should have a low error rate so that users make few errors during the use of the system and so that if they do make errors, they can easily recover from them.

- **Satisfaction:** The system should be pleasant to use, so that users are subjectively satisfied when using it; they like it.

Design activities directed towards enhancing the usability are calling for a user-centered design process that can be supported by *design rules*, which can vary in their level of generality and authority (Dix et al., 2003, p. 259). While *generality* reflects the degree of the restriction in its application, the *authority* of a rule indicates how strong a rule must be followed ranging from informal recommendations to exact specifications. Depending on their level of generality and authority, design rules can be expressed in the form of abstract *principles* (high generality, low authority), *guidelines* (general in application, low authority) or specific *standards* (limited in application, high authority).

As already emphasized by Donald A. Norman, a user-centered design is needed to be assigned to everyday things (Norman, 1988). Their design should be based on the needs and interests of the user ensuring that "(1) the user figure out what to do and (2) the user can tell what is going on" (p. 188). Its transposition is supported by following fundamental *principles of design for understandability and usability*: (Norman, 1988, pp. 13-28, 188)

- **Provide a good conceptual model:** Make it easy to determine what actions are possible at any moments (make use of constraints).
- **Make things visible:** Visibility acts as a good reminder of what can be done and allows the control to specify how the action is performed.
- **The principle of feedback:** Make it easy to evaluate the current state of the system.
- **The principle of mapping:** Follow natural mappings between intentions and the required actions; between actions and the resulting effect; and between the information that is visible and the interpretation of the system state.

To guide the design process, Norman also suggested the *seven principles for transforming difficult tasks into simple ones*: 1. use both knowledge in the world and knowledge in the head, 2. simplify the structure of tasks, 3. make things visible: bridge the gulfs of execution and evaluation, 4. get the mappings right, 5. exploit the power of constraints, both natural and artificial, 6. design for error, and 7. when all fails, standardize. (Norman, 1988, pp. 188). Similar but more restricted to computer UIs, Ben Shneiderman has introduced strategies for effective human-computer interaction (1987). His *eight golden rules of interface design* are valid for most interactive systems: 1. strive for consistency, 2. cater to universal usability, 3. offer informative feedback, 4. design dialogs to yield closure, 5. prevent errors, 6. permit easy reversal of actions, 7. support internal locus of control, and 8. reduce short-term memory load (Shneiderman et al., 2009). Therewith, Shneiderman has laid the foundation for specifying subsequent principles and guidelines (e.g., Hix & Hartson, 1993; Stone et al., 2005). Such rules are primarily intended to be used during the design, but can also be applied to the systematic inspection of a UI to identify usability problems, like Nielsen's *usability heuristics*, which provide principles to set up common properties of usability interfaces (Nielsen 1993, p. 115ff).

The above mentioned rules are universal principles with a very high generality and low authority, and thus must be adapted for each environment (Shneiderman et al., 2009). If only considering 3D applications, which also should be applicable to VR and AR applications, the usability principles have to be interpreted and extended for 3D UIs and their interaction techniques. In their book *"3D User Interfaces:Theory and Practice"*, Bowman and colleagues (2004) have summarized general guidelines for designing and developing 3D UIs, as well as guidelines for user comfort. Furthermore they also suggest various guidelines assigned to each interaction category, such as selection and

manipulation, travel, wayfinding, system control and symbolic input. So for example, they suggest that the tradeoff between the technique design and the environment design has to be consider, that the pointing technique should be used for selection, or that an appropriate spatial reference frame should be used for system control UIs.

Although very rare, some guidelines specific for designing AR applications were already proposed. While Gabbard and Hix (2001; cited after Bai & Blackwell 2012) surveyed design guidelines for VR/AR usability, Dünser et al. (2007; cited after Bai & Blackwell, 2012) were more focused on the design of AR systems suggesting guidelines as: "[...] affordance, reducing cognitive overhead, low physical effort, learnability, user satisfaction, flexibility in use, responsiveness and feedback, and error tolerance" (Bai & Blackwell, 2012, p. 456), whereby affordance indicates inherent linking between the user interface element and its functional and physical properties (e.g., tangible user interfaces). More related to guidelines considering the perceptual processing in AR system designs, critical perceptual issues were identified and classified by Drascic and Milgram (1996), and revisited by Kruijff et al. (2010). Thereby the major perceptual issues in AR are related to: the *environment* (e.g., structure, colors, conditions), *capturing* (e.g., image resolution, lens issues, exposure, frame-rate), *augmentation* (e.g., registration errors, occlusion, interferences), *display devices* (e.g., stereoscopy, field of view, viewing angle offset) and the *user* (e.g., individual differences, depth cues, disparity planes). Kruijff et al. also summarized works suggesting approaches to counteract such problems. In contrast, Li and Duh (2013) excluded technical AR issues and were solely focused on cognitive issues in mobile AR to support a Human Factors design with high effectiveness whereby "[...] the user-centred perspective should be incorporated, and more understanding on the impact of characteristics of AR system on human activities is needed" (p. 109). They also stated that "[...] our knowledge about Human Factors in mobile AR is very limited, and therefore it is crucial to comprehensively identify the opportunities and challenges posed by mobile AR interaction on the cognitive process" (p. 110). They outlined three categories of cognitive AR issues including: *information presentation* (amount, representation, placement, view combination), *physical interaction* (navigation, direct manipulation, direct hand interaction, multimodal) and *shared experiences* (social context, bodily configuration, artifact manipulation, display space). In their conclusion they emphasized that the support of users' cognitive processing "[...] has become an important part in fulfilling the potential of AR technology [...]" (p. 129).

3.3 Related Issues on Human Cognition

To analyze the user's need in the process of designing new interfaces, it is important to understand the cognitive processes demanded by the task performance in question, like SSC-interfaced payload tasks. Therefore, this section provides general fundamentals of the human perception and information processing system, and their characteristic features. Because the effects of weightlessness on the human sensorimotor system was already discussed in section 2.4., the following subsequent sections are dealing more detailed with theories focused on the cognitive load triggered by instructional design issues and on the integration of objects' features during visual search.

3.3.1 Human Perception and Information Processing

The information processing determines our understanding and interaction with the world around us. Prior to responding to an event or information, it requires its perceptual encoding, central processing

and response selection. Thus, human information processing can be considered as a three-stage model (perception, cognition, responding), whereby each stage depends upon attentional resources, which can be allocated if required and are reflecting the mental effort, which, for example, an instructional task is needed (Wickens et al., 2003, p. 101). Initially, each type of sensation (e.g., visual, auditory) is registered in the temporally limited sensory memory and transferred to the next processing stage. As shown in Figure 3.2 (left), perception can be automatic, directly leading to a response selection (e.g., hot water, ambulance siren, red alarm light), or need to be manipulated in the working memory based upon prior knowledge and experiences stored in the long-term memory (e.g., reading, searching). Relevant issues need to be explained in the scope of this thesis, like the types of perception as well as the characteristics of the working and long-term memory, are described in the following.

Perceptual Processing Types: Human perception, like the visual perception, is processed by the brain in different ways, often simultaneously and concurrently, but each has its own implication for designing user interfaces. As shown in Figure 3.2 (right), perception is influenced by *bottom-up* and *top-down* processing. Thereby the bottom-up process is driven by sensory information from the physical world, which means that the perception is reduced to its raw stimulus features (e.g., color, shape, size). In case of visually perceived events, bottom-up processing can be triggered or maximized by a cued object, arousing a *salience* effect, which catches the viewer's attention by its highlighting. Contrary, top-down processing is driven by knowledge and experiences that are relying on retrieved information from the long-term memory and processed in the working memory. The amount needed for the effort of top-down processing depends on the level of expectations (anticipated meaning), whereby high expectations require sufficient experiences or practices. Thus, perception can benefit from bottom-up processing by the salience effect or from top-down processing maximized by high expectations. Besides these two general processing types, perception processing by *unitization* can be considered as a mixture of both, but "[...] is more rapid and automatic [...]" (Wickens et al., 2003, p. 104). On the one hand, unitization is feature-driven, like bottom-up, but on the other hand it is assumed that the viewer is familiar with sets of the perceived features (e.g., words in a familiar language). This means that a high level of expectations (top-down) is required, so that its

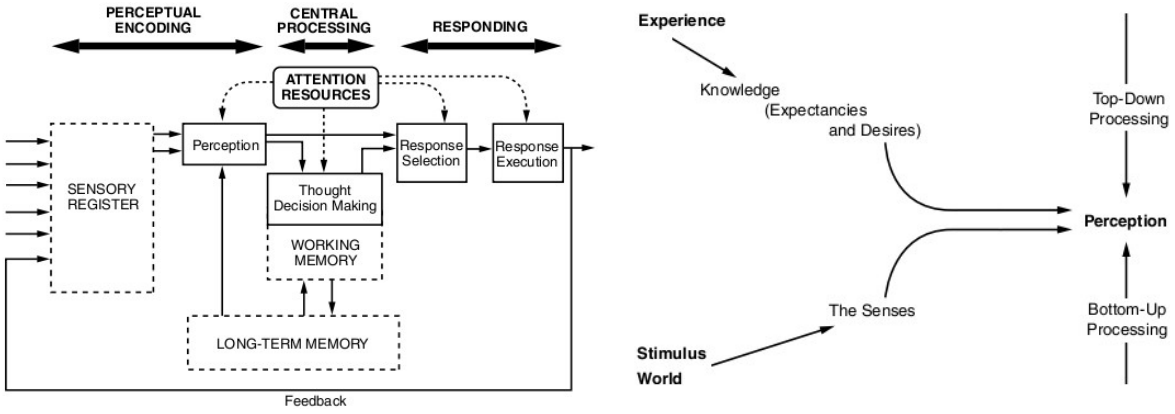


Figure 3.2: The human perception and information processing. *(left)* The information processing relies on three stages: perception, cognition and responding (from Wickens et al., 2003, p. 102). *(right)* Perceptual processing types: stimulus-driven bottom-up vs. knowledge-driven top-down (from Wickens et al., 2003, p. 54).

meaning can be directly accessed by the long-term memory as a whole unit and without effortful processing in the working memory. Ideally, perception should be relatively automatic, "[...] but becomes less so as bottom-up processing is degraded and top-down and unitization processes become less effective" (Wickens et al., 2003, p. 108). As longer the perceptual process takes, as fewer it is

perception but more comprehension, which is less automatic. To support perceptual processing, the following simple Human Factors guidelines have been proposed by Wickens et al. (2003, pp. 106-107):

- maximize bottom-up processing,
- maximize automaticity and unitization by using familiar perceptual representations, and
- maximize top-down processing when bottom-up processing may be poor.

Long-Term Memory: The long-term memory houses all knowledge, information and experiences, which an individual has previously learned or trained, and are stored in schemas related to a specific concept or topic. Such materials are in an inactive state until they are recalled by processing associated stimuli or information either directly by perception or by the working memory (schema automation and acquisition). Thus, learning or training is the process of storing information into the long-term memory by constructing or extending schemas, and can be facilitated by work examples or procedures, which in turns can be affected by their instructional design (see section 3.3.4). However, the difficult level of retrieving information from the long-term memory depends on the stored materials' *strength* and *associations* (Wickens et al., 2003). Thereby, the strength of a material is characterized by the frequency of its use (e.g., the frequency of using an instruction) and its recency indicated by the period of time between learning and recalling. For example, if a pilot learns an emergency procedure that is only required one month later, then its performance should not only rely upon the memory, but rather being supported by, for example, a checklist. Additionally, materials or items in the long-term memory can be linked or associated among each other, like a sound, symbol or name with their meanings or usages. Just like the strength of material, associations can be weak or lost over time and therefore need to be rehearsed or repeated in between. Thus, failures of retrieving information from the long-term memory are often caused by: "weak strength due to low frequency or recency", "weak or few associations with other information", and "interfering associations" (Wickens et al., 2003, p. 115).

Working Memory: The working memory, also referred to as short-term memory, can be considered as activated representation of the long-term memory, which has to be stimulated by attentional resources to bring it in the focus (Cowan, 1995). Various models have been proposed to explain how the working memory works, whereby the *multi-component model*[26] of Baddeley & Hitch (1974) is probably one of the most popular. During learning or training, the working memory is responsible to build meaningful associations between information, which are then stored in the long-term memory by means of chunks, which describes small units (e.g., words) that can be combined to larger units, such as sentences, paragraphs or stories, but also to instructions as used for payload operations. Although chunks are valuable to increase the capacity of the working memory, we are only capable to remember 7 ± 2 chunks (Miller, 1994). But more recent studies are suggesting an even lower capacity (Cowan, 2001). As for its capacity, the working memory is also limited in time, whereby unrehearsed information will begin to decay within 15-20 seconds (Goldstein, 2010, p.

[26] According to the *multi-component model* of Baddeley and Hitch (1974), the working memory consists of a *central executive*, a *visuospatial sketchpad* and a *phonologic loop*, and has been later extended by an *episodic buffer* (Baddeley, 2000). Thereby the visuospatial sketchpad is responsible for processing and temporary storing visual (e.g., color, shape) and spatial (e.g., objects' position) information in a separate manner, while the phonologic loop manipulates and temporally stores auditory or verbal information in a phonetic form, which is limited in time and thus information need to be kept active by *"rehearsal"* until it is no longer required. However, the episodic buffer is a limited multimodal storage system used to integrate units of visual, spatial, and verbal information in a chronological order, and it is assumed that this buffer has its own link to the long-term memory. All these processes are controlled and coordinated by the central executive, which is connected to the long-term memory and responsible to allocate and focus attention.

143). This not only depends on the process of periodic reactivation or rehearsal, it also depends on the number of chunks actually held in the working memory. Hence, to support a user-centered design, it is suggested to consider the limits of the working memory and to minimize its load (Wickens et al., 2003).

3.3.2 Attention and Workload

Whether by perceptual encoding, central processing in the working memory or by response selection and execution, all stages of human information processing claim the need of attentional resources (see Figure 3.2, left). Such resources are limited by nature and imply a substantial bottleneck in human information processing (Wickens & Hollands, 2000). According to the capacity model of attention proposed by Kahneman (1973), "[...] an activity can fail, either because there is altogether not enough capacity to meet its demands or because the allocation policy channels available capacity to other activities" (p. 9). Thereby the level of demand depends on the mental activities' level of difficulty, so that an easy task claims less effort than a difficult one, and thus a high level of difficulty carries the risk that the resource demand exceeds the supply and will degrade human performance. In general, attention is "[...] the ability to focus on specific stimuli or locations" (Goldstein, 2010, p. 82), and can be assigned to different modes and levels, which are outlined and described briefly in the following.

Selective and Focused Attention: *Selective attention* is allocated if we want to select one specific location, object or message while ignoring unattended factors or stimuli in the surrounding. It is a kind of a filter, acting "[...] on incoming information, keeping some information out and letting some information in for further processing" (Goldstein, 2010, p. 83). Thus, selective attention enables us to visual select on what we want to pay attention to. If a task requests to pay attention to only one of several visual signals, then it is possible that we become unaware of significant changes in the visual display. This will be caused by a lack of attention leading to visual perceptual phenomena, like *inattentional blindness* (Simons & Chabris, 1999), although we are preattentive aware of visual events, which we are not attending (see section 3.3.3). Despite this adverse consequence, the ability to ignore unattended stimuli enables an individual to maintain a certain level of performance in a stressful environment. This is also known as *focused attention*, but also as selective sustained attention, which is the capability to concentrate on one specific task for a continuous period of time without being distracted, like prolonged payload tasks aboard the ISS. This may become challenging and is often varying in its time span, whereby attention rapidly starts decay. Hence, it is important to be able to directly refocus on the task. Consequently, the ability to focus and refocus is a deciding factor for task performance, responsible for concentration to produce a consistent result over time.

Divided and Alternating Attention: *Divided attention* is needed to attend two or more things at once and is often referred to as multitasking. By nature our brain is not able to process two tasks simultaneously and only really focuses parts of its attention on to each task. However, multitasking becomes possible if one task is quite familiar and automated processed, or is extensively trained for individual tasks. Similar but not equal, we are using *alternating attention*, if we do not attempt to perform different tasks at the very same time. Instead, we rapidly shift attention from one task to another, also referred to as *task switching*. For example, while building up a presentation during office work, calls are received and mails are checked, whereby attention is alternating back and forth. Besides switching between unrelated tasks, alternating attention is also required for related tasks, such as learning or problem solving by instructional guidance material (e.g., reading an instruction) and performing it (e.g., executing the instruction), as requested by payload operations. Even if the

tasks are related to each another, they are nonetheless requiring context updating, a process needed for "[...] reconfiguration of attention to meet changing task demands [...]", which in turn "[...] requires repeated revision of one's model of the environment [...]" (Kieffaber & Hetrick, 2005). Whether unrealted or related tasks, alternating between two tasks causes significant switch costs, resulting in less accuracy and slower performance (Jersild, 1927; Posner, 1980; Rogers & Monsell, 2005).

Arousal Level: The level of arousal can be considered as state of excitedness, which "[...] may influence the amount of attentional resources available to perform a task, as well as the policy by which attention is allocated to different tasks" (Proctor & Van Zandt, 2008, p. 244). Thereby stressful circumstances can cause a high level of arousal causing an increased readiness to respond that implies physiological changes, for example, of the heart rate and its variability, blood pressure, skin impedance or the pupil diameter (Wierwille, 1979), which may have adverse effects on task performance. This complies with the inverted U-hypothesis representing the *Yerkes-Dodson law* of performance (Yerkes & Dodson, 1908) that predicts that the best quality of performance is achieved at a medium level of arousal for complex tasks, while at a higher arousal level for simple tasks (see Figure 3.3, left). Regardless from the task difficulty, both too little and too much arousal can decrease the quality of the task performance. Thereby a very low level of arousal may cause a lack of attention that is needed to detect changes during the task fulfillment, which is of utmost importance in vigilance tasks. Such tasks require sustained focused attention, although events needed to be detected, are very irregular and rare. The longer the task time, the lower the arousal level and the more the vigilance will decrease. It is exactly the other way if the level of arousal or stress becomes too high and the relevance of processing unattended stimuli changes, causing a narrowing of attention, also known as *perceptual narrowing* or *attentional tunneling*. Thereby our attentional control will decrease and we become more focused, which restricts the number of cues that we can use to guide attention (Easterbrook, 1959). As was shown by Wickens and Alexander (2009) the effect of attentional tunneling can also be induced by synthetic vision displays, like head-up displays. This will be of significant importance related to AR displays and should be considered when designing interfaces.

Workload Level: To detect attentional or, in general, cognitive impairments caused by a system's design, and thus to detect failure in performance, it is important to consider what task demands (e.g., maintenance, experiment), system requirements (e.g., payload interface) and environmental factors (e.g., absence of gravity) are imposed on the user's limited resources. According to the conceptual framework of operator workload (see Figure 3.3, middle), the performance is not only reflected by the system performance, but, and equally important, by the effort spent by the user, which is influenced by user's goals, motivational state, knowledge and skills, as well as by user's cognitive processing capabilities. Workload can also be defined as "[...] the difference between the capacities of the information processing system that are required for task performance to satisfy performance expectations and the capacity available at any given time" (Gopher & Donchin, 1986, p. 41-3). Thus, mental or cognitive workload can be considered as an "*intervening variable*" (p. 41-4), whose value can be assigned to the total amount of mental demand or effort needed by a user to perform a mental task, and the greater the complexity, the greater the effort. It is clear that if the task demands are too low or too high (e.g., increased accuracy, strict timing, increased number of tasks, or extreme environment conditions), performance will be poor. Thus the relation between performance and workload results in an inverted-U shaped function (Figure 3.3, right), similar to the *Yerkes-Dodson law*. Thereby workload extremes are related to poor performance (Region 1, 3), while "with a reasonable level of workload, performance can be expected to be acceptable as shown in Region 2" (Lysaght et al., 1989, p. 9). To determine or predict the workload level, the mental cost to perform a

task is needed to be identified. This is optimally done with a multidimensional concept, using different assessment techniques, whereby each technique may provide different values (Williges & Wierwille, 1979; Wickens, 1984; Gopher & Donchin, 1986; Hockey, 1997). Traditionally, workload is assessed by four major techniques, which analyze primary task parameters, secondary task performance, physiological measures and subjective ratings. The most common technique applied to usability engineering in AR corresponds to analyzing primary task performance, which considers measures, on the one hand, of system performance (e.g., frame rate, latency etc.), and on the other hand of user performance that can be assessed by quantitative and qualitative measures. Thus it complies with the usability criteria for effectiveness and satisfaction. The other assessment techniques can be assigned to the criteria for efficiency, providing useful feedback on the resources spent by users, although often neglected in AR and, if applied, predominantly assessed by subjective ratings (Bai & Blackwell, 2012).

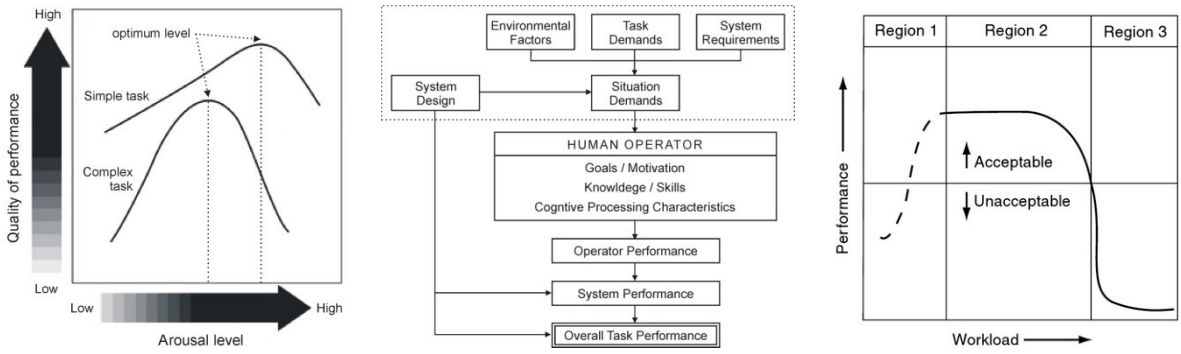


Figure 3.3: Performance related functions. *(left)* The *Yerkes-Dodson law* shows the relationship between performance and the arousal level for simple and complex task (from Proctor & Van Zandt, 2008, p. 245). *(middle)* The conceptual framework of workload and influences on operator/system performance (from Lysaght et al., 1989, p. 12). *(right)* The relation between performance and workload (from Lysaght et al., 1989, p. 10).

3.3.3 Visual Search

If we search visually for the presence or absence of a specific object or feature, which is explicitly defined by a task and held in the working memory, we actively explore our visual surrounding among an array of other objects, also referred to as *distractors*. Visual search not only causes scanning saccadic eye movements, it also implies *endogenous orientation* allocating visual attention, also referred to as voluntary shift of attention, which is top-down driven by the observer's goal, and thus the focus of attention can be manipulated by the requirements of the search task. Thus, endogenous orientation results in multiple focus steps by discrete eye movements from one location to the next, whereby temporary representations of the perceived objects or features are compared with the descriptions of the target objects stored in the working memory until a match was found that on average will happen after half of the objects have been inspected (Neisser & Lazar, 1964). However, if suddenly something changes in the visual periphery, our attention will be automatically directed by an external stimulus, activating fast bottom-up processing, which results in *exogenous orientation*, often accompanied by a reflexive saccade.

Considering goal-directed visual search only, which primarily relies on endogenous orientation and is interfered by unattended stimuli, it remains to be seen, how we process the large amount of visual information, because our information processing system is restricted, especially by limited attentional

resources. In cognitive psychology, visual search is the process of selective filtering of only the relevant information attended and it is suggested that this process claims attention in a different way. Related to this, the most popular theory is the *feature integration theory*, introduced by Treisman and Gellade (1980), which states that the perceptual image process relies on two stages related to objects' visual features implying a parallel and serial search. As shown in Figure 3.4, during the initial preattentive process the objects of the visual scene are rapidly analyzed into separate features, which are automatically registered in their feature maps (e.g., color, size, shape, depth) and therefore it is also referred as *feature search*, taking place in parallel over the entire visual display. It is not time affecting by the numbers of distractors and does not require the observer's knowledge, and thus it is efficient by increased bottom-up processing that claims attention only on a low-level. It is assumed that the feature maps "[...] preserve the spatial relations of the visual world but do not themselves make spatial information available to subsequent processing stages" (Treisman, 1986, p. 82). The theory also suggests that a subsequent *conjunction search* is needed to combine the objects' features, currently presented at a selected location of the visual scene. This conjunction search induces a step-wise serial search, whereby each step allocates focused attention "[...] whenever conjunctions of more than one separable feature are needed to characterize or distinguish the possible objects presented" (Treisman & Gellade, 1980, p. 97). Therefore, this is also referred to as *attentional spotlight*, operating a master map of locations. To combine the objects' features the observer's knowledge about these features is required, implying an effortful top-down process. Depending on how these features are unique and connectable, this process can be more or less efficient. Because focused attention is necessary to conjoin features correctly, features may be wrongly recombined when attention is diverted or overloaded (Treisman & Schmidt, 1982).

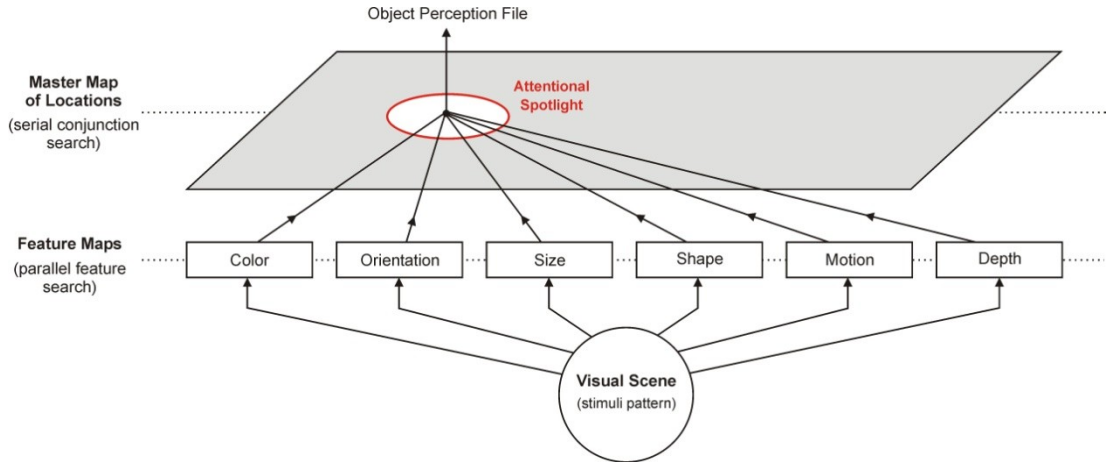


Figure 3.4: Visual search by Treisman's feature integration model (adapted from Treisman, 1986).

3.3.4 Instructional Design by the Cognitive Load Theory

Problems raised by goal-directed cognitive tasks should be solved or processed fluently through rapid acquisition of the learned schemas, which are stored in the long-term memory, ideally leading to automated processes without, or with only few conscious controls, which in turn wastes less processing resources. Consequently "without automation, performance is slow, clumsy and prone to error" (Sweller, 1994, p. 298) and thus human performance depends on the level of schema acquisition, and in cases of unexpected tasks, not learned or trained before, even on the degree of constructing, extending or linking schemas just-in-time. Solving a problem is often specified by a

hierarchical goal structure that can be supported by worked examples or guidance procedures. Thereby the overall goal can only be achieved if all sub-goals are conducted successful. Therewith, learning and performing a more complex task will consist of multiple sub-goals accumulating the costs of information processing, which may complicate schema acquisition and automation. Because the capacity of human information processing is limited by nature, failures in learning or training may be assigned to an improper allocation of attention and a high load of the working memory, which is explained through the *cognitive load theory* by John Sweller (1988, 1994), investigating the effects of cognitive load during problem solving and learning. This theory specifies two sources of cognitive load: the intrinsic and the extraneous cognitive load. It also describes effects, influencing the mental performance of the learner, and provides principles to present instructional information to reduce the working memory load. The sources of cognitive load and their effects are described in the following.

Intrinsic Cognitive Load: This source of load corresponds with the intrinsic nature of the instructional material (e.g., vocabulary, syntax, rules) and depends on the level of difficulty and element interactivity. The more difficult or complex the material is, the higher the intrinsic load becomes. Same applies to the element activity, which can be considered as the *element interactivity effect* (Sweller, 2008). Its effect size depends on the number of interacting elements that are concurrently required to understand the material. Elements that can be learned in isolation (e.g. technical terminology of a machine) contribute to a low element interactivity and thus to a low intrinsic load. Contrary, high element interactivity is caused by learning how components in a machine are linked to each other and how they interact during a machine operation, as requested by payload operations. Thus, understanding or comprehending of the learned material is benefitting from a high level of element interactivity, whereby the intrinsic load depends on the learner's prior knowledge and experiences. Because the intrinsic nature of material is fixed, its cognitive load cannot be reduced (Sweller, 1994).

Extraneous Cognitive Load: This source of load is affected by the presentation of the material to the person evaluating or solving a problem, and is determined by its instructional design or format. This load is denoted as "[...] unnecessary cognitive load and can be altered by instructional interventions" (Sweller et al., 1998, p. 259). Thus, the design of working examples or guidance procedures can be affected by their instructional design, especially those that call for integration of different and redundant multimedia contents (e.g., schematics, drawings, pictures). Such integration is often not effective, distracts attention and unnecessary loads the working memory. If the intrinsic cognitive load is low, a high extraneous cognitive load caused by a poor instructional design is acceptable. But if the intrinsic element interactivity is high, the level of the extraneous cognitive load becomes critical with complex material. In such cases, the extraneous load needs to be reduced, which can be achieved, for example, by considering the *split-attention effect*, the *redundancy effect* or by using different modalities of perception processing, increasing the *modality effect*. These three effects are influencing the design of instructions and are described in the following:

Split-Attention Effect: Such effect of divided or alternating attention presumes two related visual sources of information (e.g., text and drawing), which are presented separately and unintelligible in isolation. Such sources are not utilizable without each other until they have been mentally integrated before they can be understood, whereby both sources need to be held in the working memory (Chandler & Sweller, 1992). This can be cognitively demanding and waste unnecessary extraneous load of the working memory. To reduce the effect size of a split attention format, the level of physical integration of the sources needs to increase, whereby, in best case, both sources are fully merged into each other, (i.e. the instruction is embedded into the drawing). An example of a very low level of

integration is demonstrated in Figure 3.5, showing extracts of a Biolab procedure (see section 2.3.2). It contains a text-based instruction step *"4.2"* that references its related source *"Figure 2"*, which is separately presented two pages before. Because the text and the figure are unintelligible in isolation, both need to be mentally integrated.

Redundancy Effect: Redundant information is information, transfering the same content and is useful as long as different information is complementing each other (e.g., text and figure). As soon as each representation is showing the same content and each supporting a high level of comprehension, the redundancy effect will be triggered. Thus, this effect is generated by multiple redundant sources of visual information, which are self-contained and can be understood in isolation (Sweller et al., 1998). Its effect size depends on the level of understanding and is increased most, if a resource was fully understood. Then its integration into the working memory is ensured and is hard to ignore it subsequently. In case of multiple understood sources, all redundant information are mentally integrated and thus need to allocate cognitive load. As shown in Figure 3.5, the Biolab procedure also invokes the redundancy format. Thereby the step *"3.6"* is referencing two sources (*"Figure 2"* and *"Figure 4"*), both showing the same content. While *"Figure 4"* will be mentally integrated by a reduced split-attention format (close to each other), *"Figure 2"* will be also integrated, if it was understood in isolation. If it was not understood, it is easier to mentally ignore it.

Modality Effect: The split-attention and redundancy effects, each presumes the use of one perception processor (e.g., vision) and thus, after the multi-component model of Baddeley & Hitch (1974), such formats claim the usage of same component of the working memory (e.g., visuospatial sketchpad). It is proven that "[...] effective working memory capacity can be increased by using both visual and auditory working memory rather than either memory stream alone. Although less than purely additive, there seems to be an appreciable increase in capacity available by the use of both, rather than a single processor." (Sweller et al., 1998, p. 281). Thus, using a dual-modality presentation (e.g., visual perception of a drawing while listening to the related instruction) can reduce the level of the extraneous cognitive load.

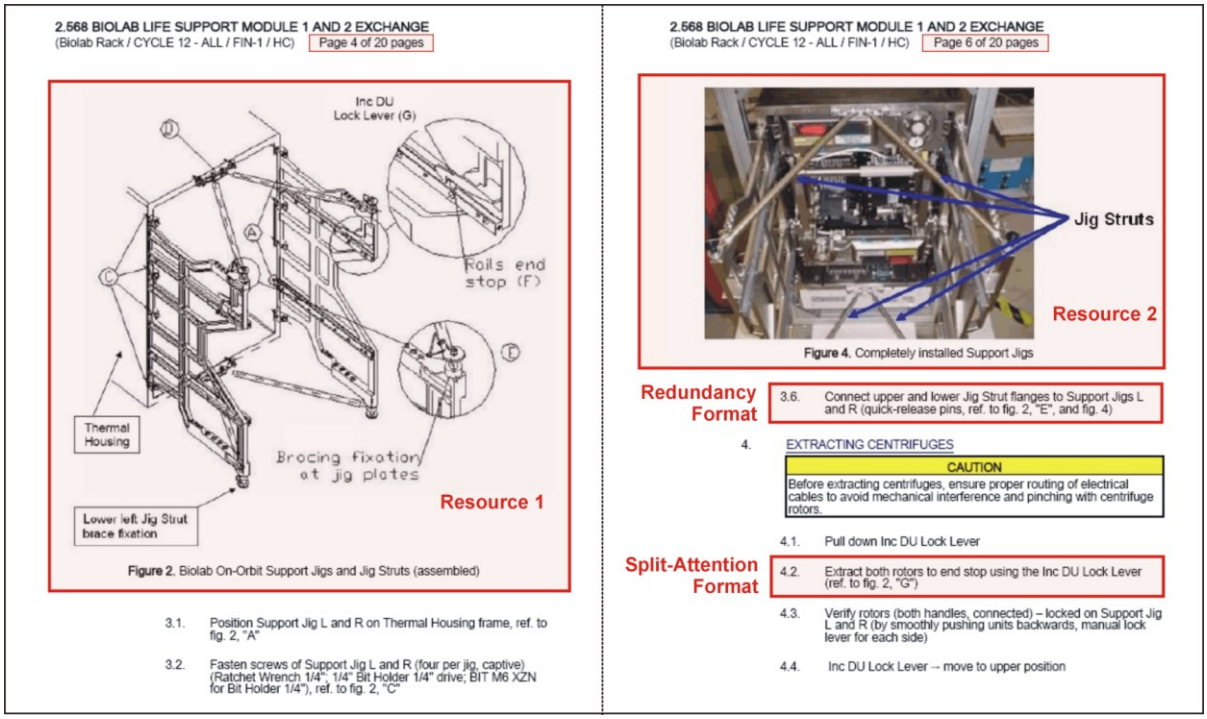


Figure 3.5: Inappropriate presentation of visual information resources in a procedure for the Biolab payload caused by split-attention and redundancy formats.

These effects mentioned above, assume that learners will be novice, which means that they have little or no knowledge about the given problem. While for novices using worked examples with instructional guidance may be essential to reduce cognitive load while problem solving, experts will benefit more from their knowledge and experiences than from studying instructional material. Such novice-expert difference is concerned with the *worked example effect* (Sweller & Cooper, 1985) indicating that learners using a worked example for the first time perform better on problem solving than learners who are experienced with the problem in question. This effect of solving an unfamiliar problem occurs because "[...] problem solvers have no choice but to use a random generation followed by effectiveness testing procedure at those choice points where they have insufficient knowledge to direct choice" (Sweller, 2008, p. 375). Thus, the benefit of instructional procedures depends on the learners' level of expertise, which in turn is associated with the *expertise reversal effect* (Kalyuga et al., 2003). This effect arises when instructional material contains redundant information or dual-modality instructions, and experienced learners cannot avoid to mentally integrate such information, meaning that they will be more capable to understand isolated sources according to the *redundancy effect* mentioned above. Then "[...] they must integrate and cross-reference this redundant information with their available knowledge schemas" which in turn claims "[...] excessive and unnecessary load on limited working memory resources" (Kalyuga et al., 2003, p. 29).

3.4 Applied Methods for Analyzing and Testing User's Needs

If a system design becomes more driven by technological aspects, it will result in poor usability and thus often will not be accepted by the end-users. This can be avoided by applying Human Factors methods, continuously considering the central importance of the users. Therefore Human Factors and usability engineering provide manifold tools and instruments to support a user-centered design process, which aims to result in a collection of guidelines to inform the user interface design. This section surveys related methods for analyzing the user needs as foundation for developing the underlying thesis's research strategy presented in Chapter 5. This section also outlines the applied methods and their characteristics to support an iterative design process and usability testing in particular.

3.4.1 Analyzing User's Needs

Because this thesis is primarily intended to redesign the current guidance interface for payload operations conducted by astronauts, it is important to analyze initially its impact on astronauts performance. In order to fully understand the users and the demands of their work situation, it is recommended by Human Factors engineering to conduct an initial *front-end analysis* (Wickens et al., 2003) that clarifies the current performance states, supporting the identification of adverse effects. The equivalent in usability engineering for virtual environments corresponds with a *user-task analysis*, which "[...] is the process of identifying a complete description of tasks, subtasks, and methods required to use a system, as well as other resources necessary for user(s) and the system to cooperatively perform tasks" (Gabbard et al., 1999, p. 8), and has been derived for AR environments (Gabbard & Swan, 2008). In keeping with the approach of a front-end analysis, the inspection of the current interface mode of payload operations was associated with answering the following questions, which are assigned to various sub-analyses: (Wickens et al., 2003, p. 17)

- User Analysis: Who are the system's users?
- Environment Analysis: What are the environmental conditions under which the system will be used?
- Function and Task Analysis: What are the major functions to be performed by the system, whether by person or machine? What tasks must be performed?
- Preference Analysis: What are the user's preferences or requirements for the system?

User Analysis: The user analysis is generally focused on identifying the potential user population and can be specified by, for example, age, gender, education, familiarity with the type of system or by task-relevant skills. If only a certain user population is assigned to a class of tasks, like astronauts to space payload operations, the criteria of its selection should be reviewed.

Environment Analysis: The *environment analysis* considers the conditions of the surrounding environment under which the tasks have to be performed. For example, such conditions could be defined by weather, type of clothing, mode of access or whether the system implies indoor or outdoor activities. Astronauts' work at ISPR payload implies space flight conditions, where the absence of gravity may impair the task performance. Thus, its effects need to be studied and considered when designing new user interfaces.

Function and Task Analysis: Most effort has to be spent on the function and task analysis that is generally based on detailed analyzing the functions accomplished by the underlying human-machine system and "[...] the tasks performed by the human to achieve those functions" (Wickens et al., 2003, p. 18). This should enable a better understanding of the interaction between an astronaut and a payload, while guided by procedures provided on the SSC-based interface (see section 2.3.4). A human-machine system is defined as "[...] a combination of one or more human beings and one or more physical components interacting to bring about, from given inputs, some desired output" (Sanders & McCormick, 1982, p. 14). Thereby the interfaces of both machine and human enabling a closed-loop system via controls and displays of the machine, as well as via the perceptual and sensorimotor system of the human, which is embedded within the larger context of the working environment. Consequently the input of the human is driven by the machine's interfaces, while the input of the machine is controlled by human's output. Hence a successful machine operation depends on human capabilities of physical and cognitively processing the perceived information. When discussing a human-machine system, the function analysis is focused on outlining the "[...] general categories of functions served by the system. [...] but not specify particular tasks" (Wickens et al., 2003, p. 18), while the task analysis identifies the actions and cognitive processes necessary for a user to perform a task or to accomplish a specific goal. A task analysis is primarily used to understand the actual work situation of the user and should be able to answer the following questions: (Hackos & Redish, 1998, pp. 8-9)

- What users' goals are; what they are trying to achieve?
- What users actually do to achieve those goals?
- How users are influenced by their physical environment?
- How users' knowledge and experience affects they thinking about their actual work?

The findings revealed from the function and task analysis conducted for the *astronaut-payload system*, should reflect the existing course of action and its characteristics and how tasks are allocated by the current payload guidance interface. To ease astronauts' work at ISPR payloads, it is vital to understand whether and how the mode of operation affects astronauts' information processing and thus their performance. Therefore a detailed analysis of the belonging perceptual, cognitive and physical activities was indispensable to identify the requirements that a redesigned guidance, for

example by AR, interface should meet. The output of the task analysis has also clarified which data or resources are involved and what are the contributions from astronauts' training.

Preference Analysis: The function and task analysis can be extended by identifying user preferences and requirements. A preference analysis is conducted by interviews and questionnaires. This enables to determine the key needs that comply with the identified user activities, but are related to issues that the users prefer, like the trade-off between functional and comfortable design. In a case where characteristics of the system are extensively preferred by the users, "[...] some attempt should be made to weight or prioritize them" (Wickens et al., 2003, p. 30). In the scope of this thesis, such analysis resulted from an initial proof-of-concept study, where application domain experts were involved.

Consequently, the initial front-end analysis in the scope of this thesis aims to understand the problem space in question (see section 5.1.1) and to justify the need for advanced user interfaces, manifesting the benefits from AR (see section 5.1.2). In consideration of the perceptual, cognitive and physical claims, it consists of various sub-analyses to outline and evaluate the involved *functions and tasks*, *resources* and the intended *user population*, as well as to discuss the impact of *training* and *environmental conditions* on task performance. Upon completing this front-end analysis, the user and system needs have been understood from a Human Factors perspective. But prior to the design process, a *competitor analysis*[27] was conducted to identify works that were or are currently performed by competitors related to this thesis's problem space. The analysis will be presented in section 5.1.3, which is focused on works intended to improve payload operations, ideally interfaced by AR. This enables to predict competitors' future intentions and more important to identify possible gaps. The output revealed from the front-end and the competitor analyses has contributed to the identification of the initial system specification, which induced an *iterative design process*.

Iterative Design Process: An iterative design process involves repetitive actions approaching a user-centered design, fulfilling Human Factors criteria and ensuring that humans intended to use the system, can perform safely and easily. It means that system's functions are tested and evaluated to make sure that their physical and cognitive demands are matching the human capabilities. In practice, the thesis's work followed such an iterative design process approach. It has continuously provided feedback that facilitates the understanding of the underlying causes of usability problems and succeeded to find alternative design solutions or countermeasures. This iterative activities included *heuristic design evaluations*, *prototyping*, *trade-off analyzing*, *usability testing* and *subsequent data analyzing*. During this process it was useful to answer the following questions: (Wickens et al., 2003, p. 30)

- Are there are any existing constraints with respect to design of the system?
- What are the Human Factors criteria for design solutions?
- Which design alternatives best accommodate human limits?
- Do the identified features and functions meet user preferences and requirements?

[27] One main issue while planning the design of new or existing systems is to position efforts made by competitors. The first question to be asked is who the competitors are. While doing so, this question is not focused on enabling technology, rather on the users involved to operate the system. Thus, systems in questions should be analyzed regarding to the attending users' needs. Although intended for business strategies, the framework of analyzing competitors presented by Porter (1998) is also usable in the context of usability engineering. It aims to answer two questions that are focused on two key aspects respectively. Thereby outlining their objectives and assumptions reflects what have motivated them, while analyzing their strategies and capabilities supports to understand what they are currently doing and are capable of.

Heuristic Design Evaluation: A heuristic design evaluation is an analytic evaluation method used to determine whether the system's characteristics comply with Human Factors criteria. Thereby the system design is systematically inspected for its usability to make sure that it meets not only Human Factors principles and guidelines, but also other theories, like from cognitive psychology. Typically a heuristic evaluation is conducted by usability experts and without participation of the system's users. In usability engineering, the goal of an heuristic evaluation, also referred to as *expert guidelines-based evaluation* (Gabbard et al., 1999; Gabbard & Swan, 2008), is the iterative process of finding problems in usability of a UI design and thus it is a method of "[...] systematic inspection of a user interface design [...]" (Nielsen, 1993, p. 155). In the context of this thesis, heuristic design evaluations have contributed to the identification of possible gaps and potential countermeasures related to Human Factors challenges, relevant for AR assisted space payload tasks, which are presented in section 5.2.

Prototyping: From the very beginning, an iterative design process benefits from early stage prototypes, which should exemplify the interface and interaction design. Prototypes can vary in their level of fidelity, ranging from paper-based to fully functional prototypes. While paper-based prototypes are very effective at the beginning because they can be redrawn with little costs, functional prototypes have more of the look and feel of the final system, supporting usability testing with real users. A functional prototype may, but does not necessarily have (like the *Wizard-of-Oz* technique, Dow et al., 2005), to support the full variety of functions that makes ideas concrete and support "[...] usability testing by giving users something to react to and use" (Wickens et al., 2003, p. 37). Prototypes implemented for the work conducted in this thesis, complied with fully functional prototypes, which have effectively supported the iterative design process. Because as already mentioned by Schwerdtfeger et al. (2009): "The problem is, to be able to elicit such user needs from future users, we need high-fidelity AR prototypes [..]. Those mature AR systems help elicit much more realistic user needs than low-fidelity prototypes." (p. 16).

Trade-Off Analyzing: A trade-off analysis is needed to perform if a design feature (e.g., an interaction modality) can be implemented in different ways and there are no data or guidelines available supporting the decision between alternative solutions. Then "[...] a small-scale study is conducted to determine which design alternative results in the best performance (e.g., fastest or most accurate)" (Wickens et al., 2003, p. 33), which is also referred to as *trade-off* or *trade study*. The iterative design process of this thesis has included multiple trade-off studies, which resulted from heuristic design evaluations.

3.4.2 Usability Testing and Data Analyzing

This subsection outlines the methods applied to the usability testing, which was requested by the thesis's inherent iterative design process. This includes the descriptions of important criteria needed to be considered when preparing usability tests. It also clarifies what types of data were collected and how they were statistically analyzed.

Usability Testing: If an element of a system requires the interaction with a human, its usability should be tested by real users and representative tasks. A usability test corresponds to either a *formative*[28] or a *summative evaluation* [29], varying in their goal settings and the level of experimental

[28] According to Nielsen (1993), a *formative evaluation* aims to find out "[...] which detailed aspects of the interface are good or bad, and how the design can be improved" (p. 170). It is typically conducted as a *thinking-aloud test*, whereby users are asked to verbally comment what they are doing, what are the reasons or what they would expect. Such a test

assessment of usability measures. Assuming a summative evaluation, usability testing is an empirical methodology claiming the validation of research questions or hypotheses to optimize interactive system elements and to subsequently provide design guidelines. Thereby a hypothesis is a prediction of the findings, which "[..] is framed in terms of the independent and dependent variables, stating that a variation in the dependent variable will cause a significant difference in the independent variable" (Dix et al., 2003, p. 330). While *independent variables* reflect the manipulated interface conditions (e.g., display type, interaction style), whereby different levels can be assigned to each variable, *dependent variables* are measured and affected by the independent variables, but should be unaffected by other confounding factors. Besides identifying such variables, another important issue is concerned with the way the test users are employed, resulting either in a *between-subject*[30] or a *within-subject design*[31]. Other criteria are directed towards the test users and the study environment. In general, the users recruited for usability testing should be as "[...] representative as possible of the intended users of the system [...]" (Nielsen, 1993, p. 175). At least, it should be ensured "[...] to evaluate the most diverse user population possible in terms of age, gender, technical ability, physical characteristics, and so on [...]" (Bowman et al., 2004, p. 363). Equally important is the study environment, which can vary in its level of realism and control, ranging from laboratory to field studies, whereas laboratory studies indicate a low level of realism but a high level of control and vice versa for field studies, both implying their own pros and cons in their pure form. Field studies are primarily used to get knowledge about the conditions of the target environment and users' behavior in it. Using such environment under daily business conditions usually implies continuous interruptions by, for example, phone calls or colleagues, and thus it is difficult to get a certain level of control over the experiment. However usability laboratories indicate specially equipped interruption-free environments offering better conditions for users' observation and data collection, but it is also "[...] possible to convert a regular office temporarily into a usability laboratory [...]" (Nielsen, 1993, p. 200). Although such environments enable controlled experimental conditions, the users' are taken out of the target environment, which may constitute a confounding factor, but can be compensated by the dual-task metaphor, increasing the level of realism (see section 3.4.3). "Ideally the design process should include both styles of evaluation, probably with laboratory studies dominating the early stages and field studies conducted with the new implementation." (Dix et al., 2003, p. 358).

In the scope of this thesis, each iterative design process has implied a summative evaluation using a within-subject design. Because the system's application domain encompasses space payload operations, the intended users are easy to identify, but astronauts are limited in their population and thus hard to get them to spend time on usability testing. With the exception of an initial

---

contributes mainly with subjective data collected either qualitative by a set of reported observations or quantitative by ratings scales.

[29] More results-driven and intended for comparative purposes, a *summative evaluation* is used to either evaluate the final design version or to support the decision between alternative design solutions comparing two or more user interface configurations within a trade-off study. It is usually conducted by a *measurement test* (Nielsen, 1993) collecting quantitative data (objective, subjective) and qualitative subjective data under controlled experiment conditions.

[30] A *between-subject design* ensures that a user or group of users is only assigned to one level of an independent variable, and is used when it is undesired that a user should perform more than one condition. Possible problems are assigned to "[...] the huge individual variation in user skills [...] Therefore, it can be necessary to run a very large number of test users in each condition in order to smooth over random differences between test users in each of the groups." (Nielsen, 1993, p. 178).

[31] A *within-subject design* indicates that each user receives all levels of the experiment, which not only implies the need of less users, it also contributes to an automatic control for individual variability, and thus it is "[...] easier to find statistically significant differences between experimental conditions" (Wickens et al., 2003, p. 495). Problems are related to inter-fering learning effects, which however can be reduced by a counterbalanced presentation of the independent variables.

proof-of-concept study, the performed tests correspond to trade-off studies related to relevant interaction issues. Thus, the test users recruited for the trade-off studies have not been necessarily needed to comply with the intended user group, but requirements to select a diverse user population were considered. Such studies were conducted under controlled laboratory conditions using either regular offices or environments enabling conditions for studying possible effects caused by altered gravity, as provided by parabolic flight or human centrifuge. Unlike the trade-off studies, the proof-of-concept study was intended to initially prove that payload tasks can benefit from AR. Thus it aimed at a high level of realism, which has called for a field test implying representative users and tasks within a payload environment, but also ensuring controlled experiment conditions.

Usability Measures: According to the concept of usability (see section 3.2), the usability tests performed in the scope of this thesis, were intended to assess the *effectiveness*, *efficiency* and *satisfaction* by collecting and statistical analyzing of usability measures. Thereby, user experience related to aspects of effectiveness was considered as a function of task performance feedback and measured by means of objective data, reflecting performance indicators, such as completion time, search time, response time, response frequency, error rate, speed and accuracy, but also considered tracking data to analyze head and goal-directed pointing movements. However, efficiency in general reflects the cost of human information processing related to the amount of attentional resources claimed by task performance. Thus, effectiveness was considered as a function of workload feedback that was measured by multidimensional indices as outlined and described in the next subsection. Besides effectiveness and efficiency, satisfaction was concerned with users' subjective feedback about the comfort and acceptability of the work system. Such feedback was particularly relevant in the scope of the proof-of-concept study, analyzing the preference and acceptance of application domain experts, whereby both quantitative and qualitative data were collected by ratings scales, questionnaires and interviews. However, for studying trade-offs between competing interaction modalities, the aspect of satisfaction was of secondary importance, because the associated research was focused on visuomotor skills and thus was more related to effectiveness and efficiency.

Data Analyzing: After quantitative data have been collected from objective and subjective feedback, the research questions or hypotheses needed to be evaluated by determining whether the dependent variables have changed as a function of the independent variables. Quantitative output can be classified as either discrete or continuous data. Discrete data were obtained by dependent variables implying simple counting of occurrences, such as pointing frequency, while continuous variables held measured values, like completion time, which could be mapped on an infinitely divisible scale. In the scope of this thesis, such data were analyzed by descriptive and inferential statistics using SAS®9.4[32] in its advanced mode supplied by the SAS® programming language, which includes, among others, predefined procedure statements (SAS® PROC), providing reports, graphics, statistics and analyses. Before analyzing, the raw sample data were pre-processed by the SAS® DATA step command. In cases where the sampling distributions exhibited skewness, it was adjusted by a natural logarithmic transformation. The following outlines the measures and methods used for descriptive and inferential statistics. Additional statistical analyses performed in the scope of this thesis, such as a resource trend analysis, a motion analysis and an analysis to evaluate the speed-accuracy tradeoff according to Fitts' law, required an extended data pre-processing, which will be described in the relevant chapters.

[32] Statistical Analysis Software (SAS), SAS Institute Inc., Cary, NC. URL: http://sas.com, last visit: 28.09.2016

Descriptive Statistics: Differences between experimental groups are descriptively reported by measures of central tendency and variability. While in the chapters' text passages the differences in the observations (*N*) are described by the *mean* (*M*) as arithmetic average and the *Standard Deviation*[33] (*SD, Stddev*) as a measure of the amount of variation, the chapters' appendices provide a more detailed description in tabular form (SAS® PROC TABULATE), additionally presenting the *median* as robust middle value, the *Standard Error*[34] (*SE*, *Stderr*) as a measure of the precision of the sample mean, although not descriptive, as well as the sample minimum (*Min*) and maximum (*Max*). In addition, distributions are graphically presented (SAS® PROC SGPLOT) by either box or series plots, but mostly by mean value plots that are marked with a confidence interval (CI) of 95%.

Inferential Statistics: Unlike descriptive statistics, inferential statistics enabled to make generalizations about the populations from which the samples were collected and thus was used to test hypotheses and validate research statements for statistical significance. Such tests were used to analyze the differences between group means and their associated interactions, using the SAS® LSMEANS statement to compute least squares means of fixed effects. Thereby the significance or confidence level α was set at .05 for all tests presented in this thesis. In cases of multiple pairwise comparisons ($m \geq 2$), it is important that the tests maintain the specified value of α, which is then also referred to as multiple level or experimentwise error rate. For example, using a *t-test* is not suitable, because the value of α would increase relative to the number of inherent comparisons and thus α would have assigned a much higher value than initially set. Hence, multiple pairwise comparisons, such as conducted in the scope of this thesis, claim to concurrent confidence intervals or tests, ensuring a reasonable multiple coverage probability (Schumacher & Weimer, 2006). Tests such as the ones according to *Bonferroni*, *Sidak* and *Scheffe*, although supporting multiple comparisons, are considered as conservative, because they usually do not use up the multiple level α. The same applies to the *Tukey-Kramer* test, but only for asymmetric data, as well as to the *Dunnett* method, which exactly maintains the multiple level, but only in tests using a single control. Thus, multiple pairwise comparisons of group means, requested by usability testing in the scope of this thesis, were conducted by the *Simulate* method (Edwards & Berry, 1987). This method is exclusively provided by SAS® and guarantees the specified significance level. The LSMEANS tests with Simulate adjustment were conducted, on the one hand, by the SAS® procedure PROC MIXED for continuous data and, on the other hand, by the SAS® procedure PROC GLIMMIX for count data. Models for count data were additionally fitted assuming a Poisson distribution and a logarithmic link function. Regardless from the used procedure, its output will be tabulated reported by the degrees of freedom of the test (*DF*), the normalized estimated *t Value*, the unadjusted *p* value ($Pr > |t|$) and the simulated adjusted *p* value (*Adj. p-Val*). This course of action described so far was applied to objective repeated measures, calling for parametric tests. In contrast, quantitative subjective data, obtained from rating scale scores, like the NASA TLX (Hart & Staveland, 1988), were statistically analyzed by non-parametric testing using the SAS® procedure PROC NPAR1WAY, requesting an analysis of the Wilcoxon scores as ranks of the observations. In cases of only two random samples, the Wilcoxon option produces the Wilcoxon rank-sum test, which does not require the assumption of normal distributions and is equivalent to the Mann-Whitney-U test. In the scope of this thesis, the Wilcoxon rank-sum test was used to assess whether the population mean ranks of two samples differed. Its output will be reported by the mean score *M*, the *Z* score, the *p* value and the rank correlation *r*, which reports the

[33] The *Standard Deviation* (SD) measures how close are samples to each other in same population and will not be affected by the sample size. While a small SD is indicating a uniform population, a large one indicates a more diverse population.

[34] The *Standard Error* is the standard deviation between the different sample's means and thus shows how close or accurate the sample mean is to the population mean. Typically it decreases if the sample size increases. If it is nevertheless large relative to the overall sample mean, then it indicates that the sample might not be representative of the population.

effect size of the test. However, in cases of more than two random samples, the Wilcoxon option induces the Kruskal-Wallis rank test, which investigates whether independent samples have equal medians in accordance with the null hypothesis of $H_0: m_1 = m_2 = \cdots = m_i$. To show differences between independent populations, H_0 was needed to be rejected. Therefore the resulting test statistic H was approximated by the chi-square (χ^2) distribution with $(i-1)$ degrees of freedom. If the test statistic H was greater than the χ^2 upper-tail critical value, H_0 could be rejected that has indicated a main effect between the independent samples. The output from a Kruskal-Wallis test will be reported by the value of the test statistic H including the degree of freedom and the number of observations, the p value indicating whether H is significantly greater than χ^2, as well as the mean ranks of the independent groups. In cases of a proven main effect, subsequent post-hoc tests for multiple pairwise comparisons were performed using the SAS® procedure PROC GENMOD applying the LSMEANS statement with the Simulate adjustment. Because the data complied with response data, taking values from a number of categories, the models were fitted assuming a multinomial distribution with a cumulative logit link function.

Note: In the scope of writing this thesis, all statistical tests were repeatedly executed. In cases of adjusted *p* values obtained by the SAS® Simulate method, this yielded divergent results than previously published, but not changed its main findings. This has resulted from not setting the SEED option, which specifies an integer used to start the pseudo-random number generator for the simulation. If such a value is not specified, "[...] the seed is generated from reading the time of day from the computer clock"[35].

3.4.3 Dimensioning the Workload Assessment

Besides effectiveness and satisfaction, efficiency as indicator for usability reflects "the resources expended in relation to the accuracy and completeness of goals achieved" (ISO 9241) or, in other words, "efficiency takes into account the costs of delivering effective action [...]" (Hockey, 1997, p. 83). As have been already mentioned in the last subsection, the efficiency of task performance in usability testing is related to the amount of attentional resources claimed by human information processing, and thus, in the scope of this thesis, efficiency was considered as a function of mental workload (see section 3.3.2). With a fast growing application of Augmented Reality (AR) to support instructional tasks, the importance to assess its efficiency by workload measures is becoming increasingly significant, but usability testing is mostly restricted to subjective ratings (Bai & Blackwell, 2012), although "[..] workload is generally considered to be complex and a multidimensional concept" (Lysaght et al., 1989, p. 16). As already mentioned in section 3.3.2, workload can be assessed by *primary task performance*, *subjective ratings*, the dual-task metaphor using a *secondary task* and by *physiological measures*. If a system is tested related to its usability, the performance using this system is needed to be evaluated at first, because this is the goal of interest. Primary task measures are most commonly used in AR studies to evaluate the effectiveness of systems or interaction techniques, but may not always be sufficient to verify clearly the benefits of an optimized system. Thus, usability testing conducted in the scope of this thesis has considered workload assessment in all its dimensions and sometimes using a twofold or threefold approach. The

[35] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mixed_sect014.htm, last visit: 28.09.2016

applied measurement methods are described in the following. A list providing the full degree of the workload measurement techniques are provided by Cain (2007).

Subjective Workload: Assessing the subjective workload complies with self-reported effort spent on task performance and is the most simple method to assess workload. Because such subjective measures did not consider issues of the task environment, "[...] it may be best to couple their use with primary task measures [...]" (Proctor & Van Zandt, 2008, p. 254). Further limitations are concerned with the difficulty to report perceived efforts and with the inaccessibility of loading factors. Subjective workload is usually assessed by using a rating scale. The most prominent standardized rating scales are the Cooper-Harper scale for assessing the overall workload (Cooper & Harder, 1969), the Subjective Workload Assessment Technique (SWAT, Reid & Nygren, 1988), the NASA Task Load Index (NASA TLX, Hart & Staveland, 1988) and the Workload Profile (WP, Tsang & Velasquez, 1996), whereby the latter three comprise multiple subscales. Because the NASA TLX is the most widely used rating scale, it was applied to assess the subjective workload for usability testing in the scope of this thesis. It consists of six loading factors, requesting to rate the mental demand, physical demand, temporal demand, performance, effort and the frustration level. The rating is done for each loading factor on a continuous scale with two bipolar descriptors (high/low) at each end. A weighting scheme can be used to consider individual differences while computing an overall workload score for each experimental condition. It can also be used in its raw nature (RTLX), without weighting the factors (Byers et al., 1989), whereby "[...] the ratings are simply averaged or added to create an estimate of overall workload" (Hart, 2006, pp. 906). In the scope of this thesis, the NASA RTLX was used multiple times for usability testing (Chapter 6, 7, 8), and is presented in Appendix A.

Secondary Task: The secondary task technique is based on the dual-task metaphor, whereby the user is requested to perform a second task, concurrently to the primary one. Capacity theories as theories of attention, such as discussed by Moray (1967), Kahneman, (1973) and Wickens (1984), point out that the workload is affected by limited capacity resources. Thus, the "secondary task performance is assumed to be inversely proportional to the primary task resource demands" (Wickens & Hollands, 2000, p. 462). This means if the capacity is limited to the primary task, additional resource allocation will reduce the amount of resources that are available for the concurrent task, and consequently, the measures of the secondary task will be more sensitive than that of the primary task. Commonly used tasks for concurrent performance are simple reaction time tasks, choice reaction time tasks, monitoring tasks and mental arithmetic tasks. Besides assessing the secondary task measures, the dual-task paradigm can also be used to analyze the cost of dual-task processing, assuming that the primary task measures were also collected in a single task condition. The dual-task interferences can then be declared as percentage shift between the single-task and dual-task scores. In the scope of this thesis, the dual-task metaphor was applied to one trade-off study (Chapter 7). Therefore a simple reaction time task was used for dual-task performance with related single-task tests to additionally assess the dual-task costs.

Physiological Workload: The physiological assessment of the workload is used to measure arousal, whereby "the general level of arousal is presumed to increase as mental workload increases [...]" (Proctor & Van Zandt, 2008, p. 253). It enables to detect workload changes in a way different from the above mentioned techniques, because it provides online data about the underlying physiological processes involved in task performance, which are usually not directly accessible. Although it is a widely common practice in workload research, care should be taken whether the used method is, on the one hand, appropriate to detect workload changes and, on the other hand, sensitive

to the variations in the workload that the user will experience (Lysaght et al., 1989). Besides analyzing the blood pressure, skin impedance, body fluid and many other indices (Wierwille, 1979), the most promising candidates to assess the workload physiologically are the heart rate and its variability and the pupil diameter. So for example, the pupil diameter reflects the amount of attentional resources that are spent on task performance, whereby a larger size of the pupil indicates an increased workload. Using the pupil diameter is a reliable assessment method, although revealing only small changes. In contrast, cardiac responses, like the heart rate or the heart rate variability (HRV), are influenced by the autonomic nervous system (ANS) and thus can reflect physical and emotional states. Because the heart rate is "[...] related to the amount of physical activity, respiration and thermal regulation" (Lysaght et al., 1989, p. 141), effects changing mental activities are also noticeable within the heart rate. Contrary to the heart rate, which increases if the workload increases, but not always provide changes in the workload, the HRV is more suitable to assess mental load, as it describes the range in which heart rate changes in a certain time (Proctor & Van Zandt, 2008). Thus, it is probably the most commonly used technique to physiologically assess the mental workload (Meshkati, 1988). In the scope of this thesis, the physiological workload was assessed by cardiac response, analyzing the HRV, which can be considered as an immanent expression of sympathetic and parasympathetic influences of the heart function (Task Force, 1996), whereby the sympathetic nervous system helps us to deal with stress, while the parasympathetic nervous system controls our bodies' responses during rest. The HRV provides a set of parameters, which can be assigned either to the frequency or the time domain (Task Force, 1996). The parameters used to analyze the physiological workload in the scope of this thesis will be described in the relevant chapters, but in general comply with the LF/HF ratio (Chapter 7), the Stress Index (Chapter 8) and the R-R distance (Chapter 9). The HRV was recorded by Electrocardiograms (ECGs), electrical signals originating from the heart, which was measured by two wireless electrodes placed at the participants' thorax, which detect muscle activity on the skin surface.

As mentioned above, assessing the mental workload in AR usability testing is almost restricted to subjective ratings (Bai & Blackwell, 2012), predominantly using the NASA TLX. Only few studies have also considered physiological measures for workload assessment (Tümler et al., 2008; Schwerdtfeger et al.; 2009; Grubert et al.; 2010), but have also shown how challenging the evaluation of physiological data turns out. In general, such data are featuring high variations that needed to be limited or partially eliminated, for example, by reducing the participants' inter-individual differences (e.g., age, gender, BMI), increasing the sample size and by statistical normalization using baseline measures. Even less studies have applied the dual-task metaphor assessing secondary task measures (Grubert et al., 2010; Kishishita et al., 2014), although Grubert et al. applied the dual-task metaphor only during pre- and post-test-phases and not during the experimentation phase.

Chapter 4

Related Work on Instructional AR Beyond Spaceflight and Advanced User Interfaces for Spaceflight Operations

This chapter provides a survey of selected works, previously done in the field of Augmented Reality that should illustrate the continuous progress on its individual adaptation to instructional tasks beyond spaceflight. This survey also considers works on related issues, like task localization and the placement of interactive AR control interfaces, both relevant and important to assist instructional tasks using AR. Finally, and most related to this thesis's problem space, this chapter provides a review of works on advanced user interfacing intended to support astronauts' work and projects specifically claiming the usage of AR will be presented more detailed.

4.1 Relevant Issues for Instructional Augmented Reality Beyond Spaceflight

After Ivan Sutherland had laid the foundation for Augmented Reality (AR) by releasing the relevant technology in 1968, first applications appeared more than twenty years later to support tasks for medical (Bajura et al., 1992) and manufacturing purposes (Caudell & Mizell, 1992), both belonging to the most meaningful areas of applications supported by AR. Relevant delimitations and definitions for AR were only defined a few years later. While Milgram and Kishino (1994) classified mixed reality approaches to their ratio between the physical and virtual reality, Azuma (1997) defined AR by its three crucial characteristics stating that AR interactively combines real and virtual worlds by 3D-registered information in real-time. Since then until today, these specifications and the upcoming level of awareness for AR have motivated numerous works in various fields, which have driven a remarkable technological progress which in turn resulted in a fast growth in applications (Azuma et al., 2001; Zhou et al., 2008; Schmalstieg et al., 2010, Krevelen & Poelman, 2010). Besides medicine and manufacturing, the application of AR covers various other areas, like architecture, entertainment, sports and education. Thereby most applications share the common aim of increasing the perception of the surrounding to assist or support a given task. Thus the value of AR yields the greatest potential for task support and mostly contributes to aid a human to more easily carry out complex tasks such as maintenance, assembly, repair or service, which are usually guided by instructions, constituting a manual, commonly assigned to a technical device and in most cases explaining how to assemble or operate it. Regarding the values of AR, Azuma (1997, p. 359) stated that "Instructions might be easier to understand if they were available, not as manuals with text and pictures, but rather as as 3-D drawings superimposed upon the actual equipment, showing step-by-step the tasks that need to be done and how to do them.".

In this section, the importance of AR for instructional tasks beyond spaceflight is emphasized by reviewing remarkable works done in this field showing a beneficial application of AR, works investigating the spatial orientation of AR instructions to enhance human attention control and works

analyzing the placement of interfaces that are needed to control such tasks. Even if the medical area offers a potential place to support tasks guided by AR, this review does not touch this application area, because of their more challenging requirements (e.g., dynamical aspects for 3D-registration) and the fact that physicians are more willing to accept deficits in handling and visualization when the application's value significantly contributes to their work (Markov-Vetter et al., 2009). To make instructional Augmented Reality (AR) executable in a smart technical environment and before beginning research on usability aspects on, for example, task localization and interaction techniques, there are a number of fundamental research issues on display and processing hardware, 3D pose tracking, calibration and real-time rendering, which are of significant importance, but outlining their progresses would go beyond the scope of this section and is not in the main focus of this thesis. Outstanding advances in the field of AR can be annually followed at the *International Symposium on Mixed and Augmented Reality* (ISMAR[36]) as successor of the *International Workshop on Augmented Reality* (IWAR) launched in 1998 (Zhou et al., 2008). As already mentioned in the last chapter, considering Human Factors by concepts of usability engineering is critically important for users' acceptance. Hence, the works presented in the following subsections, will be reviewed and discussed regarding their usability testing , if available.

4.1.1 Instructional AR for Maintenance, Assembling and Service Tasks

Since the advent of Augmented Reality in 1992, countless works and numerous research projects have been carried out, which have applied and studied the potential of instructional AR to support assembly, constructional, repair, maintenance or just service tasks. It can be stated that several works have been published every year that reflect the potential of AR to support instructional operations in such task areas, most of them surveyed by Ong and colleagues (2008). Also the increasing incidence of research consortiums or academic industry-coupled projects, like STARMATE (Schwald, et al., 2001), ARVIKA (Friedrich, 2002; Weidenhausen et al., 2003), ARTESAS[37], STAR (Raczynski & Gussmann, 2004), AVILUSplus[38] or ManuVAR[39], are emphasizing the interest in this application area. Regenbrecht and colleagues (2005) have discussed the rare advent of AR in an industrial context and reviewed R&D (Research and Development) projects in the area of automotive and aerospace industry, as well as for production support. Even the oil and gas industry investigated the potential of virtual and mixed reality on mobile data visualization for their maintenance systems (Espindola et al., 2013). Another work aimed to improve the efficiency and safety of maintenance tasks for nuclear power plant (Ishii, et al., 2007). A survey of industrial AR applications was provided by Fite-Georgel (2011), describing the life-cycle that was used as a taxonomy to categorize and present a selection of papers. Due to the abundance of works in the field of instructional AR, this review is limited to a selection of prominent works, which have essentially contributed to advances in instructional AR and can be considered as representatives for all other works. Thereby special attention is paid on works, which have solved physical-world problems benefitting from AR. Specific research conducted from Henderson and Feiner, (2009, 2011a, 2011b) will be discussed more detailed, because they are most related to the work on instructional AR presented in this thesis.

As mentioned above, pioneering with their early work in 1992, Caudell and Mizell have supported tasks for aircraft manufacturing. On behalf of Boeing they should assist and speed up the task of

[36] ISMAR (International Symposium on Mixed and Augmented Reality). http://ismar.vgtc.org/, last visit: 28.09.2016
[37] ARTESAS project (2004-2006). URL: http://www.wzl.rwth-aachen.de/de/aa272c5cc77694f6c12570fb00676ba1.htm, last visit: 28.09.2016
[38] AVILUSplus project (2008-2011). URL: http://www.avilusplus.de, last visit: 28.09.2016
[39] ManuVAR project (2009-2012). URL: http://www.manuvar.eu/, last visit: 28.09.2016

installing electrical wirings in the aircraft fuselage. As a result, the technicians were equipped with a head worn display enabling the projection of wiring diagrams superimposed on parts of the aircraft on which the technicians were working (see Figure 4.1). In the scope of their work, Caudell and Mizell have not only presented the first prototype for a manufacturing task, they also were the first to introduce the term *Augmented Reality* by specifying it as overlay on top of the real world with digital content and they soon realized the necessity of a high accuracy during the 3D registration process to ensure reliable operations. Related field testing conducted by Curtis et al. (1999) revealed a functional prototype but lacked adequate acceptance because of practical issues. Other works in the realm of aviation have shown the importance of reliable interfaces for the performance of AR guidance during pre-flight inspection of a *Cessna 172* (Ockerman & Pritchett, 1998), and discussed that AR could essentially contribute to aviation maintenance.

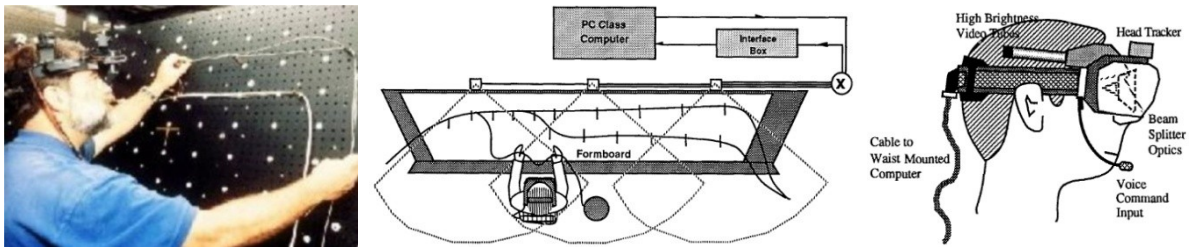


Figure 4.1: Pioneering work by Caudell and Mizell for aircraft manufacturing (from Caudell & Mizell, 1992).

Similar work was conducted by De Crescenzio et al. (2011) also investigating the benefit of AR for daily inspection tasks of a Cessna 172 focused on the maintenance check, which is required before the first flight and commonly supported by an aircraft maintenance manual and a flight manual (see Figure 4.2). Their AR system consisted of an adjustable headset equipped with a see-through Liteye LE750 monitor and a Logitech webcam, as well as a notebook computer. To avoid the use of fiducial markers for 3D pose tracking, they applied markerless tracking, using natural features using the SIFT[40] and SURF[41] algorithm. The system was successfully validated with ten participants in a field test, which included instructions for using the AR system and checking the oil, but it was not directly compared with the traditional method. They evaluated the mental workload by subjective experiences assessed by the NASA TLX. The average workload (see Figure 4.2, right) was acceptable rated in respect of the mental, physical and temporal demands as well as for the effort needed to accomplish the tasks, while the performance and satisfaction were rated poorly.

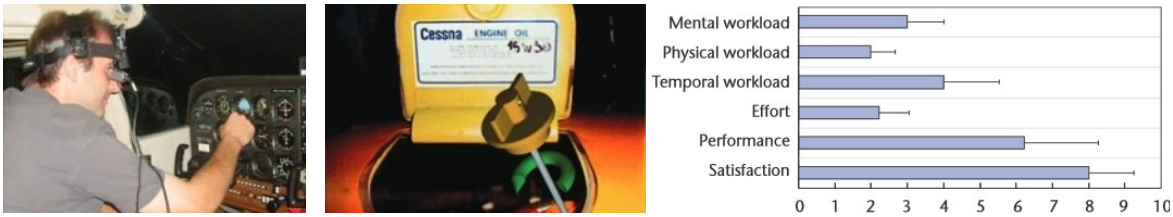


Figure 4.2: AR for daily inspection tasks of a Cessna 172 (from De Crescenzio et al., 2011). *(left)* The see-through Liteye LE750 equipped with a Logitech webcam. *(middle)* A dipstick model indicates the task while a 3D arrow shows the direction. *(right)* Participants' average workload rated by the NASA TLX.

[40] SIFT - Scale-Invariant Feature Transform (Lowe, 2004)
[41] SURF - Speeded-Up Robust Features (Bay et al., 2008)

Beside aviation, AR also has promising potential to support tasks in the automotive industry. For example, Reiners et al. (1999) have early introduced their system for doorlock assembly in the auto-motive industry, which was detailed discussed towards challenging system issues, like fiducial marker tracking or see-through calibration, but lacked in usability testing, although causing great interest when it was presented at the Hannover Industrial Fair. Similar to this work, Wiedenmaier et al. (2003) applied AR also to the automotive domain by presenting AR support for the assembly of an auto-mobile door, but evaluated the effect of instructional media on door assembly by comparing a paper instruction with AR and an expert tutorial. The long-term experimental task (mean paper time: ~ 1h 30min) included three different levels of difficulty (mounting window regulator, wiring and fixing clips). Analyzing the total assembly time revealed that the AR support could reduce the required time related to the paper instruction by about 35 percent, but guidance by an expert attained a reduction of about 65 percent. The reason for that was that difficult tasks benefit more from AR than intuitive and repetitive tasks, like fixing the clips. Although not for the auto-motive industry, similar conclusions have been made by Syberfeldt et al. (2015) stating that "the complexity of the assembling task must be high enough for the user to feel that it is worth using the AR system". Another study published by Platonov et al. (2006) showed interesting work for the maintenance of a BMW7 Series engine (see Figure 4.3). Even though this work was mainly focused on the development of a suitable tracking algorithm using a markerless, model-based approach based on 2D point features, it was finally tested successful by different BMW staffs. Similar work in the automotive sector was done by Gay-Bellile et al. (2012), also contributing to the development of 3D pose tracking by introducing a generic tracking framework based on model constrained SLAM[42] capable of handling various automotive AR applications with high accuracy and robustness. Using a Renault Clio, the experimental results have shown that their flexible framework can be adapted to a wide range of automotive applications, like the maintenance and repair of the engine or virtual customization of the car bodywork or dashboard.

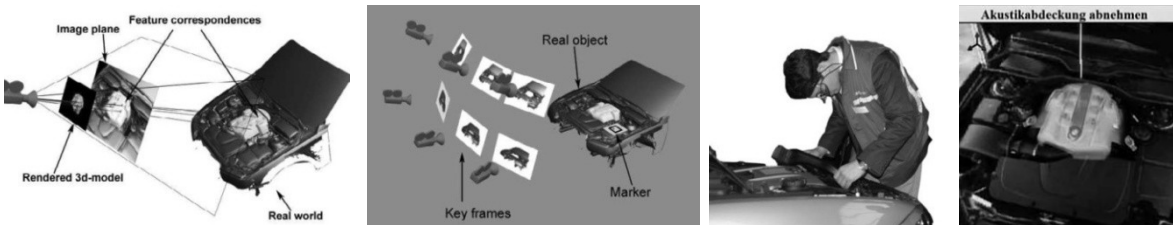


Figure 4.3: A mobile Augmented Reality system for automotive applications (from Platonov et al., 2006).

Another domain of researching AR for manufacturing tasks was the food and beverage packaging industry, such as the Sidel company, a manufacturer of machinery for PET bottling and packaging. In the scope of the project SKILLS[43], which aimed towards the transfer of human skills by means of multimodal interfaces, AR was studied for industrial training addressed to assembly and maintenance tasks (Webel et al., 2011a; see Figure 4.4). By direct and indirect visual hints the trainee was stepwise guided through the training procedure and could, if needed, request remote support by an expert. Instead of using a head-mounted display (HMD), they favored handhelds and monitor based devices, because these are more suitable for industrial environments (Bischoff & Kazi, 2004). For multimodal purposes and to unburden the human visual channel, a vibrotactile bracelet was used to direct the hand movements by haptic rotational or translational stimuli. The system was tested by four expert

[42] SLAM - Simultaneous Localization And Mapping (Durrant-Whyte & Bailey, 2006)
[43] SKILLS project (2006-2011). URL: http://www.igd.fraunhofer.de/en/Institut/Abteilungen/VRAR/Projekte/SKILLS, last visit: 28.09.2016

trainers, who were requested to assemble a valve with the AR system (Webel et al., 2011b). Evaluating subjective data showed that the training of such tasks benefits from AR support.

Figure 4.4: The project SKILLS for "Multimodal Interfaces for Capturing and Transfer of Skill" (from Webel et al., 2011). *(left)* A service technician is using the AR-based Training at Sidel training center. *(middle)* AR-Training of an Assembly Task using the indirect visual aid. *(right)* Direct visual aid for a maintenance task.

However, taken the view to the category of instructional tasks that should be assisted by AR, it stands out that most of the work was done in the assembly task domain. Beside the above mentioned works for door assembly, Baird & Barfield (1999) investigated the benefit of AR support on assembling a computer motherboard. Therefore they conducted a usability study comparing different guidance methods using a paper-manual, a computer-aided manual and an AR considering video (VST) and optical (OST) see-through HMDs. Related to assembly times and errors, AR improved this task, whereby the OST display supported the fastest assembling. Also in 1999, Boud and colleagues studied VR and AR as a training tool for a water pump assembly and compared conventional 2D engineering drawings with two VR desktop conditions, one immersive VR condition using a HMD and one context-free AR condition, which have displayed a static pictorial representation of the water pump's assembly sequence. All VR conditions and AR enabled faster assembling than using the drawings, whereby the assembling task mostly benefited from AR, but as concluded by the authors, AR is disadvantages by the physical objects needed to be present during the training, which was not the case for VR and "[...] in some cases it may not be practical to interact with the real objects due to their non-availability and their associated costs" (Boud et al., 1999, p. 36). Another work, presented by Zauner and colleagues (2003), considered the potential of AR applied on furniture assembly (see Figure 4.5). Although they work was predominantly focused on authoring issues, this work contributed with a hierarchical data structure enabling a state mechanism for the assembly plan and various visualization effects (e.g., highlighting, 3D objects, placement animations), assignable to specific tasks.

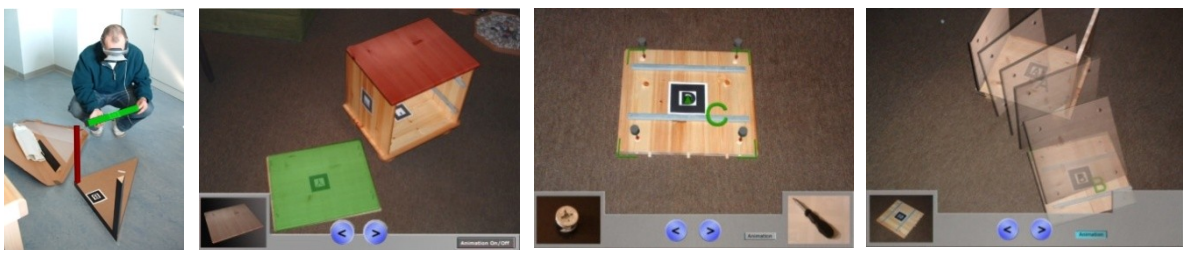


Figure 4.5: Assembling furniture by AR to install table-legs or build wooden chests supported by highlighted components, indicated positions for nails, or by animations for the placement (from Zauner et al., 2003).

Works not considering a specific task area discussed mostly the key features of instructional AR systems or aimed to implement more generic approaches. Thus, Schwald and Laval (2003) discussed hard- and software issues and introduced an AR prototype to support diverse tasks belonging to the

maintenance domain, but lacked usability testing. The same applies to work done by Pang et al. (2006), but they implied that "more interaction tools need to be developed to support the complex assembly design efficiently" (p. 34) and therefore introduced their virtual interactive tool for assembly guidance (Yuan et al., 2008). Similar to Zauner et al. (2003), Knöpfle and colleagues (2005) were focused on template-based authoring of AR manuals, but was intended to assist mechanics during various maintenance and service tasks. Therefore they analyzed typical service tasks to classify the related operations for generalization purposes, but only discussed implementation details and also lacked in usability testing.

In case of lacking a real-world problem needed to be solved, assembling toy blocks or puzzles have often been utilized as suitable examples of use cases. Thus, Tang and colleagues (2003) investigated the effectiveness of AR in assembling toy blocks by carrying out a between-subject user study performed by 75 participants, in which they compared four different guidance techniques using a printed media, a LCD monitor, as well as unregistered and 3D-registered AR provided by an OST HMD (see Figure 4.6). Besides evaluating performance measures, like completion time and accuracy assessed by the number of error, they also investigated the effect of instructional media on mental workload assessed by subjective experiences using the NASA TLX. They have shown that 3D-registered AR led to significant faster performance than using the LCD, but not revealed timely effects between the other conditions. The results have also shown that 3D-registered AR caused less errors at all. Evaluating the ratings scores of the NASA TLX verified their hypothesis, stating that the type of instructional medium significantly affects the mental workload. Works by Robertson et al. (2008) and Khuong et al. (2014) also used toy blocks, but investigating localization in the scope of AR, which will be reviewed in next subsection. However, Salonen & Sääski (2008) used a 3D puzzle to emulate an assembling task. They showed that AR assistance can reduce the assembly time and has a high effectiveness ratio on such tasks.

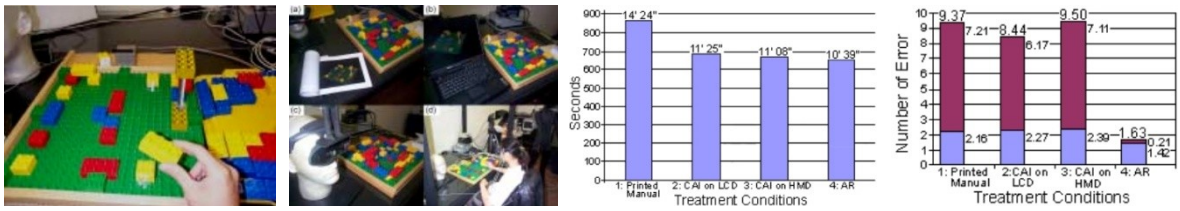


Figure 4.6: AR for toy blocks assembly (from Tang et al., 2003). *(left)* HMD view showing virtual toy blocks indicating the correct location. *(middle)* The studied conditions: (a) paper-manual, (b) LCD, (c, d) AR conditions. *(right)* Assembling time per condition: the LCD and the two AR methods only differed slightly.

In the scope of the COGNITO[44] project, Mura et al. (2012) used puzzle assembly to study the effect of display hardware by comparing monocular and binocular HMDs. They showed that puzzle assembling was not affected by the display, although the participants have significantly preferred the handling by the monocular HMD, because it was significantly faster ready for use than the binocular HMD. In the same publication they reported on a second study that investigated the effect of the type of instruction on the performance of two assembling tasks differing in their task difficulties (see Figure 4.7). Thereby the simple task was realized by attaching two batons on a base plate (nails and screws), while the other tasks required the disassembling of a ball valve. These tasks were guided by a paper manual (baseline), AR support (Wizard-of-Oz) and a video instruction watched by participants directly before starting the performance. The results revealed that the video support

[44] COGNITO project (2010-2012). URL: http://www.ict-cognito.org, last visit: 28.09.2016

enabled the significantly fastest support compared to the paper manual for both tasks (see Figure 4.7, right). The more complex task could also significantly benefit from the AR support than the simple task, but did not outperform the video condition. These findings are consistent with observations from Wiedenmaier et al. (2003) and Syberfeldt et al. (2015) stating that difficult tasks can benefit more from AR than simple tasks.

Figure 4.7: Exploring the effect of instruction type on performance of two assembling tasks differing in their task difficulties (from Mura et al., 2012). *(left)* Simple task: attachment of two batons on base plate (nails and screws). *(middle)* Complex task: disassembly of a ball valve. *(right)* Assembly times by instructional media.

From 2008 until 2010, an interesting series of work was presented at the ISMAR conference, investigating the potential of AR to support the order picking process in logistic applications focused on user-related aspects (Tümler et al., 2008; Schwerdtfeger et al, 2009; Grubert et al., 2010). Thereby each work conducted long-term tests comparing the common method with AR guidance and measuring the physiological workload by the heart rate and its variability (HRV, see section 3.4.3), but assessing different parameters. The initial work carried out by Tümler et al. (2008) is actually a continuation of Schwerdtfeger et al. (2006), which firstly provided location-based AR wayfinding and picking information on mobile displays intended to support logistic systems. Because this work was more focused on presentation schemes for task localization, it will be presented in the next subsection. Coming back to the work of Tümler et al. (2008), it was one of the first applying HRV to investigate the physiological strain in mobile AR technologies, but doing so for a manual task under long-term condition. They used a reference environment based on an industrial warehouse picking scenario arranged within an area that was prepared for optical marker tracking (see Figure 4.8, left). The apparatus consisted of a monocular OST HMD, equipped on top with a camera for tracking purposes. The study was conducted with twelve participants, who were requested to perform multiple picking tasks for two hours. Tümler and colleagues evaluated the physiological strain by analyzing the heart frequency (HF) and the standard deviations of the R-R intervals (SDNN) assessed by HRV, which was successive obtained from a long-term ECG recorder (MT-101, Schiller GmbH). The HRV data were continuously registered during several ten minutes time phases at the begin, in the middle and at the end of the work task, as well as during pre- and post-HRV-tests in lying and standing postures as baseline measures. The results showed that the HF (see Figure 4.8, middle) was not affected after two hours of work with both AR and paper, while the SDNN (see Figure 4.8, right) revealed significant differences between AR and the paper manual at the beginning work phase, resulting in a higher SDNN with AR implying a higher strain, but returned closer to the paper-manual values towards the end of the test. The authors assumed that this was caused by initial familiarization during the use of the AR setup. This suggests a lack of a prior short training session. Tümler et al. also evaluated the subjective strain by discomfort and sensitivity questionnaires, but did not reveal differences between the guidance methods. Thus, this study has shown that the "change of strain caused by an AR system developed similar strain changes caused by a common non-AR system (paper list)" (Tümler et al., 2008, p. 90).

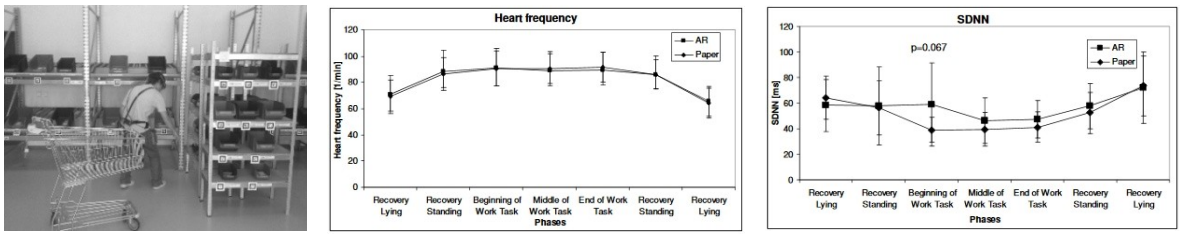


Figure 4.8: AR for order picking tasks (from Tümler et al., 2008). *(left)* Subject with shopping cart in rack area. *(middle)* The heart frequency (HF) through the test phases. *(right)* The HRV SDNN through the test phases.

Besides using a different warehouse setup, the follow-up work conducted by Schwerdtfeger et al. (2009), modified the AR conditions using infrared marker tracking and their redesigned Pick-by-Vision visualization (see Figure 4.9, left), whereby the picking process was guided by an attentional funnel, first introduced by Biocca et al. (2006). They also changed the HRV setup, using a Polar RS800 CX Multi pulse recorder that was applied the same way as before. This time they only evaluated the SDNN to analyze the physiological strain. Within the follow-up study eight participants performed order picking tasks, also for two hours, comparing paper and AR guidance. The results (see Figure 4.9, right) revealed that the development of the SDNN through the test phases did not show significant differences. They also evaluated the subjective strain as was done by Tümler et al. (2008), and also did not reveal differences between the guidance methods. This time the subjective experiences were supplemented by analyzing the task load (NASA TLX), which resulted in a higher workload using the AR system with an average rating of 87.81 compared to the paper condition (71.98), but did not revealed significant differences. Besides strain, performance was evaluated by the picking time and error rate. The results showed that using the AR system caused a higher error rate, but enabled a slightly faster performance (7.6%), although not reporting any significance tests. Among other issues, like problems using the HMD, the authors discussed the results related to the experience levels of the participants, showing that real picking workers performed slower than novices. Perhaps this may be explained by the worked example effect derived from the cognitive load theory (see section 3.3.4), which states that that learners using a worked example for the first time perform better on problem solving than learners who are experienced with the given problem.

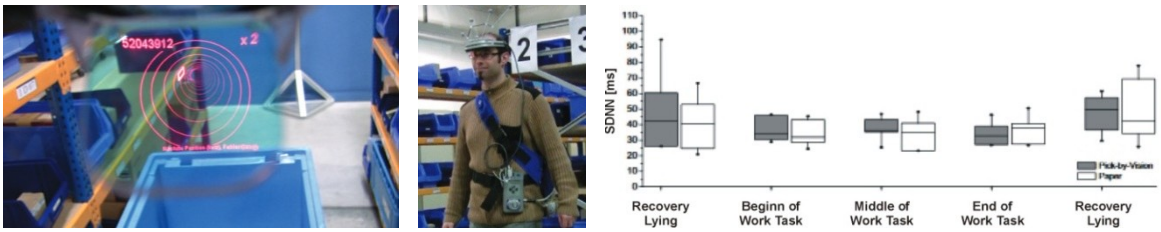


Figure 4.9: Pick-by-Vision (from Schwerdtfeger et al., 2009). *(left)* Tunnel visualization for AR picking. *(middle)* The AR equipment. *(right)* Development of the HRV SDNN through the test phases.

During a following study conducted by Grubert et al. (2010), nineteen participants performed picking tasks, but this time for four hours, whereby the test setup and procedure, as well as the target visualization using the attentional funnel metaphor (see Figure 4.10, left and middle) were similar to Schwerdtfeger et al. (2009). To evaluate the physical strain, they also only analyzed the SDNN parameter of the HRV. The subjective strain was assessed in the same way as was done by Tümler et al. (2008). They additionally applied a secondary task, requesting simultaneously responding to an auditory stimuli, but this had only to be executed in a pre-test and post-test phase and related data were only collected in a subjective way obtained with a questionnaire assessing strain indicators.

During the test, they additionally quantified the performance by the overall completion time and type errors. Besides a significant improvement of the performance by the AR system (see Figure 4.10, right), the results did not show any effects of long-term operation with AR or paper manual on the physical strain, neither in HRV nor by secondary task performance. The authors reasoned the better AR performance with system optimizations, especially by using the attention funnel, although such technique was also used by Schwerdtfeger et al. (2009), not causing improvements there.

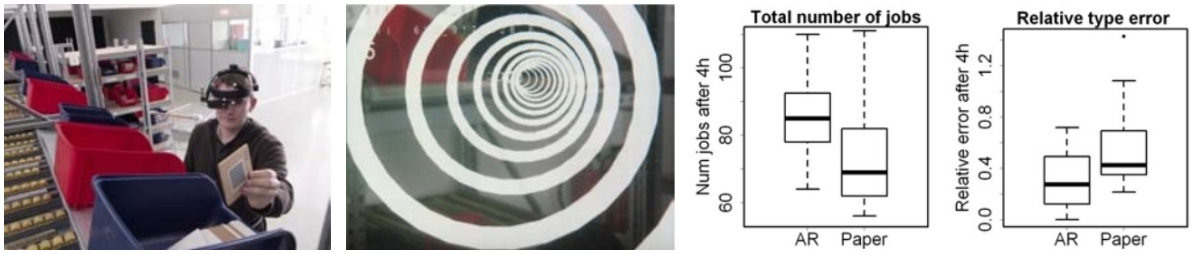


Figure 4.10: Extended study for order picking (from Grubert et al., 2010). *(left)* Participant performs picking task. *(middle)* The attention funnel navigation. *(right)* AR yielded more picked items and less type error.

Also before Azuma defined AR in a distinct manner in 1997, the members of the *Computer Graphics and User Interface Lab*[45] (CGUI) of the Columbia University (NY, U.S.), led by Steve Feiner, had already started their research on AR supported tasks and have contributed to issues on instructional AR several times until today. Already in 1993 and besides their work on *"Windows on the World"* exploring 2D windows for 3D Augmented Reality (Feiner et al., 1993b), they launched their first work towards instructional AR in the scope of the project for *Knowledge-based Augmented Reality for Maintenance Assistance* (KARMA), which aimed to design an AR guided assistance system to support maintenance tasks for a laser printer (Feiner et al., 1993a). Their system consisted of a see-through HMD and several Logitech 3D trackers attached to the key components of the printer, enabling an overlaid view of 3D-registered graphics and simple textual callouts (see Figure 4.11). They main contribution was an interactive rule-based management system, which, for example, had to autonomously decide whether an object is needed to be rendered depending on its visibility in the real world.

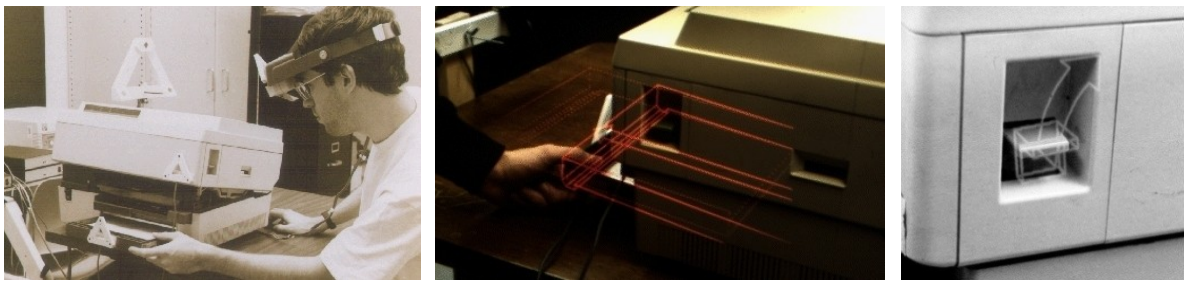


Figure 4.11: Knowledge-based AR for maintenance at a laser printer (from Feiner et al., 1993a). *(left)* The user wears a see-through HMD, while the laser printer is equipped with Logitech 3D trackers. *(middle)* The virtual world shows the user how to remove the paper tray. *(right)* Virtual instructions to pull up the lid's lever.

The most prominent work of the CGUI group on instructional AR was done in the scope of the ARMAR[46] project, which started in 2007 sponsored by the Air Force Research Laboratory. This project aimed to investigate the benefits of Augmented Reality to support maintenance and repair

[45] Computer Graphics and User Interface Lab (CGUI). URL: http://graphics.cs.columbia.edu, last visit: 28.09.2016
[46] ARMAR Project. URL: http://graphics.cs.columbia.edu/project/armar, last visit: 28.09.2016

tasks inside an armored vehicle turret. Because this work has mostly influenced the modus operandi of research on instructional AR presented in this thesis, it approaches will be discussed in detail. The main purpose of this research was to design and develop an AR prototype to support maintenance tasks. It's feasibility was evaluated with regard to the question "[...] how real time computer graphics, overlaid on and registered with the actual equipment being maintained, can significantly increase the productivity of maintenance personnel, both during training and in the field" (Henderson & Feiner, 2007, p. i). Therefore a prototype was implemented and tested aiming at support routine maintenance tasks inside an armored vehicle turret normally needed to be conducted by military mechanics (Henderson & Feiner, 2009, 2011a). The prototyped system (see Figure 4.12) consisted of a video see-through (VST) HMD (Headplay gamin display) equipped with two Point Grey Firefly MV cameras (640x480 resolution), an outside-inside tracking system (NaturalPoint OptiTrack, 10 cameras) and a wireless wrist controller for interaction purposes, which was an Android G1 phone using 2D controls operated by the touch screen. This setup revealed two weak points. Once, the tracking system was not suitable for practical field testing and, on the other hand, the controller required an additionally device while wearing an HMD. It should be noted that they have tackled the last mentioned problem by introducing their Opportunistic Controls (Henderson & Feiner, 2008, 2010) that will be discussed in section 4.1.3.

Figure 4.12: AR maintenance for an armored personnel carrier turret (Feiner & Henderson, 2009). *(left)* An application domain expert performs a maintenance task guided by AR. *(middle)* A view through the HMD. *(right)* The wrist worn controller used for showing the next instruction in a maintenance sequence.

The prototype was tested in a controlled field study conducted by application domain experts (military mechanics). In a within-subject design, three display conditions were studied by representative maintenance tasks, including the conventional guidance method adopted by an external LCD and used as baseline condition that was compared with the AR prototype and an HUD condition, which complied with a none 3D-registered version of the AR system. All three conditions were operated by the wrist controller. Besides performance measures, they also evaluated head movements and subjective experiences. The findings showed that the overall completion time was not affected by the guidance conditions (see Figure 4.13, left), but analyzing the task localization time revealed a significant faster localizing with AR, while using the HUD led to the significant slowest (see Figure 4.13, middle). The overall slow completion times caused by the AR and HUD conditions were reasoned with post-localization processes related to the workpiece portion of a task (Neumann & Majoros, 1998), whereby the HMD did not contribute with superimposed information, but delimitated the view during task executions. Evaluating the translational head movements showed that AR and HUD led to significantly fewer movements than LCD, while the HUD condition caused the significant lowest velocities compared to AR and LCD (see Figure 4.13, right). It was noted, that the AR condition required smaller ranges of head movements, which was emphasized as an outstanding ability of AR "[...] to potentially reduce overall musculoskeletal workload and strain

related to head movements during maintenance tasks" (Henderson & Feiner, 2009, p. 143) and, thus, assessing the HRV was proposed to verify strain reductions accomplished by decreased head and neck movements coupled with the effort of wearing a HMD. The qualitative results revealed that the ease of use was criticized by a lack of comfort in wearing the HMD and that the AR condition was rated equally in satisfaction and intuitiveness as the conventional method. Overall this study showed that AR was beneficial during task localization, also called the informational phase, but could not contribute in overall completion and qualitative aspects.

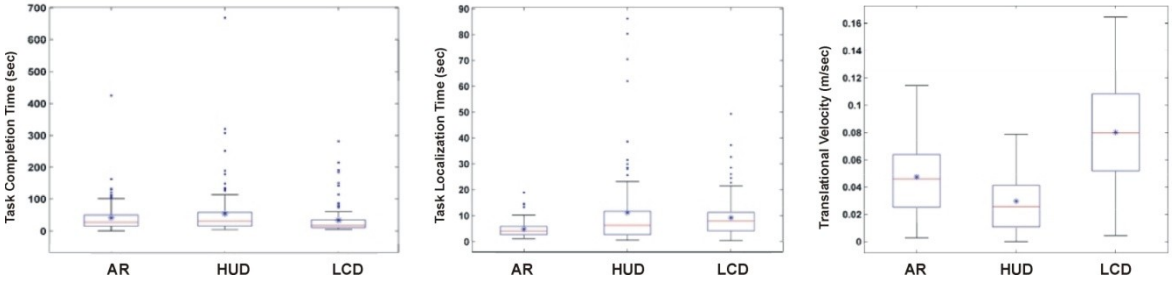


Figure 4.13: The results of the ARMA usability study, comparing three guidance methods (from Feiner & Henderson, 2009). *(left)* Task completion time. *(middle)* Localizing time. *(right)* Velocity of head movements.

In a subsequent work, Henderson & Feiner (2011b) were focused on the stage during the workpiece portion, also referred to as the psychomotor phase, which resulted in slower performances while wearing a HMD. This time they studied instructional AR applied to a Rolls-Royce Dart 510 turboprop engine and contributed with AR assistance during the psychomotor phase providing dynamic visualizations, like 3D arrows, billboarded labels and 3D highlights (see Figure 4.14, *a*), that has complemented the AR assistance during the informational phase adopted from their previous work (Henderson & Feiner, 2009). The research contribution of this work was reflected by validating their approach in a controlled laboratory user study conducted from participants mainly acquired from the department of computer science, who were not familiar with the engine. By performing an assembling task, the AR condition was compared with a control condition using an LCD device installed in front of the working area (see Figure 4.14, *b*). The study task required the assembling of a combustion chamber divided into two locations, three position and one alignment task. The test apparatus consisted of an optical-see-through (OST) HMD (NVIS nVisor ST60, 60° diagonal FO, 1280x1024 resolution), and two kinds of optical tracking (11 NaturalPoint OptiTrack infrared cameras, marker-based VTT ALVAR[47] tracking). The results showed that the AR condition led to

Figure 4.14: The ARMA project - Instructional AR for the psychomotor phase (from Feiner & Henderson, 2011b). *(a)* 3D-registered AR data overlaid on components of an aircraft engine. *(b)* Participant performs with the LCD condition. *(c)* The completion times of the different activities for the AR and LCD condition.

[47] VTT ALVAR tracking library. http://virtual.vtt.fi/virtual/proj2/multimedia/alvar, last visit: 28.09.2016

significantly fastest performance during the alignment activity, while the location and position activities did not differ from the LCD condition in completion times (see Figure 4.14, *c*). The accuracy was only measured for the alignment activity and revealed that AR led to significantly most accurate performance. Keeping in mind that the usability test was conducted under laboratory conditions, this study contributed by proving that alignment activities during such assembling tasks are benefiting significantly from AR support in completion time and accuracy. As in the previous study, there was no assessment of mental or physiological workload, which could be relevant for long-term assembling.

Summary: So far as presented in this section, the application of AR can vary in its contribution to assist instructional tasks, showing that the task requirements and environmental conditions are crucial factors to adopt AR for a particular problem. It is not always clear whether an instructional task has significantly benefitted from AR, and if so, which criteria were responsible for a profitable application. Besides technical issues (e.g., display, 3D pose tracking), such criteria could reflect the *type of subtasks* (e.g. in a simple way considered for furniture assembly by Zauner et al., 2003), the *level of task difficulty* (Wiedenmeier et al., 2003; Mura et al., 2012; Syberfeldt et al., 2015), the *task duration* (Tümler et al., 2008; Schwerdtfeger et al., 2009; Grubert et al., 2010), as well as the *amount of body and head movements* (Henderson & Feiner, 2009, 2011a). So, for example, Henderson & Feiner evaluated head movements and concluded that AR can reduce head and neck movements during repair tasks (Henderson & Feiner, 2009, 2011a). However, in cases of developing generic approaches (e.g., Schwald & Laval, 2003; Pang et al., 2006), such criteria should be identified and classified, which in turn could be reflected in individual strategies for suitable AR support. On the other hand the question arises whether the effort of generic approaching is reasonable, or if it would be more efficient to consider one certain task problem. Besides using toy blocks or puzzles as alternative use cases for assembly tasks (Tang et al., 2003; Robertson et al., 2008; Salonen & Sääski, 2008; Mura et al., 2012; Khuong et al., 2014; Syberfeldt et al., 2015), only few works were done to investigate the effect of AR support on a practical real-world problem in ratio to the amount of challenges (Caudell & Mizell, 1992; Feiner et al., 1993a; Ockerman & Pritchett, 1998; Reiners et al., 1999; Zauner et al., 2003; Platonov et al., 2006, Ishii, et al., 2007; Webel et al., 2011a; Gay-Bellile et al., 2012; Espindola et al., 2013) and performed sufficient usability testing considering quantitative and/or qualitative indicators (Curtis et al., 1999; Boud et al., 1999; Baird & Barfield, 1999; Wiedenmaier et al. 2003; Henderson & Feiner, 2011b). In an optimal case the physical task environment is considered either by using elaborated recreated reference scenarios (Tümler et al., 2008; Schwerdtfeger et al, 2009; Grubert et al., 2010), or ideally by controlled field testing (De Crescenzio et al., 2011; Henderson & Feiner, 2009, 2011a, Webel et al., 2011b). This will be of particular importance for the research presented in this thesis, evaluating the benefit of AR assistance for instructional ISPR tasks.

Although there is an apparent increase in research on user-related aspects of AR systems (Swan & Gabbard, 2005; Bai & Blackwell, 2012), evaluating the usefulness of AR relies mainly on analyzing performance indicators, such as completion time, error rate and accuracy. But maintenance, assembling or servicing tasks may take long times and may imply stressful environments. Because this carries the risk of losing or tunneling attention caused by fatigue or an increased level of arousal, analyzing the workload will be equally important[48]. Only few of the presented studies have

[48] "Designers and operators realize that performance is not all that matters in the design of a good system. It is just as important to consider what demand a task imposes on the operator's limited resources" (Wickens & Hollands, 2000, p.459)

considered this, but mainly subjectively assessed by the NASA TLX (Tang et al., 2003; Robertson et al., 2008; Schwerdtfeger et al., 2009; De Crescenzio et al., 2011). However, few other works were primarily focused on investigating the effect of long-term operations on users' strain, evaluated by biofeedback using HRV (Tümler et al., 2008; Schwerdtfeger et al.; 2009; Grubert et al.; 2010). Finally it should be mentioned, that in an ideal case the workload is assessed by multi-dimensional indicators (see section 3.4.3), as was done by Schwerdtfeger et al. (2009), evaluating the strain physiologically by HRV and subjectively by the NASA TLX. Also Grubert et al. (2010) assessed the workload two-folded (HRV and secondary task), but applied the dual-task metaphor only during pre- and post-test-phases.

4.1.2 Presentation of Target Cues for AR Task Localization

When carrying out instructional tasks, a crucial aspect of human performance is related to the process of visual search, which optimally ends with the detection of objects in question (Wickens et al., 2003, p. 60). In Augmented Reality the visual search process mainly complies with the localization of off-screen targets. Once the target area is in the field of view, it is required to confirm or detect the target of interest. At this point, task localization should be defined as a sequential process comprising the initial visual search of an off-screen target resulting in its subsequent detection. While the previous section reviewed works, introducing AR applied to the instructional task domain, this section looks more at research on issues related to task localization, which generally aims to improve the visual attention control by investigating modalities to present cues for the off-screen search and to tag targets.

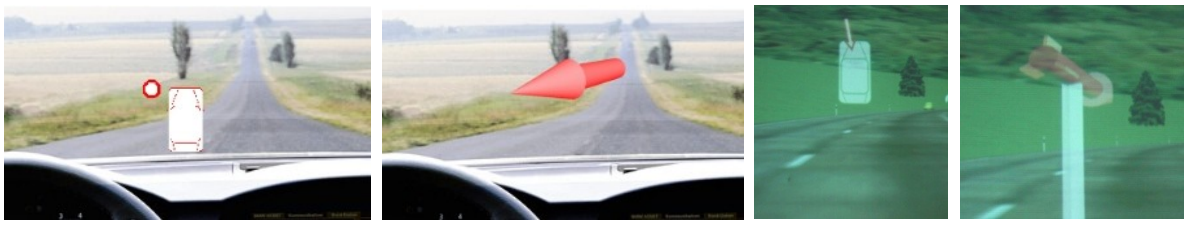
Directing users' attention towards the position of off-screen target objects is generally addressed towards the question how efficiently users' navigation can be supported by visual cues. This is closely linked to users' ability to process spatial information about their environment and their orientation within it. In general, spatial representations of environments can be classified by the used reference frame in which the objects are coded, whereby their spatial information can be represented by an egocentric or exocentric reference frame. Target cues located in the egocentric frame of reference are coded in intrinsic coordinates using the body or body parts as point of reference, while exocentric cues are coded in extrinsic world coordinates relative to surrounded objects or landmarks and without reference to their relation to the user. Thus, the spatial orientation of the required information to reach objects not being in the field of view can affect the workload and performance during task localization in AR guided operations. In AR environments, an *egocentric presentation* complies with the fundamental AR characteristic of conformal 3D registration and is coded in user's frame of reference. Thereby the user simply pursues a direction indicator, such as an arrow or tunnel, which starts from the user's view point keeping her/his viewing direction that in turn could preserve the attention of execution. So, for example, Henderson and Feiner (2009) have used but not studied, different kinds of egocentric visualizations (see Figure 4.15) to direct the visual attention during the localization sequence of their AR prototype. Thereby the users were directed by a screen-fixed green arrow towards targets located behind their back, while a tapered red semitransparent 3D arrow navigated the user towards the target if the target was located in the natural peripheral field of vision and gradually faded out when the target was in the display's field of view. In contrast, *exocentric presentations* of target cues are registered anywhere at the screen and detached from user's foveal vision. Such presentations schemes usually use overview maps, in which the target location is indicated. The overview maps are often realized by a bird's eye perspective, but other presentation schemes indicate the target location by, for example, using a compass or radar

metaphor. An extended implementation of the bird's eye map is the concept of the *world in miniature* (WIM) (e.g., Stoakley et al., 1995; Höllerer et al., 2001; Mulloni et al., 2011), which uses a 3D copy of the object of interest or the surrounding environment. Regardless of the used presentation scheme, such target cues always force the user to an exocentric navigation strategy that distracts users focus, but can enable customizable performance in remembering locations (Barfield et al., 1995) or taking short-cuts.

Figure 4.15: The ARMA project - Localizing off-screen targets using a screen-fixed arrow if the target was behind the user *(left)* and a semi-transparent arrow if the target was in the peripheral field of vision *(middle)* that gradual faded out if the target was in the range of the display *(right)*. (from Feiner & Henderson, 2009).

Some works, not considered here, compared exclusively device conditions (e.g., Billinghurst et al., 1998; Wither et al., 2007) showing that visual search tasks generally are benefiting from egocentric displays (e.g., HMD, handheld displays). However, until now, the findings of related research were mixed regarding the question from which reference frame the user can benefit most. Although egocentric visualizations do not require spatial transformations, the question arises whether an egocentric presentation always advantages localization tasks most efficiently. More intended for VR purposes but also relevant for AR, Chittaro and Burigat (2004) introduced an egocentric 3D arrow pointing technique to support the localization of objects inside a large-scale virtual environment (VE). The 3D arrow approach was compared with two conditions using 2D radar and 2D arrows, and a baseline condition providing no navigation aids. The search task that was investigated required the localization of different targets in specific order using two environmental conditions (urban, abstract). Searching in the urban condition was significantly slower with the control condition compared to the other conditions, which in turn did not differ from each other. In contrast, searching in the abstract environment revealed that the 3D arrow condition significantly outperformed all other conditions and was significantly preferred by participants. This study showed that target localizing in an abstract VE benefits from guidance by a 3D arrow, while localizing in an urban VE is not affected by any kind of visual support. Similar work, but in the scope of AR, was done by Tönnis and colleagues (2005) investigating spatial localizing to direct car driver's attention towards the position of imminent danger. Similar to Chittaro and Burigat (2004), they compared visualization schemes using a 2D sketch of the car from an exocentric bird's eye perspective (see Figure 4.16, *a*) and a 3D arrow displayed in the egocentric frame of reference of the driver's field of view (see Figure 4.16, *b*). Using a BMW driving simulator, the usability testing was conducted by twelve participants requested to look as quickly as possible in an indicated direction while driving. The results revealed that the 3D arrow caused significantly slower responses and less reliable localizations, while the bird's eye view significantly led to more lane deviations, which could be caused by the demanding process of spatial transformation. In a continuing work (Tönnis & Klinker, 2006) the visualization schemes were improved. While the localization hint of the exocentric bird's eye view was replaced by a 2D arrow indicating the direction of the upcoming danger (see Figure 4.16, *c*), the 3D arrow was extended by attached fins at the rear-side to dissolve ambiguous

directing during forward or backward pointing (see Figure 4.16, *d*). The same driving simulator was used for usability test comparing the bird's eye perspective and 3D arrow, both without and with additional 3D sound. Contrary to the previous study (Tönnis et al., 2005), the comparison between the presentation schemes revealed significantly faster responses using the 3D arrow under both sound conditions. The driving behavior was similar affected as in the last study, showing an increase in both lane deviation and speed, by using the birds' view.



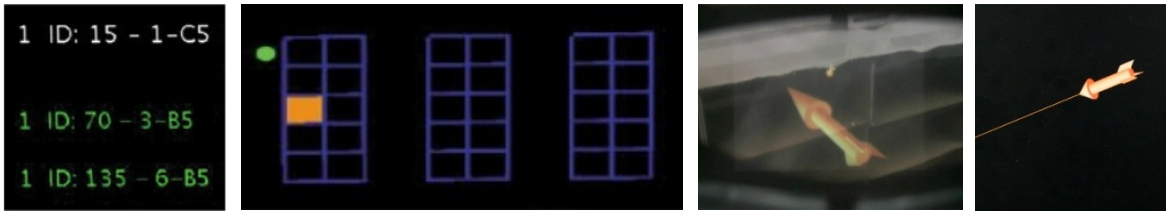
(a) **(b)** **(c)** **(d)**

Figure 4.16: Evaluating AR visualizations for directing car driver's attention. Localization hints investigated by Tönnis et al. (2005): *(a)* 2D bird's eye sketch, *(b)* egocentric 3D arrow. Localization hints investigated by Tönnis & Klinker (2006): *(c)* 2D bird's eye sketch with a 2D arrow, *(d)* egocentric 3D arrow extended by attached fins.

Also intended to support car driver, the work from Medenica et al. (2011) was focused on driver assistance for navigation. Using a common navigation device forces the driver to look away from the road, affecting the driving performance by the absence of visual attention. This work investigated the interrelation between the spatial orientation of the display and the visualization. Therefore three conditions for navigation support were compared (see Figure 4.17), using AR, which provided an egocentric display and visualization, and an exocentric display that differed in the spatial orientation of its visualization, enabling an egocentric street view and an exocentric bird's eye view. The study was conducted with a high fidelity driving simulator, displaying a simulated city surrounding, and performed by eighteen participants that where requested to drive a route with two unexpected events. The visual attention was evaluated by performance measures (dwell time, cross-correlation peaks in viewing away from the road, absolute values of first difference and variances of lane position) and by the subjective workload, assessed by the NASA TLX. The results showed that egocentric AR led to the most visual attention related to the significantly shortest dwell times and the smallest effect size in cross-correlation peaks, compared to both exocentric conditions, whereby the egocentric visualization caused longer dwell times than the bird's eye perspective, but also caused the largest effect size in cross-correlation peaks related to viewing away from the road. The lane position was not affected by the navigation method. Analyzing the NASA TLX yielded least workload with AR, which significantly differed from the egocentric street view, resulting in the highest workload.

Figure 4.17: Car navigation comparing an egocentric AR display and visualization *(left)* with an exocentric display with egocentric visualization *(middle)* and an exocentric bird's eye *(right)* (from Medenica et al., 2011).

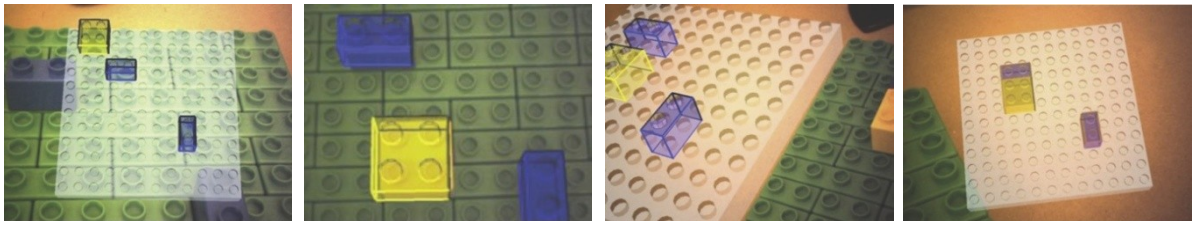
The same 3D arrow visualization used to direct car driver's attention (Tönnis & Klinker, 2006), was applied by Schwerdtfeger et al. (2006) for investigating visualization schemes to support wayfinding for order picking tasks across several types of displays. Thus, the related study design consists of two independent variables (visualization scheme and display device) with three levels respectively. As shown in Figure 4.18, the levels of the visualization scheme were assigned to a 1D textual list, a 2D *you-are-here* strip-map and the 3D arrow, which was extended by a rubber band, being a dynamic line connecting the arrow's tip with the target in question as long as it was not in the field of view. Using a within-subject-design and 18 participants, the three visualization conditions were compared for one display device at a time (remote screen, PDA, HMD). The results did not show any significant difference in wayfinding times between 1D and 2D for neither device, but revealed significant slower performance using the 3D arrow compared to the others. However, analyzing only the picking times revealed no differences at all, but the error rate showed that the AR-based visualizations caused up to 10 times more mistakes than with lists and maps. The authors assumed that this was a problem with depth perception, because the items were often picked too high or too low.



(a) **(b)** **(c)** **(d)**

Figure 4.18: Comparing visualizations to support order picking tasks (from Schwerdtfeger et al., 2006): *(a)* a 1D textual list, *(b)* a 2D *you-are-here* strip-map, *(c)* a 3D arrow and *(d)* a 3D arrow with a rubber band.

More relevant work related to instructional AR guidance was conducted by Robertson et al. (2008) and Khuong et al. (2014) evaluating the impact of situating localization assistances during LEGO assembly. Therefore Robertson and colleagues (2008) studied four conditions for target localizing, comparing in-view, in-situ and off-screen support (see Figure 4.19). Thereby the in-view support complied with an unregistered WIM interface, which was centrally projected in the user's field of view, always partially blocking the field of view, while the in-situ support corresponded to 3D-registered AR overlays. Contrary, the off-screen conditions corresponded to WIM interfaces placed outside of user's field of view and differed in their degree of registration, whereby a virtual base plate was positioned parallel to the side of the physical base plate or an unregistered virtual plate, appearing when the head orientation exceeded 30 degrees. Results showed that LEGO placement



(a) **(b)** **(c)** **(d)**

Figure 4.19: The localizing conditions for LEGO assembly evaluated by Robertson et al. (2008): *(a)* in-view unregistered WIM, *(b)* in-situ 3D-registered AR, *(c)* off-screen virtual plate parallel positioned to the left side of the physical base plate, *(d)* off-screen virtual plate displayed at least 30 degrees changes of head orientation.

tasks benefited from the in-situ 3D-registered AR overlays with respect to speed, errors and subjective cognitive load assessed by the NASA TLX. Although always hindering the view, results also revealed that the in-view unregistered WIM interface was not advantageous over both off-screen conditions. This suggests that the needed head movements may have disadvantaged the assembly of LEGO blocks. Also Khuong and colleagues (2014) used LEGO assembly, but studied task localization by investigating the effect of two in-situ AR visualization conditions on context-aware assembly support (see Figure 4.20). Thereby the assembly was compared between guidance using conformal 3D-registered AR, which was overlaid with a partial wireframe, and guidance provided by hints rendered on a dynamic virtual model, positioned next to the real model. It was hypothesized that the side-by-side mode will decrease the completion time compared to the AR mode. In addition, two modes were provided to evaluate the effectiveness of error detection, hypothesizing that assembling without error detection will decrease the performance in completion time, but will also cause a higher error rate compared to the performance with support of error detection. The results of the within-subject user study conducted by 24 participants, revealed a significantly faster assembly using the side-by-side mode in general, regardless of the error detection condition (see Figure 4.20, right). While assembling with error detection support significantly increased the completion for the side-by-side mode, assembling using the AR was not affected by the mode of error detection. With respect to the mean number of errors, the side-by-side model produced significantly fewer errors than AR in general. Qualitative results related to self-reported experiences revealed a significant preference for the side-by-side mode.

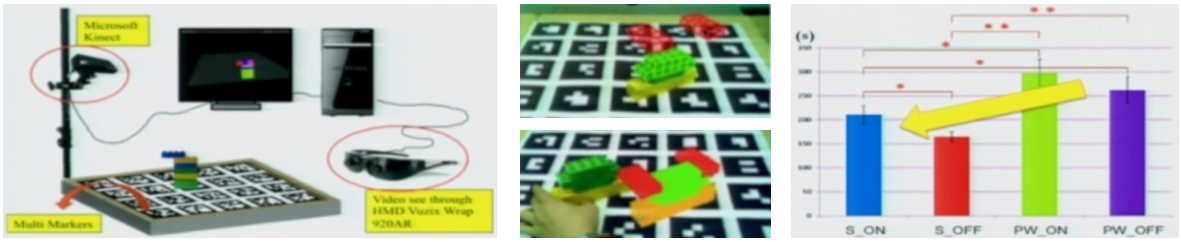


Figure 4.20: Evaluation of AR-based context-aware support for LEGO assembly (from Khuong et al., 2014). *(left)* AR setup using marker tracking and VST HMD. *(middle)* The visualization conditions: *top* - the partial overlaid AR mode (PW); *bottom* - the side-by-side mode (S). *(right)* Task completion time per visualization mode, showing significant faster performance for the side-by-side mode.

Without a specific application, but rather intended for search purposes in the immediate environment, Biocca et al. (2006) studied an egocentric presentation scheme to guide the visual attention in AR. They introduced their omnidirectional attentional funnel (see Figure 4.21) addressed to the enhancement of the attention management during handling mobile AR interfaces for physical and virtual object selection, visual search in a cluttered space or for navigation purposes in near and far space. Thereby the attentional funnel uses strong perspective cues to directly connect the head-centered coordinates of the egocentric AR view to the coordinate space of the object by a dynamic virtual funnel, which enables to tunnel attention. Along a curved path, it is fragmented in holed planes tapered towards the object. Usability testing revealed that the attentional funnel tended to increase the performance at decreased subjective workload (assessed by NASA TLX) during visual search and retrieval of objects compared to conditions using visual highlighting and audio instructions, although visual highlighting exclusively complies with the process of detecting an object and does not support the visual search process. Thus it did not provide an equivalent condition for the comparison with audio instructions and the attentional funnel.

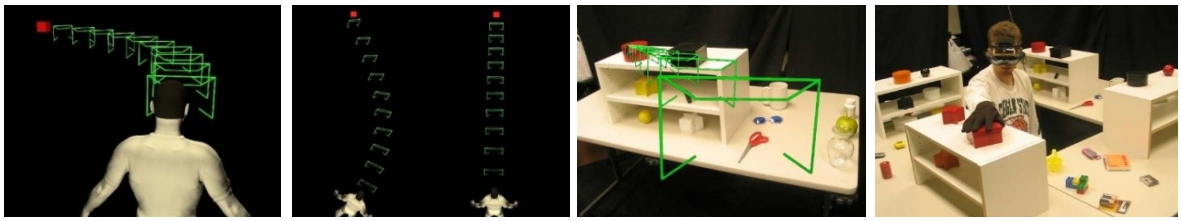


Figure 4.21: The Attention Funnel and the environment during usability testing (from Biocca et al. (2006).

The attention tunnel metaphor was also studied by Schwerdtfeger and Klinker (2008). In this work they studied the impact of target visualization on order picking tasks in the scope of AR. To investigate the localization performance, different methods for the picking visualization were introduced and compared, like using a 3D arrow, a highlighted frame and the tunnel technique (see Figure 4.22). While the tunnel technique already combines both visual search and detection in ones, the arrow and the frame only support the detection of an object. Hence, both techniques were combined with a rubber band technique providing navigation for the initial visual search process, as was already used in a prior work (Schwerdtfeger et al., 2006). Once the target was in user's field of view, the rubber band was replaced with the studied visualization (frame, arrow). The results of the comparison showed that the picking task was timely affected by the visualization, but only revealed significantly faster localization using the highlighted frame compared to the arrow, while the tunnel visualization tended to be equal in time as the arrow. Analyzing the frequency of picking errors also revealed a main effect for the target visualizations, whereby the 3D arrow caused significantly the most errors. Contrary to the findings from Biocca et al. (2006), this study has shown that using a 3D arrow will be improper for order picking tasks, while a highlighted frame seems to be more suitable. Of course, in this case it would have been useful to distinguish more clearly between visual search and detection, but the study design lacked such observations. The used Wizard-of-Oz technique did not allow to make distinctions between the search and detection process.

Figure 4.22: Evaluating picking visualizations by Schwerdtfeger and Klinker (2008). *(left)* The 3D arrow in front of a box. *(middle)* The 3D frame highlighting a box. *(right)* The tunnel visualization.

Summary: The works presented in this section investigated issues on task localization in different ways. Some works studied issues on spatial situating of localization hints related to their position, orientation or the degree of dimensionality. Works comparing egocentric directing using 3D arrow techniques with exocentric 2D presentation schemes, like arrows, radar or bird's eye view (Chittaro & Burigat, 2004; Tönnis et al., 2005, 2006), revealed unambiguous results. While Chittaro & Burigat (2004) have shown that a 3D arrow technique can advantage target localization in a VE, but only in an abstract environment, works by Tönnis et al. have shown that directing car driver's attention can benefit either from exocentric 2D bird's eye view (Tönnis et al., 2005) or from egocentric 3D pointing (Tönnis & Klinker, 2006). In contrast, Medenica et al. (2011) have shown that visual attention during car driving while way finding, could best be kept up using egocentric AR navigation aids compared to exocentric displays, whereby an exocentric bird's eye view astonishingly supported

visual attention more than the egocentric street view, which we find actually in our car navigation systems as the predominant metaphor. However, relevant works investigating 3D presentation schemes for target localization during assembly tasks have shown that AR overlays cause better performance than in-situ WIM or off-screen interfaces (Robertson et al., 2008). The consideration of in-situ locations only has shown that a 3D visualization located side-by-side can outperform AR overlays (Khuong et al., 2014). In conclusion, it can be stated that task localization benefits predominantly from egocentric guidance, but is not generally valid as shown by Tönnis et al. (2005) and Khuong et al. (2014). This suggests that the localization performance can depend on the task, which needs to be fulfilled, or the quality of the visualization resources. Other works were focused on introducing and studying 3D-registered visualization techniques to feature the position of the targets that needed to be localized. Therefore Biocca et al. (2006) shown that visuospatial searching can benefit more from their introduced attentional funnel visualization than from highlighting, while, in contrast, Schwerdtfeger and Klinker (2008) have demonstrated that using a highlighted frame is more suitable for order picking tasks compared to a tunnel visualization, as well as to a 3D arrow technique. However, looking more closely at the egocentric nature of the attentional tunnel metaphor revealed that the works studying such attentional tunnel for task localization (Biocca et al., 2006, Schwerdtfeger & Klinker, 2008) or only applying it for logistic picking tasks, as presented in last section (Schwerdtfeger et al., 2009; Grubert et al. 2010), lacked adequate prior heuristic design evaluation. This would have enabled the consideration of fundamentals and theories from cognitive psychology, such as perceptual narrowing causing the tunnel effect (see section 3.3.2), although at least finally discussed by Grubert et al. (2010) referencing the work from Wickens & Alexander (2009), which investigated display-induced attentional tunneling, such as in head-up displays. Instead of comparing different picking methods only in picking time, number of picked items or error rates, it would have been useful to investigate the resulted level of arousal and workload in addition. This could be done, for example, by detecting unattended or peripheral stimuli using the secondary task technique (see section 3.4.3). Although Grubert et al. (2010) applied such dual-task metaphor, they only collected related data in a pre-test and post-test phase and only subjectively with a questionnaire assessing strain indicators.

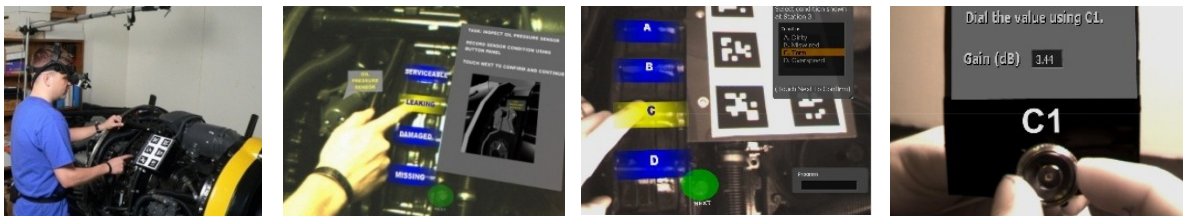
As was already discussed in the last subsection, assessing the workload can contribute to the knowledge about the attentional level and control, which, for example, would be of vital importance for studying the last discussed tunnel effect. Considering the studies presented in this subsection so far, only a few works investigated the effect of target cueing on the workload level. Thereby, works that have done this, assessed the mental workload by subjective ratings using the NASA TLX (Biocca et al., 2006; Robertson et al., 2008; Mulloni et al., 2011; Medenica et al., 2011). There was no study assessing the workload by physiological parameter or by the effort spent on dual-task processing, although the latter one was applied from Kishishita et al. (2014), but analyzing the effect of wide field of view displays on the visual perception of augmented objects for view management and not studying presentation schemes for task localization. In general, these two indicators are meaningful to measure stress yielded by inadequate target cueing, causing for example spatial ambiguity, which can lead to confusion or frustration that in turn can increase the level of arousal.

4.1.3 Placement of Interactive AR Control Interfaces

Triggering actions, while working with laptop or desktop computers, predominantly implies the usage of classical input devices (e.g., keyboard, touchpad, or mouse). However, in the scope of AR assisted instructional operations, the associated control interface has to provide interaction modalities

for system-control or symbolic-input tasks in a reasonable manner so that the operator is able to communicate efficiently with the AR system. In an ideal case, such interfaces are quickly accessible and require no particular effort to handle devices out of reach or sight, drawing off the focused attention. In turn, this states that the fundamental criterion for interactive control interfaces is the placement condition needed to be met. While the previous sections were focused on issues related to the viewing of content provided by AR, this section reviews and discusses works related to issues on the spatial placement of interfaces needed to operate AR control interfaces, but not considering issues regarding the design of control items nor the modality to interact with them, assuming that such interfaces are operated by direct touch input (Buxton et al., 1985) using the hand or fingers as most natural input device to efficiently perform selection or manipulation tasks.

As mentioned in section 4.1.1, the setup of the AR system developed by Henderson and Feiner (2009), included a wireless wrist controller for interaction purposes, which requires an additionally device while wearing a HMD (see Figure 4.12, right). This will not be efficient in AR environments. One of their works introduced the Opportunistic Controls (see Figure 4.23) as class of interaction techniques to avoid repositioning the hands, or distracted head movements (Henderson & Feiner, 2008, 2010). Thereby button-based controls were positioned on convenient physical surfaces in the AR working area, while movable controls were using physical objects to adjust parameter ranges. Thus they can be considered as tangible user interfaces (Fitzmaurice et al., 1995; Ishii & Ullmer, 1997; Poupyrev et al., 2001), which enable a "[...] direct control of virtual objects through physical handles [...] making virtual objects physically graspable" (Ishii & Ullmer, 1997, p. 3). Due to their natural handling, physical objects always provide tactile sensations, it is assumed that tangible interfaces are the most intuitive way of interacting in AR. Henderson and Feiner (2008) could show that button-based controls revealed a benefit in time and acceptance than handling virtual buttons on a single surface, installed above the task area. Within a follow-up study they investigated the most efficient control setting, complying with the task-related requirements (Henderson & Feiner, 2010). They evaluated a broad range of potential surfaces and objects used to apply their button-based and movable controls. The results showed that button-based controls were preferred (67 %) over movable controls (33 %) and that surfaces at eye level and located in reachable distance by arm's length were favored.



(a) **(b)** **(c)** **(d)**

Figure 4.23: The Opportunistic Controls for instructional Augmented Reality (from Feiner & Henderson, 2008). *(a,b,c)* Button-based controls projected in the area of the task domain, while receiving tactile feedback from the underlying physical surface. *(right)* A movable control to adjust a parameter range.

While the Opportunistic Controls, and other virtual or ad hoc controls (Weimer & Ganapathy, 1989; Fails & Olsen, 2002; Porter et al., 2010; Marner et al., 2014) are resulting in interfaces spatially referenced to an object in the physical world, interfaces can also be referenced to the head or body. Similar reference types were already classified for a 3D AR surrounding by Feiner et al. (1993b) iden-tifying the placement conditions of 2D windows, which can be surround-, display- or

world-fixed. Works investigating projected interfaces for control or symbolic input tasks mainly assume a world-referenced physical surface to present, for example, a virtual keyboard (Kölsch & Turk, 2002; Roeber et al., 2003; Tomasi et al., 2003). The interaction with world-referenced interfaces are coded in extrinsic world coordinates and thus requires an additional transformation during spatial orientation processing, while head or body referenced interfaces provoke an interaction coded in the intrinsic frame of reference, which does not require any spatial coordinate transformation. The interaction with body-referenced interfaces (attached to limbs), can also benefit from the proprioceptive sense, which can promote direct manipulations by an enhanced sense of object locations, physical mnemonics to find and select control items, or gestural actions to trigger control activities (Mine et. al., 1997). Works considering body-fixed interfaces have used areas of the body for interaction purposes (Cho et al., 2002; Kim et al., 2012; Weigel et al., 2014) or the skin as projection surface (Harrison et al., 2010, 2011; Harrison & Faste, 2014), whereby arms and hands are preferred. For example, Harrison and colleagues (2010) used the body as input surface named Skinput, a wearable, bio-acoustic sensing array built into an armband coupled with a pico-projector, which if attached to the upper arm enables the display of interactive control items on the skin of the forearm. Other work by Harrison et al. (2011) introduced a wearable projection system, referred to as the OmniTouch system, that can be installed, for example, at the shoulder, enabling the projection of control or input interfaces on body parts, physical hand-held surfaces or surfaces located in the external world reference frame (see Figure 4.24, top). The projection system was coupled with a depth camera for finger tracking providing interactive, multitouch input on arbitrary surfaces. The OmniTouch system was evaluated in respect of the accuracy of aimed pointing movements (finger touch selection) by comparing different placement modalities (see Figure 4.24, bottom). Thereby the same interface was projected onto the hand, the arm, a hand-held panel, and the wall in front of the user. Related to the focus of this section, these modalities constitute three placement conditions coded within the human internal coordinate system (hand, arm, hand-held), while the wall projection resulted in a world-referenced interface. Contrary to authors' expectations, the results showed that the wall-projected interfaces provided most precise pointing, while pointing on the arm yielded most imprecise pointing. Pointing towards the palm of the hand and the hand-held panel did not really differ, because both were referenced to the user's hand.

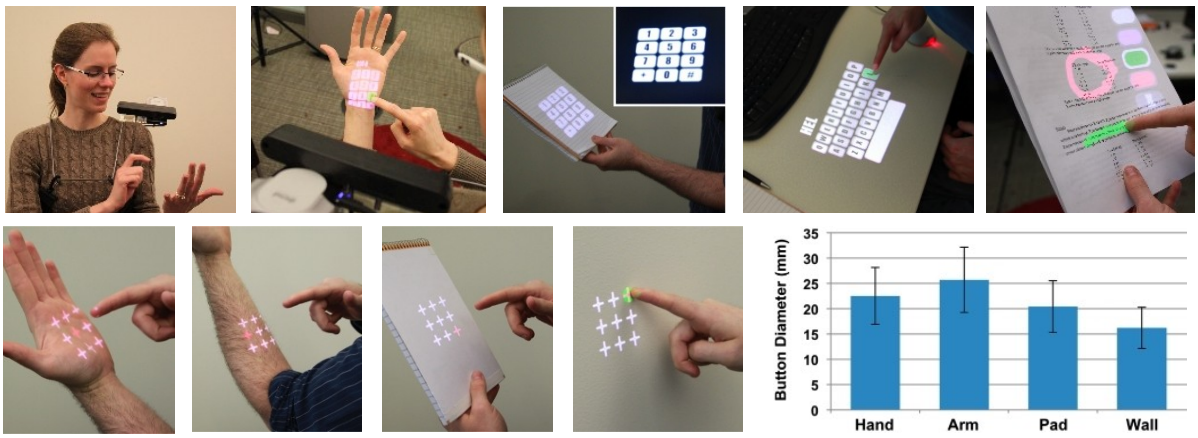


Figure 4.24: The OmniTouch interface (from Harrison et al., 2011). *(top)* The projection system and its applications. *(bottom)* Comparing the accuracy by the button diameter needed to encompass 95% of touches.

Not intended for AR usage but also using the body as a reference to place interactive interfaces, Bowman and Wingrave (2001) designed and evaluated their hand-based TULIP menu (Three-Up, Labels in Palm) aimed at system control tasks in immersive VEs. The TULIP menu (see Figure 4.25,

left) is a menu technique using Pinch Gloves™ whereby the menu items are virtually attached to the fingers or palm, each assigned to a certain menu level. The action is triggered by pinches of the related hand's thumb, while both hands are not interacting with each other. Within a user study, the TULIP menu was compared with two other system control techniques (see Figure 4.25, middle) using a head-referenced floating menu not providing tactile sensation and another hand-referenced interfaces provided by a hand-held physical surface. Although all three interface conditions were coded in the human internal coordinate system, the interaction with the floating menu was disadvantaged by the absence of tactile sensation, while the interaction with the physical surfaces was handicapped by holding the surface with the non-dominant hand hampering hands-free operations. The results (see Figure 4.25, right) revealed that using the hand-held tablet enables significantly fastest performance and fewest selection errors, while handling the TULIP menu causes slowest selection times, which was explained by the selection costs resulting from traversing different levels of the menu hierarchy. Initial difficulties in handling the interfaces could be dissolved with the increase of trials.

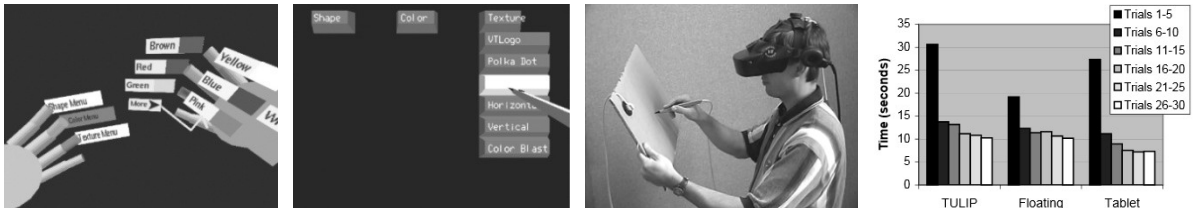


Figure 4.25: The TULIP menu for immersive VEs (from Bowman & Wingrave, 2001). *(from left to right)* The TULIP menu; the head-referenced floating menu; the hand-held menu; and the results of the usability study.

Similar work, but for AR surroundings, was done by Piekarski and Thomas (2003) introducing their ThumbsUp interface technique for mobile outdoor AR (see Figure 4.26), which also uses a tracked set of pinch gloves providing the selection of system control items and 3D manipulation in one single device, whereby the operations are defined by natural movements of the head and hands. Thereby the control items are hierarchically arranged and handled by mapped pinch gestures, while 3D manipulations are performed by the thumbs relative to the head position. Like the TULIP menu, each finger is assigned to a menu option that can be triggered by pinching the particular finger with the thumb.

Figure 4.26: The ThumbUp user interface for mobile outdoor AR from Piekarski and Thomas (2003).

The work of Prätorius and colleagues (2014) presents the DigiTap prototype, a wrist-worn device for symbolic input in VR and AR surroundings (see Figure 4.27). The device consists of an accelerometer and a camera with an LED flash to detect whether and where the user has tapped. The tapping locations are assigned to the fingertips and the finger joints, whereby symbolic input is

performed by tapping with the thumb on it. This enables, for example, to trigger different actions during a point-and-click interaction or grasping a color panel for symbolic input to set the attribute values. Such interfaces are operated with the thumb of the same hand where the interface is assigned to. This enables unimanual handling, which supports hands-free operations in a better way.

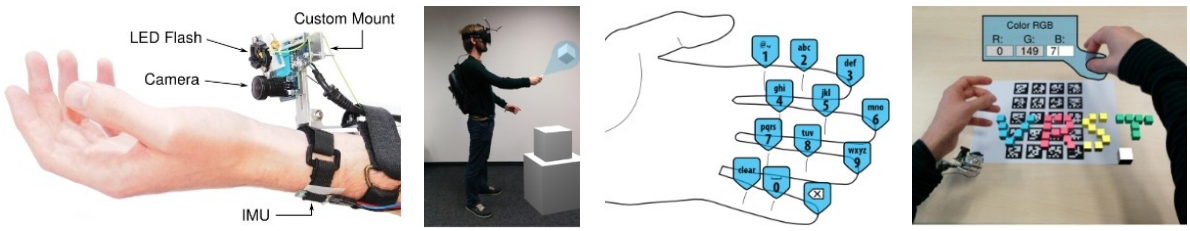


Figure 4.27: The DigiTap input device (from Prätorius et al., 2014). *(from left to right)* The DigiTap prototype; user performs point-and-click interaction; a multi-tap keyboard is mapped to the fingertips and finger joints; and user performs symbolic input to set color attributes.

Besides unimanual work performed with one hand, a working task can also be conducted bimanually with both hands with either symmetric actions, or asymmetric actions whereby both hands are assigned to different actions, but are synchronized to achieve one goal that is commonly required for bimanual everyday skills (Guiard, 1987) and depends on proprioception to update the internal forward model for movement control (Kazennikov & Wiesendanger, 2005; Yavari et al., 2015). As shown by Buxton and Myers (1986), asymmetric bimanual operations can significantly outperform the unimanual method in finding and selecting words in a document. This was also confirmed in the results of the study conducted by Bowman and Wingrave (2001), showing that the interaction with the bimanually handled tablet interface outperformed the unimanually handled TULIP menu. Using hand-referenced widgets for control or symbolic input tasks in AR or VR surroundings can also imply asymmetric bimanual operations, whereby the tracked non-dominant hand is used as interactive surface and the dominant hand is responsible for the selection of the desired task, mainly performed by direct touching with the tracked forefinger (see Figure 4.28). This led to works realizing a palm-based TV remote control (Dezfuli et al., 2012), a browser for unfamiliar imaginary interfaces (Gustafson et al., 2013) or an interactive keyboard for smart wearable displays (Wang et al., 2015). Other works used a physical hand-held panel to display control widgets (Poupyrev et al. 1998; Coquillart & Wesche, 1999; Lindeman, et al., 1999; Schmalstieg et al. 1999) or a virtual hand-held surface overlaid on the hand (Kohli & Whitton, 2005), which were operated by tracked pens or styli (Poupyrev et al. 1998; Coquillart & Wesche, 1999; Schmalstieg et al. 1999) or by the tracked forefinger (Lindeman, et al., 1999; Kohli & Whitton, 2005).

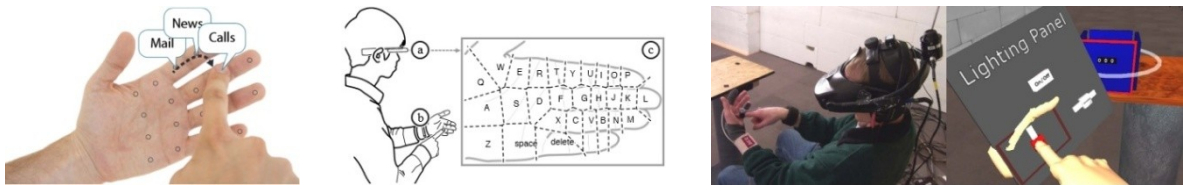
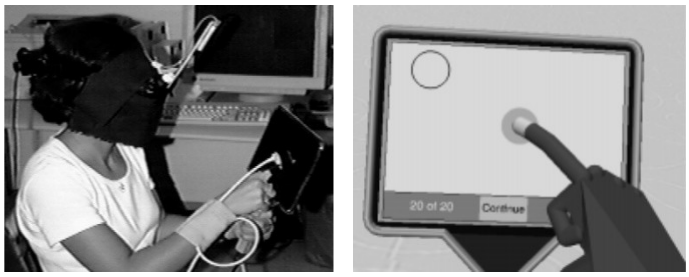


Figure 4.28: Hand-referenced interfaces requiring bimanual handling. *(left)* Browsing unfamiliar imaginary interfaces (Gustafson et al., 2013). *(middle)* Interactive keyboard for smart wearable displays (from Wang et al., 2015). *(right)* The Haptic Hand for displaying a virtual control interface (from Kohli & Whitton, 2005).

Whether by physical surfaces or by body parts, most works mentioned above designed or investigated control interfaces providing tactile sensation during the release of an control or symbolic input action.

Although the non-tactile floating menu condition used in the study from Bowman and Wingrave (2001) indicated better performance than the TULIP menu interaction, especially during the initial phase, other works have proven the importance of haptic feedback for the ability to manipulate objects (Lindeman et al., 1999) or the sense of presence. The study by Lindeman and colleagues (1999) was investigating the effect of haptic feedback on sliding and pointing tasks performed with a world- and a body-referenced interface, whereby the reference types were compared additionally (see Figure 4.29, right). While the body-referenced interface condition implied their bimanually handled HARP system (Haptic Augmented Reality Paddle, see Figure 4.29, left), the world-referenced interface condition was realized by a panel with same dimensions as the HARP surface mounted on a rigid frame standing in front of the user at the level of the dominant pointing hand. In the case of the absence of haptic feedback, only the grip of the HARP system was physically presented, while the paddle surface was replaced with a virtual representation. The same was applied to the world-referenced interface implying that the user needed to point towards a virtual panel fixed to the rigid frame. A usability study compared the four interface conditions during the performance of a docking task (see Figure 4.29, middle) and a shape selection task. While the docking task required sliding movements, the selection task has to be done by a pointing action. The results revealed that haptic feedback significantly slowed down the sliding and the pointing performance regardless of the spatial reference, while the accuracy only suffered significant impairment in the sliding performance, but this for both types of reference. The comparison of the performance between the interface references showed that the hand-held condition caused significantly slower but more precise pointing than the world-referenced condition for sliding activities during the docking task, while the selection task benefited from significantly faster pointing using the hand-held interface.



	Hand-Held (H)	World-Fixed (W)
Passive Haptics (P)	HP Treatment	WP Treatment
No Haptics (N)	HN Treatment	WN Treatment

HP = Hand-Held Window, with Passive Haptics.
WP = World-Fixed Window, with Passive Haptics.
HN = Hand-Held Window, No Haptics.
WN = World-Fixed Window, No Haptics.

Figure 4.29: The Haptic Augmented Reality Paddle (HARP) (from Lindeman et al., 1999). *(left)* User interacting with the HARP system. *(middle)* The docking task of the experiment. *(right)* The independent variables of the experiment, investigating the effect of haptics and spatial references on sliding and pointing.

Summary: As was shown by the presented works, the interaction with interfaces used for control or symbolic input tasks can be influenced by their placement, differing in its spatial reference (world, body and head) and by their handling modality (unimanual, bimanual). While Henderson and Feiner (2008, 2010) only investigated unimanually handled world-referenced interfaces, Harrison et al. (2011) compared body- (hand, arm, hand-held) and world-referenced interfaces, showing that unimanually pointing towards world-projected targets can contribute to most precise pointing, while pointing towards targets projected onto the forearm leads to less precise pointing. The bimanually handling used to point towards hand-projected targets and targets projected on a hand-held surface is not showing significant variations in accuracy, but tend to be slightly more precise using the hand-held interface. In contrast, Bowman and Wingrave (2001) compared three conditions coded in the internal frame of reference, whereby two conditions were referenced to the hand, both providing haptic feedback, and another one to the head implying the absence of tactile sensation during unimanual pointing. The hand-referenced interfaces differed in respect to their handling. While the

TULIP menu condition was unimanually operated by pinching a certain finger with the thumb of the same hand, the hand-held condition required bimanually handling. The results showed that the bimanually operated hand-held interface lead to significantly faster pointing. Other works applicable in AR surroundings have designed interesting approaches, like the ThumbsUp interface (Piekarski & Thomas, 2003) or the DigiTap device (Prätorius et al. 2014), but were not studied related to their spatial reference or handling condition. Beside the placement and handling modalities, body- and world-referenced control interfaces are benefiting significantly from haptic feedback (Lindeman et al., 1999). Overall it can be concluded that when studying placement conditions related to interfaces for control and symbolic input in AR surroundings, their *spatial reference*, the *handling modality* and the presence of *haptic feedback* need to be considered. The above mentioned works designed and studied the placement of interactive interfaces usable for control or symbolic tasks under Earth-condition, but until now there was no work in the scope of AR, investigating the effect of altered gravity on aimed pointing towards world-, body- and head referenced interfaces.

4.2 Projects on Advanced User Interfacing to Support Space Operations

Before presenting works investigating advanced user interfaces to support astronauts' space operations, it should be mentioned that Virtual Reality (VR) was already used as a mission rehearsal tool in the early nineties. Directly after launching the Hubble telescope in 1990, problems on its vision system were detected. To fix this problem, the installation of two devices was required that could only be realized by multiple extravehicular activities (EVAs). Motivated by this needed repair task, NASA built up its first VR laboratory at the Johnson Space Center. It was intended to make astronauts familiar with the upcoming repair situation and to detect risks at the corresponding task procedures (Craig et al., 2009, p. 257). After the repair mission could be successfully conducted in 1993, VR was broadly adopted to plan operational tasks and for mission training, especially for EVAs, and is until today an important part of the education for every astronaut. While wearing a HMD and special gloves, astronauts can virtually explore various locations on the space station hardware which they will be working with (see Figure 4.30, left). The VR lab also provides training for mass handling using the Charlotte[49] robot (see Figure 4.30, middle), which was adapted to simulate haptic feedback during EVA operations. Thereby the physics of objects is emulated, while seeing the imitated object through an HMD (Craig et al., 2009, p. 261). Besides astronauts' training, VR is also used to evaluate ergonomics at the Human Engineering Modeling and Performance (HEMAP) motion capture laboratory hosted at the Kennedy Space Center. The HEMAP lab is used to capture biomechanical motions, analyzing ergonomic designs or high-risk operations based on the feedback of virtual avatars (Osterlund & Lawrence, 2012; Stelges & Lawrence, 2012). This enables, for example, the study of ergonomic impacts of alternative clothing on astronaut's performance (see Figure 4.30, *right*), or the analysis of collaborative operations to support engineers with the Orion[50] design process.

[49] The Charlotte robot is a wire-driven haptic I/O device to emulate physics of large objects. It consists of eight separate motors, whereby each motor controls one wire connected to the centered device box, which mimic the emulated object. All motors are operated by a central computer, which calculates the movements of the emulated object depending on its size and mass. The robot has a motion platform with six degrees of freedom with a motion range of several feet. (URL: http://gizmodo.com/the-nasa-playground-that-takes-virtual-reality-to-a-who-1658637427, last visit: 28.09.2016)

[50] The Orion is a manned spacecraft under the development of NASA and was announced in 2011. It is intended to carry astronauts into deeper space, like to asteroids and Mars, and return to Earth. It consists of two major modules, the crew and the service module, whereby the latter one was provided by the ESA. The first unmanned flight test was conducted in 2014, while the NASA aims to bring astronauts to an asteroid in the 2020s and is planning the first journey to Mars in the 2030s. (NASA, URL: https://www.nasa.gov/exploration/systems/orion/index.html, last visit: 28.09.2016)

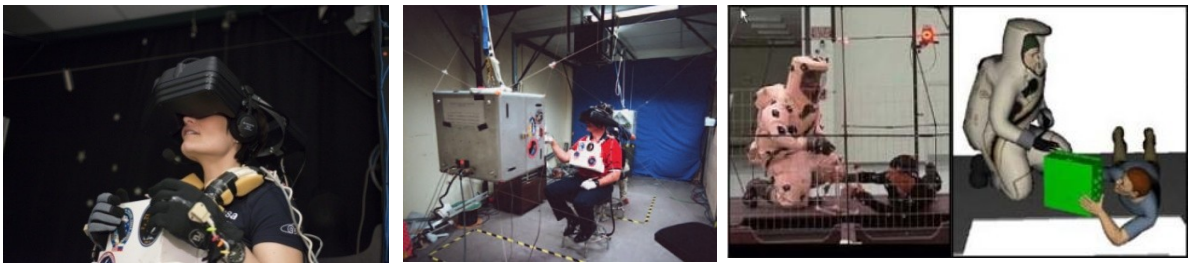


Figure 4.30: Virtual Reality for astronaut training and ergonomic evaluations. *(left)* ESA astronaut Samantha Cristoforetti using the VR hardware at the Johnson Space Center. [Photo © ESA–S. Corvaja, 2014] *(middle)* NASA astronauts practice mass handling using the haptic system and the Charlotte robot at the VR lab of the Johnson Space Center (from Craig et al., 2009, p. 262). *(right)* Analyzing of alternative clothing by the HEMAP system using biomechanical motion capturing at the Kennedy Space Center (from Stelges & Lawrence, 2012).

Since the late 1980s until today, space agencies, like NASA and ESA, are also interested in advanced user interfacing to ease astronauts' activities during their missions, initially focused on the improvement of EVA operations. In 1989 NASA has already started research on integrating visual information displays inside or in front of the helmet of the Extravehicular Mobility Unit (EMU), which is worn by astronauts conducting an EVA operation. Therefore four display prototypes (Marmolejo et al., 1989; Marmolejo, 1994) were investigated, but turned out not to be suitable for operations, like high power consumptions or were disproportionate with the low profile helmet of the EMU (see Figure 4.31a). Another work related to EVAs was conducted by the Massachusetts Institute of Technology (MIT) in cooperation with Boeing and NASA, developing a wearable situational awareness terminal (WearSAT) "[...] that provides text, graphics, and video to an astronaut via a near-eye display, and acts as a client on a wireless network that could be deployed external to the ISS" (Carr et al., 2002, p. 23). The display was placed inside the helmet and connected with a Personal Digital Assistant (PDA) built into the EMU, providing non-spatial registered digital information. Due to the bulkiness of the display, putting on and off the EMU helmet could only be done with assistance. Usability testing (see Figure 4.31b) revealed that the display provided an unhampered field of view, while the display was easily visible when positioned near the top of the field of view. The display did not interfere with nodding or translational movements, but hindered head's lateral rotations, especially to the side where the display was attached. Although none of these works were considered to be implemented for spaceflight tests, they have contributed to identify challenges faced by sizing and ergonomic issues for the designing of advanced interfaces for EVAs. In contrast, an electronic cuff display developed by NASA (Marmolejo, 1996) was evaluated during flight tests aboard the Space Shuttle Orbiter between 1994 and 1996. The display (see Figure 4.31c) should replace the wrist-mounted paper checklist commonly used by EVA tasks, but the flight tests revealed usability problems, such as glare, lack of contrast, small font size and cold intolerance. Similar research was conducted by the ESA in 2006, discussing in detail design issues for a Direct Visualization Display Tool (DVDT) intended to be used for EVA operations during Mars and Moon explorations, as well as during operations in orbit outside a spacecraft (Pereira Do Carmo et al., 2006). Following the approach of the WearSAT system (Carr et al., 2002), the DVDT system should consists of a monocular OLED display placed inside the EMU helmet (see Figure 4.31d) and a DELL AximX50v PDA as control unit. The authors mainly contributed by specifying the optical, mechanical and ergonomic requirements for such display tools, as well as by identifying numerous application scenarios, like displaying biomedi-cal or space suit telemetry data, providing navigational support or replacing the current cuff checklist.

(a) (b) (c) (d)

Figure 4.31: Advanced displays to support EVAs. *(a)* Display prototypes explored by the NASA (from Pereira Do Carmo et al., 2006). *(b)* Subject testing the WearSat prototype developed by the MIT with Boeing and NASA (from Carr et al., 2002). *(c)* The electronic cuff checklist developed by NASA (from Marmolejo, 1996). *(d)* 3D artistic illustrating the DVDT system discussed by the ESA (from Pereira Do Carmo et al., 2006).

While the above mentioned works aimed to extend astronauts' training with virtual environments or to improve EVA operations by near-eye or wrist-worn displays, other works, more relevant in the scope of this thesis, were and still are conducted to enhance astronauts' performance during intravehicular activities (IVAs), especially by applying Augmented Reality (AR). In 1997 NASA initiated a project discussing the concept of their first Wireless AR Prototype (WARP) for IVA operations. The WARP system was intended to be a personal communication interface for real-time display of information related to health and safety issues of the space station personnel (Agan et al., 1998). Once again in 2006, Boeing, the U.S. prime contractor for the ISS, reported on potential applications of AR for space training and operations. They said that "[...] AR can save money by reducing paper procedures while significantly decreasing training, operations and logistics requirements" (Memi, 2006, p. 21). Especially in the future, when astronauts will travel to Mars, "[...] astronauts will face the daunting task of operating and maintaining numerous systems that might unexpectedly break, or maybe having to perform life-saving surgery. Because of time delays of more than 15 minutes, communications back to earth can be a problem. And it's difficult for a crew member to remember every detail of every system. That's where augmented reality can help reduce training requirements and communicate complicated technical instructions." (Memi, 2006, p. 21).

In the meantime ESA studied the potential of PDAs (Martignano, 2006; Garcia, 2011) that were tested aboard the ISS until 2006. It was reported that PDAs could be used as a complementary tool for astronauts' operations or as a alternative platform for current applications, like the international Procedure Viewer (iPV) displaying Operation Data Files (ODFs). To further improve ISS operations commonly guided by the iPV (see section 2.3.4), the ESA has commissioned the development of a voice activated Procedure Viewer (vaPV, see Figure 4.32, left), which was part of the Crew usability demonstrator (Crusade) proposed in 2011 (Wolff, 2011). Beside the hands-free vaPV condition, Crusade also provided guidance by Procedural Displays (PDs) that extended ODF procedures by embedding displays to carry out commands and control operations directly within a procedure (see Figure 4.32, right). In December 2012, Crusade was studied aboard the ISS by two astronauts in the scope of the CRUISE[51] (CRew User Interface System Enhancement) experiment investigating the effect of microgravity on procedure execution guided by the vaPV and PD conditions (Smets & Neerincx, 2013). The study revealed that using the vaPV caused an increase in mental effort and rate of failures, but led to a higher degree of situation awareness, while the PD condition benefited from higher subjective ratings related to satisfaction and trust. While the CRUISE experiment investigated control conditions of the iPV, whereby the procedure content was still presented on the Station

[51] ESA's CRUISE (CRew User Interface System Enhancement) experiment aboard the ISS.
URL: http://www.nasa.gov/mission_pages/station/research/experiments/1225.html, last visit: 28.09.2016

Support Computer (SSC), in 2014 the ESA initiated the mobiPV project intended to enhance the display modality of the viewer. This project was intended to demonstrate new operational concepts for ISS task support by a mobile Procedure Viewer (mobiPV), which provides the presentation of procedure instructions and related multimedia resources on mobile devices using commercial off-the-shelf (COTS) equipment. The mobiPV system was tested during the NEEMO 19 and 20 missions, as well as aboard the ISS in September 2015 (Chintamani et al., 2013; Boyd et al., 2016).

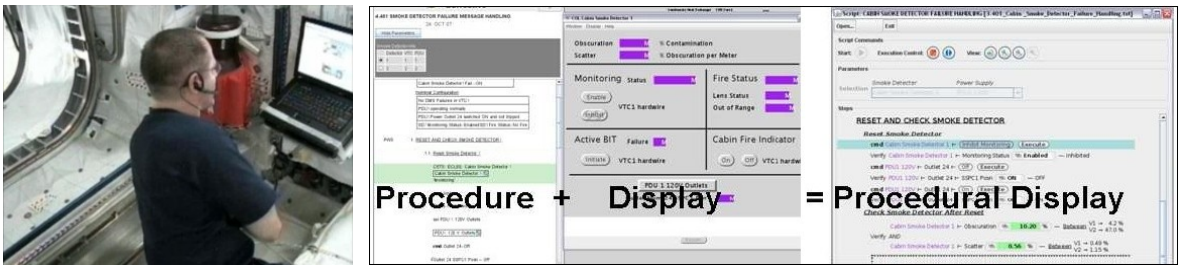


Figure 4.32: The Crew usability demonstrator providing task guidance by a voice activated Procedure Viewer (vaPV) and Procedural Displays (PDs). *(left)* CSA astronaut Chris Hadfield using the vaPV aboard the ISS (from Chintamani et al., 2013). *(right)* A PD merging a common ODF procedure and display items to carry out commands and control operations (from Smets & Neerincx, 2013).

More related to Augmented Reality considering Azuma's criteria (Azuma, 1997), ESA initiated and funded few projects investigating AR technologies for medical diagnosis as well as on-board and remote task assistance. In 2009, the ESA started the project CAMDASS (Computer Aided Medical Diagnostics and Surgery System) that should provide just-in-time medical assistance during long-term spaceflights (Fritz et al., 2009; Nevatia et al., 2011). It was mainly focused on AR assisted ultrasound examinations, so that astronauts will be able to examine each other while performing medical ultrasound procedures and diagnosis with the help of ground-based specialists. Work more closely linked to the research presented in this thesis, was intended to assist astronauts' in their performance of maintenance tasks commonly guided by ODF procedures. Therefore, in 2008, ESA funded the WEAR project (WEarable Augmented Reality) that was tested aboard the ISS in 2009 (Arguello, 2009; Scheid et al., 2010). This project aimed at a wearable AR system, but was mainly focused on issues related to the see-trough display and spatial pose tracking. In 2013, ESA together with the Delft University of Technology conducted research on AR-based collaboration investigating virtual co-location to remotely support intravehicular maintenance tasks in ground training (Datcu et al., 2014). In the beginning of 2014, the NASA has also shown interest in applying AR technologies for ISS operations and announced a partnering opportunity that included topics like authoring and viewing of AR-based procedures, just-in-time training, distributed AR team training, 3D pose tracking and interactive interfaces to control real and virtual objects with gesture, voice and haptic devices (NASA JSC, 2014). All these projects are revealing a strong interest to improve spaceflight operations by applying innovative interface technologies, like Augmented Reality. As chronologically presented in Figure 4.33, a selection of these projects claiming the usage of AR for intravehicular space operations (WARP, WEAR, CAMDASS, Virtual Co-Location) or intending to enhance the display modalities of the iPV (mobiPV) will be presented and discussed more detailed in the next subsections.

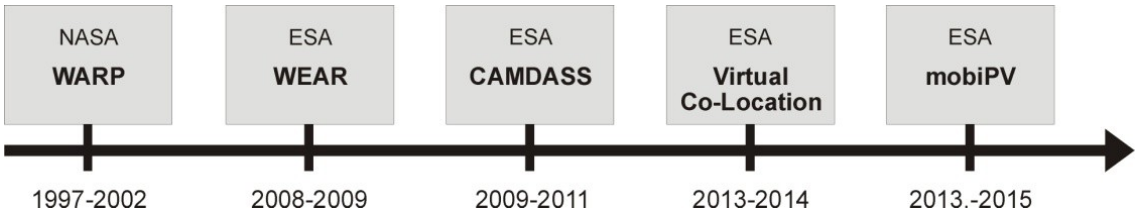


Figure 4.33: Timeline showing the most important projects focused on AR interfaces or enhancing the iPV to assist astronauts during intravehicular space operations.

4.2.1 WARP (1997-2002)

Title:	**W**earable **A**ugmented **R**eality **P**rototype (WARP)
Purpose	Conceptual design of a wearable personal communication device enabling real-time access to voice communication, pictures, video, documents and biofeedback.
Associated partners:	NASA, Jet Propulsion Laboratory (JPL)
Tested platforms:	-
References:	Agan et al., 1998; NASA JPL, 2002

In 1997 the NASA was already interested in equipping astronauts with lightweight wearable computers enabling hands-free operations while providing information about crew members' health and safety. The system should allow astronauts to wireless communicate audio and video, to view documents via a near-eye display as well as to capture and display biosensor data. Contracting with the NASA, the JPL[52] drafted a conceptual design how the WARP system may look like (see Figure 4.34) and how it can link to the space station communication infrastructure. Thus, this work was mainly focused on specifying technical issues, whereby high-level requirements for the headset including a miniature HMD, the codec for the audio and video sources, the assessment of biofeedback, the wireless communication system transmitting the audio, video and biosensor data, as well as voice controlling to operate the WARP system were discussed and defined.

Figure 4.34: NASA's Wireless Augmented Reality Prototype (WARP). *(left)* Conceptual sketch of the WARP system (from Agan et al., 1998). *(middle)* Ann Devereaux, JPL staff working on WARP, wearing a dummy of the WARP system (from NASA JPL, 2002). *(right)* Mockup of the WARP headset equipped with a near-eye display and a noise canceling microphone (from NASA JPL, 2002).

[52] The Jet Propulsion Laboratory (JPL) mainly build and operate satellites and space probes for the NASA and is part of the California Institute of Technology (Caltech). URL: http://www.jpl.nasa.gov/index.php, last visit: 28.09.2016

Therewith first prototyping should have considered then-widespread compression standards for video (H.261), audio (G.722) and multiplex coding (H.221), as well as a Pulse Oximeter sensor providing biofeedback of the pulse rate and oxygen blood saturation. The headset should have been equipped with stereo audio, a noise canceling microphone and a flip-up video display providing NTSC video, while the VGA standard was intended for a later stage. To reduce the size of the control hardware, the transceiver, and the encoder/decoder should be implemented on a custom-built computer using COTS components. Because there are no reports about the progress of a prototype nor corresponding usability tests, this work has mainly contributed to identify challenges for the deployment of wearable systems supporting astronauts aboard a space station. Although the project's name claim the usage of AR, the WARP system belongs more to the field of wearable computing.

4.2.2 WEAR (2008-2009)

Title:	**WE**arable **A**ugmented **R**eality (WEAR)
Purpose	Design, implementation and onboard tests of a lightweight, modular and wearable AR system supporting ISS operation.
Associated partners:	ESA, Space Applications Services, Catholic University of Leuven (KU Leuven)
Tested platforms:	Ground (EAC), ISS
References:	Arguello, 2009; Scheid et al., 2010; Tingdahl et al., 2011

To safe crew time and increase efficiently working onboard the ISS, the ESA funded the project WEAR, that aimed to assist astronauts' performance of ISS procedures enabling hands-free operations. The WEAR system was intended to provide context-sensitive information presented in a static or spatially registered way, and displayed on an optical see-trough (OST) HMD, providing visual support for task localization with 3D-registered resources, like labels, annotations or highlighted frames (Figure 4.36, right), which enabled hands-free operation while reading manuals and checklists, viewing multimedia references, like images and video. Furthermore, the system should be controlled by voice commanding and capable of supplying synthesized voice output. It should also provide barcode reading to visually or acoustically inform an astronaut about the referenced meaning of a used ISS tool or item. To facilitate decision-making, to choose most suitable technologies to apply WEAR onboard the ISS, the initial step required the identification of the specific working conditions (inside a space vehicle; orbiting in LEO; microgravity; no GPS signal; weak, noisy and varying magnetic field; and known environment enabling to use CAD models, IT infrastructure). This has induced the choice to use, for example, computer vision based techniques combined with an Inertial Measurement Unit (IMU) to track the 3D pose to enable AR user interfacing, or to use the Microsoft Speech API (application programming interface) for speech recognition and synthesizing. As the primary contractor for this project, Space Applications Services[53] was responsible to conceptually design, implement and validate a running demonstrator built on R&D activities. The resulted head-referenced prototype (see Figure 4.35) consisted of a head- and microphone, and an adjustable head harness equipped with a tracking and a barcode camera, a plug connector for the motion sensor and an adjustable monocular see-through display. All these hardware components were connectable via VGA and USB to an IBM A31p laptop available onboard the ISS, while the needed software components where delivered on a replaceable hard disk compatible with the laptop used.

[53] Space Applications Services is an Belgian company founded in 1987 and focused on innovative technology, solutions and services for the aerospace sector. URL: http://www.spaceapplications.com, last visit: 28.09.2016

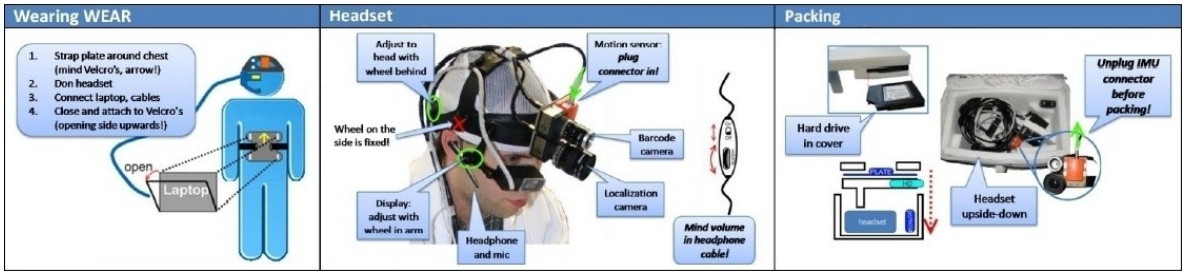


Figure 4.35: The ESA's WEarable Augmented Reality (WEAR) system (from Scheid, et al. 2010).

The vision-based tracking algorithm was developed by Tingdahl et al. (2011) from the KU Leuven, a subcontractor for the WEAR project. Because it is barely possible to equip the ISS with fiducial markers and a detailed virtual 3D model of the Columbus lab was available, the camera's 3D pose was obtained from a hybrid approach combining the motion data captured by the IMU and markerless model-based tracking (Lepetit & Fua, 2005) considering offline generated keyframes (Lepetit et al., 2003; Platonov et al., 2006). The keyframes containing training images of the Columbus lab, their 2D features and 3D correspondences, were generated on ground before the Columbus module was launched. While the visual-based tracker achieved real-time performance during ground tests on a Intel Core i7 2.8 GHz CPU, the performance in space on the IBM A31p laptop equipped with a Mobile Intel Pentium 4 (1GB RAM) was considerably slower. Therefore the IMU was used for real-time measurements, while the visual-based tracker initiated the 3D pose tracking and corrected the IMU data periodically, but still revealed bottlenecks in performance caused by the number of keyframes needed to cover the Columbus lab. For future work it was suggested to use a semi-SLAM, whereby areas of interests are tracked with high accuracy while moving and could be updated by Simultaneous Localization and Mapping (SLAM).

Figure 4.36: ESA astronaut Frank De Winne testing the WEAR system. *(left)* Ground training in the Columbus mockup at the EAC in Cologne. [Photo © ESA–M. Soppa, 2009] *(middle)* Test aboard the ISS showing the WEAR components (from Scheid et al., 2010) *(right)* Mockup of a WEAR user interface showing a procedure step enriched with static and 3D-registered context-sensitive information (from Chintamani et al., 2013).

The WEAR system was tested onboard the ISS in 2009 performed by ESA astronaut Frank de Winne (Scheid et al., 2010), who was previously trained on ground in the Columbus mock-up at the EAC in Cologne (see Figure 4.36, left, middle). The procedure being performed was a maintenance task to inspect the Desiccant Module of the Condensate Water Separator Assembly (CWSA) located in the Columbus module. The first attempt took place in September 2009, but the test was not performed successfully, because the visual tracker did not work correctly, which was caused by significant discrepancies between the Columbus configuration used to previously generate the keyframes and the actual onboard configuration. A recovery plan was developed to update the keyframes by acquiring and downloading new images of the Columbus to create new keyframes, which in turn were uploaded

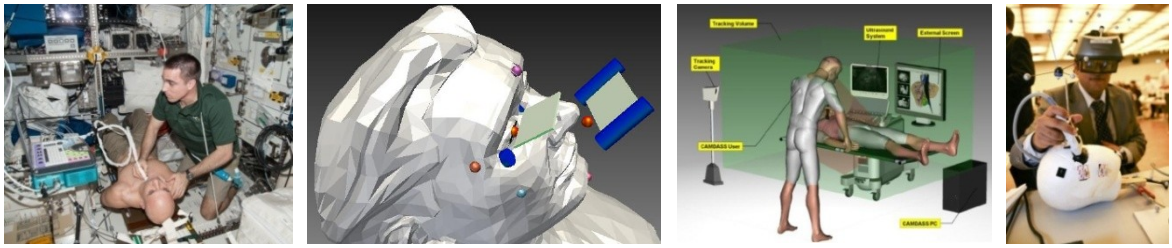
to the onboard WEAR system together with a marker that was placed on a payload rack in the Columbus lab to train parts of the keyframes. In November 2009, a second attempt using the new keyframes was conducted showing that the WEAR system worked as expected and the test session was completed successfull. Besides performance bottlenecks of the visual tracker, the performed test corresponded to operational tests and lacked reasonable usability research, but it was reported that a post-session questionnaire and interview with Frank de Winne exposed that "[...] WEAR-type technology could provide significant help in future onboard operations" (Scheid et al., 2010, p. 8).

4.2.3 CAMDASS (2009-2011)

Title:	**C**omputer **A**ided **M**edical **D**iagnosis and **S**urgery **S**ystem (CAMDASS)
Purpose	Design and implementation of an AR based assistance system to guide non-experts through ultrasound examinations during long-term space missions.
Associated partners:	ESA, TU München, DFKZ, Space Applications Services (SAS)
Tested platforms:	Ground-based demonstrator
References:	Fritz et al., 2009; Nevatia et al., 2011

Not for supporting payload operations, but intended to guide astronauts through medical procedures, the project CAMDASS, funded by ESA, was targeted to design and develop an AR system enabling medical assistance to be used by astronauts on prolonged spaceflight missions. In cases where no medical crew officer is available onboard the spacecraft and emergencies caused by, for example, injuries, diseases or experimental physiological measurements, the crew members normally contact the ground control to get professional support in form of instructions given by an physician. Because space missions beyond Low Earth Orbit can cause long delays restricting the communication, the ESA was interested in identifying the challenges of developing an assistance system using AR, which should provide medical expertise enabling non-experts to perform examinations among each other, initially investigated for ultrasound as imaging modality. Therefore, CAMDASS should be able to track the ultrasound probe, while the non-expert examiner should be visually guided to place the probe correctly, whereby the most critical issue in medical AR is the accurate 3D registration to the patient's body. In general such registration processes are challenging, whereby individual differences in human physiology need to be considered, for example, by applying a personalized reference model, which in turn needs to be adapted to dynamic processes, like breathing movements. The CAMDASS system provided a Patient Body Registration module, which in its initial state offered a simple registration technique, locating a set of marker at a certain region of patient's body and a respective virtual reference model. To support the acquisition of appropriate sonographic slices, the CAMDASS prototype provides the display of featured reference images, but optimal handling of an ultrasound probe, only by means of reference images, requires extensive training, which on its own can benefit from AR, as was shown for neonatal cranial sonography (Markov-Vetter et al., 2009). To present the visual guidance cues, it was intended that CAMDASS provided a VR representation displayed on a screen (see Figure 4.37b) or an OST HMD, which in turn called for an accurate eye-display calibration that was realized by a semi-automatic approach using the Single Point Active Alignment Method (Tuceryan, 2002). The CAMDASS prototype comprised COTS components, like the NVIS nVisor ST to display context-sensitive guiding instructions in the user's field of view, either by static or contact analogue presentation registration, whereby the latter one was realized by tracking the 3D pose of user's head and the ultrasound probe, obtained by passive infrared-reflective markers using the Polaris Spectra system (Northern Digital Inc.).

Evaluating the usability by means of a between-subject study, using a task procedure for muscle mass measurements assessed by ultrasound, was planned, but never conducted. However, a CAMDASS prototype was demonstrated in 2011 during the demo session of the International Symposium on Mixed and Augmented Reality (ISMAR) in Basel, Switzerland. The demonstrator as shown in Figure 4.37 (d) provided a simulated cranial sonography, using a tracked plastic head and ultrasound probe, while the user could move the probe, seeing the according reconstructed ultrasound image in real-time through an optical see-through HMD (nVisor ST).



(a) **(b)** **(c)** **(d)**

Figure 4.37: The CAMDASS project. *(a)* Common ultrasound scan aboard the ISS performed by NASA astronaut Chris Cassidy examining ESA astronaut Luca Parmitano [Photo © NASA, 2013]. *(b)* AR guidance to place the probe (from Nevatia et al., 2011). *(c)* Artistic representation of the prototype (from Nevatia et al., 2011). *(d)* Demonstrator during the ISMAR 2011 [Photo © ESA– Space Applications Services, 2012].

4.2.4 Virtual Co-Location Assistance (2013-2014)

Title:	Virtual Co-Location Assistance
Purpose	Exploring the potential of AR on remote collaboration to support astronauts' task performance by virtual co-location.
Associated partners:	ESA, Delft University of Technology (TU Delft)
Tested platforms:	Ground (EAC)
References:	Datcu et al., 2014; Datcu et al., 2014a

Intended to support astronauts' in emergency situation during task performance onboard the ISS, but also initially considering the appropriate ground training, the ESA was interested in exploring the benefit of applying AR to remotely support maintenance operations. Therefore the TU Delft, the prime contractor for this project, developed and evaluated an AR assistance system enabling remote collaboration by virtual co-location. Hence a trainer can provide visual support, which was virtually co-located at the side of the trainee in real-time (see Figure 4.38a,b). This enabled a shared view, whereby the trainer was able to interactively augment video sequences at the working space of the trainee, who was wearing an OST HMD. The video stream was captured from the built-in camera sensor at the trainee's HMD. For placing and viewing visual annotations in real-time, the virtual co-location assistance supported marker-less 3D pose tracking using RDSLAM, providing Robust Dynamic Simultaneous Localization and Mapping (Tan et al., 2013).

The usability of the system was evaluated for a training scenario by two users, performing nominal procedures for the Smart Gas Sensor (SGS) payload, conducted in the Columbus mockup at ESA/ESTEC in Noordwijk, Netherland. The experiment task included the connection of the SGS to a laptop computer, changing a fuse and verifying the states of switches and LEDs. The experimental

setup consisted of the Spaceglasses HMD from Meta[54] and a laptop computer worn by the trainee (see Figure 4.38c), while the trainer operated a real-time authoring tool on a laptop computer to augment the trainee's scene with a common mouse device. Both laptops, of the trainee and the trainer, were wired connected. The task performance was guided by text-based instructions central displayed on the trainee's HMD in a non-registered manner, while the trainer enriched the scene by placing arrows used for task localization and indicating a direction in real-time. Although the task performance was described as good, there was no information given on a quantitative assessment. However, subjective experiences revealed that AR has the potential to support remote collaboration to train onboard procedures. The study was negatively impacted by incorrectly placed annotations, caused by faulty tracking calibration, the wired connection between both laptops, a missing option for voice communication between the collaborators and the restricted space to place text into the HMD's field of view. Beside these issues, future work will also consider a lightweight AR HMD, the improvement of 3D pose tracking, hand-based interaction and the collaboration of multiple trainers and trainees.

(a) **(b)** **(c)** **(d)**

Figure 4.38: The Virtual Co-Location Assistance for remote collboration. *(a)* Artistic representation of a collaboration scenario between the trainer and the trainee (from Datcu et al., 2014). *(b)* Shared view of the trainer and the trainee (from Datcu et al., 2014). *(c)* Participant of the usability study while working at the SGS payload inside the ESA/ESTEC Columbus mock-up in Noordwijk (from Datcu et al., 2014a). *(d)* Augmented scene showing a HUD text instruction and a colocated arrow indicating the direction (from Datcu et al., 2014a).

4.2.5 mobiPV (2014-2015)

Title:	**mobi**le **P**rocedure **V**iewer (mobiPV)
Purpose	Introducing and testing of multiple mobile devices to enhance the display modality of the international Procedure Viewer (iPV).
Associated partners:	ESA, Space Applications Services, Skytek
Tested platforms:	Ground (EAC), NEEMO, ISS
References:	Chintamani et al., 2013; Chintamani et al. 2015; Boyd et al., 2016

Taking into account previous works conducted by the ESA, like CRUISE and WEAR, both already aimed at improving or replacing the iPV, displaying ISS procedures, the ESA funded the project mobiPV to explore the potential of mobile devices for collaboration and complementary support. Beside text-based instructions, ODF procedures can be complemented by special hints for notes, cautions and warnings, as well as visual resources, like pictures, drawings or illustrations. In the future, such procedures should also be extended by rendering 3D content for animating CAD models and with streaming video sequences to enable collaborative work for ground-based support. Commonly, iPV guided procedures are displayed on the SSC laptop and controlled by inputs obtained

[54] Meta is a manufacture of AR displays and was founded by Meron Gribetz in December 2012. Meta is supported by Steven Feiner as lead advisor and Steven Mann as chief scientist. URL: http://www.getameta.com, last visit: 28.09.2016

from the keyboard and mouse device. Initially the ESA and partners identified the key features for new user interfaces (e.g., wearable computers, head-worn displays) and interaction modalities (e.g., voice commanding, touch screen, collaborative working), which may be provided by mobiPV (Chintamani et al., 2013). They also drafted several design options, whereby the initial concept should comprise a customized wearable computer connected to a monocular OST HMD, an audio headset and a head-worn camera. Thereby the HMD was intended to be used as primary display to present the procedure instructions following the approach from the WEAR system (see section 4.2.2), while larger procedure content not suitable for the limited space of the HMD's field of view, should be displayed on an assistive handheld tablet. Together with crewmembers, flight controllers and training specialists, this design concept was discussed and revised, which resulted in an updated concept now considering COTS hardware. The final mobiPV should constitute a configurable system comprising the LG Nexus 5 smart phone, the Apple iPad and the Google Glass (see Figure 4.39, left), as well as a head-worn camera (not been designated) used for the generation of a shared space for real-time collaboration between astronauts and their ground teams and for mutual communications between the members of the flight crew. Thereby procedure viewing should be provided on a wrist mounted smart phone, a Google Glass and a iPad, while speech recognition should be used for procedure navigation. Besides images, the procedure was complemented with video and text notes. For collaborative purposes, mobiPV supported remote note sharing and the control of procedures by live streaming audio and video between low orbit and ground.

A prototype of mobiPV was initially used aboard the underwater research laboratory Aquarius during the NEEMO 19 and 20 missions (see Figure 4.39, middle). In September 2015, ESA astronaut Andreas Mogensen tested the mobiPV system aboard the ISS in the course of the IRISS short-duration mission (see Figure 4.39, right). As was reported by ESA and its partners, the onboard demonstration did not work out as planned (Chintamani et al., 2015). Due to the mission priorities, the onboard session was limited to only one hour. Furthermore, only the backup configuration of the mobiPV system was tested, which consisted of the smart phone and iPad. The operational test showed that procedure viewing, sync and note sharing worked, while video streaming only partially worked, but lacked of functioning audio and had limited network bandwidth. From the tests in NEEMO and on the ISS, the findings revealed a personal preference for the tablet display compared to the smart phone, while the Google Glass was criticized regarding its limited display space and its heat generation. The feature providing quick syncs of procedures and the opportunity to share pictures and videos received positive feedback. Intended for procedure navigation, voice recognition still remains a major issue relating to its reliability.

Figure 4.39: The mobile Procedure Viewer - mobiPV. *(left)* The display modalities (from Chintamani et al., 2015). *(middle)* ESA astronaut Andreas Morgensen testing mobiPV during the NEEMO 19 mission in September 2014 (from ESA, 2015e). *(right)* Andreas Morgensen using mobiPV onboard the ISS in September 2015 (from Skytek, 2015).

Chapter 5

Human Factors as Challenge on Instructional Augmented Reality for Space Payload Operations

"For any question there is always a corresponding method of finding. Or you might say, a question denotes a method of searching."

(Ludwig Wittgenstein, cited from T. Burke: *Dewey's New Logic: A Reply to Russell*, 1994)

Based on the theoretical foundation and reviewed works presented in the previous chapters, this chapter clarifies the need for AR as advanced interface for space payload tasks by analyzing their functions and tasks, the provided resources and training, as well as discusses the effect of weightlessness on astronauts' performance that needed to be considered when designing interactive AR interfaces. Because the efforts that have been made to ease astronauts' activities so far, were mainly focused on technological aspects, the lack of Human Factors will drive the identification of challenges, which builds the foundation for the research presented in the subsequent thesis chapters.

5.1 Need for Advanced Payload Interfacing

During their average six month stay, astronauts' schedules aboard the ISS are tight, managing their tremendous amount of daily work activities. Thereby a majority of their working time is intended for maintenance and experimental tasks at research payloads, housed in International Standard Payload Racks (ISPRs, see section 2.3.1). Such tasks have to be carried out within a predefined timeframe using an assigned standardized procedure, presented in a Procedure Operation Data File (PODF, see

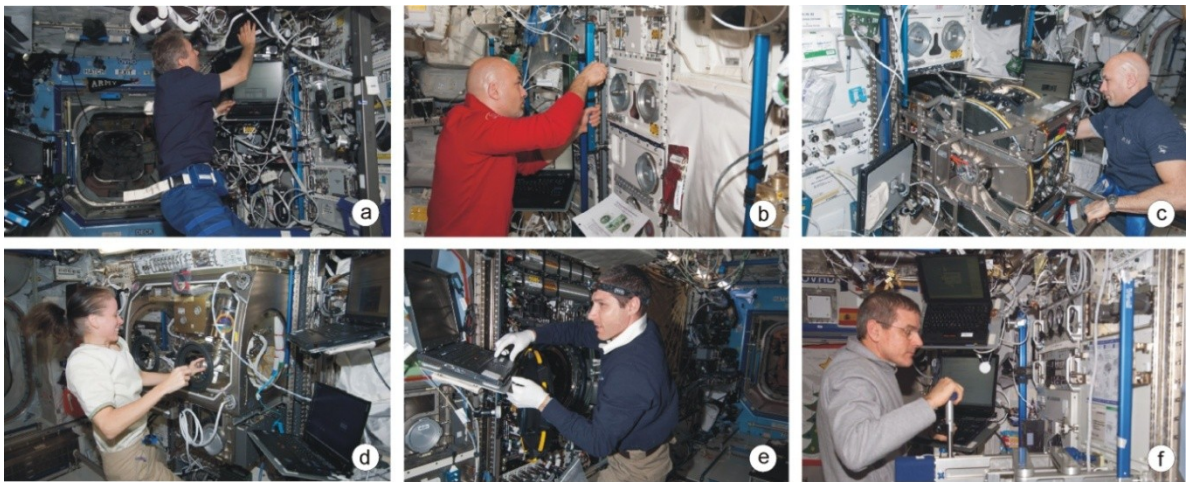


Figure 5.1: ISS payload operations guided by instructions on the SSC laptop: *(a)* ESA astronaut Thomas Reiter prepares the HRF-2 rack for next experimentation in the Destiny lab. [Photo © NASA, 2006] *(b,c)* ESA astronaut Luca Parmitano performs maintenance for the Biolab in the Columbus lab. [Photos © NASA, 2013] *(d)* NASA astronaut Shannon Walker works on an experiment inside the Microgravity Science Glovebox (MSG) in the Columbus lab. [Photo © NASA, 2010] *(e)* NASA astronaut Mike Hopkins services research hardware in the Destiny lab. [Photo © NASA, 2013] *(f)* NASA astronaut Bill McArthur sets up the Space Linear Acceleration Mass Measurement Device in the Destiny lab. [Photo © NASA, 2006]

section 2.3.4), which guides through the requested payload task in step-by-step instructions often complemented by visual reference resources and should guarantee the reliability of each operation. Thereby instructions can vary in their level of difficulty ranging from simple (e.g., *"sw RACK POWER → OFF"*) to more complex instructions (e.g., *"Screw Support Jig L and R captive screws (four per jig) to Thermal Housing frame [...]"*). Beside paper-based media, viewing the procedures is interfaced with the international Procedure Viewer (iPV, see section 2.3.4) provided on the Station Support Computer (SSC), a laptop computer laterally fixed to the working area (see Figure 5.1). Additionally, the SSC laptop also provides displays for telemetry and states payload, because most of them, like the Biolab payload (see section 2.3.2), don't have embedded displays. Hence, the guiding instructions for the astronaut and the output of the research payload, both are interfaced with the SSC laptop, which furthermore is used as an input- and control device to perform procedure-related actions, such as navigating through a procedure and entering command-based instructions.

These issues constitute the problem space in question, calling for a redesign of the current guidance interface to ease astronauts' working at ISPRs. Before clarifying and discussing related challenges, this section analyzes in detail the problem space, taken the knowledge presented in the last chapters into account, pointing out how the current mode of payload operation can benefit from AR and finally analyzes competitive works to identify issues by considering aspects of Human Factors.

5.1.1 Analyzing the Problem Space

To avoid a technology-driven system design, it is important to consider the aspects of Human Factors from the beginning (see section 3.1). Therefore, it is fundamental to gain a deeper understanding of the problem space by front-end analysis. Consequently, the design of the current operation mode is reviewed in reference to common inspection methods for analyzing user's needs (see section 3.4.1). Following the brief outline of the problem space above, this section presents the results of various analyses (see Figure 5.2). Taking perceptual and cognitive aspects of human information processing into account (see section 3.3), these analyses are related to the function and tasks of SSC-interfaced payload operations, the provided resources, the available payload training, the user population in question and to the environmental condition, in which payload operations have to be carried out.

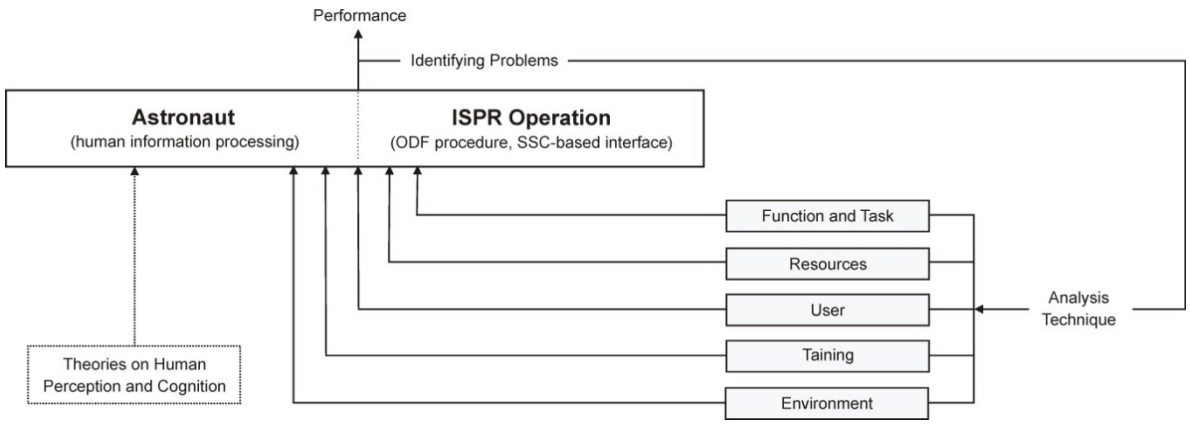


Figure 5.2: Conceptual analysis framework used to identify possible problems affecting astronauts' performance of payload operations (adapted from the Human Factors circle from Wickens et al., 2003, p. 3).

Function and Task Analysis

Applying the conceptual framework of the underlying human-machine system (HMS) to SSC-interfaced payload operations, it is conspicuous that astronauts not only need to cognitively process the state of the payload elements, they also need to perceive and process the PODF instruction, on which the next operation relies on. Thus, the PODF procedure and its SSC based interfaces are becoming an integral part of the closed-loop system of the payload HMS, mapping the continuous interaction between an astronaut and the payload. On closer inspection, it is noticeable that the SSC not only serves as interface of the payload HMS, the interaction between the astronaut and the laptop device maps a fully independent human-machine system, which in this case can be considered as human-computer interaction (HCI) implicating its own closed-loop system. Thus the additional effort astronauts make to receive the operational task instruction is influencing their work at the payload, which is illustrated in Figure 5.3. Assuming that the iPV interface is already started, the needed procedure loaded and the header information read, the astronaut navigates initially through the procedure to the instruction in question and visually perceives the current payload instruction and views its associated visual reference sources, if available. The perceived information must be cognitively processed, understood and memorized. Depending on their complexity and presentation, the perception can be driven by effortful top-down processing, while understanding and memorizing an instruction can lead to a high load of the working memory (see *"Resource Analysis"*). Resulting from this perceptual-cognitive process, the astronaut selects the required sensorimotor action. In accordance with the instruction content, the astronaut is either entering a command on the SSC's keyboard or turns to the payload. Once he or she is directed toward the payload, then the target object in question needs to be localized by goal-directed visual search. Thereby the process to localize the target area is supported by a location code (e.g., "COL1A2_J1"), which is assigned to each instruction. In contrast, the position information to subsequently detect the target object in question,

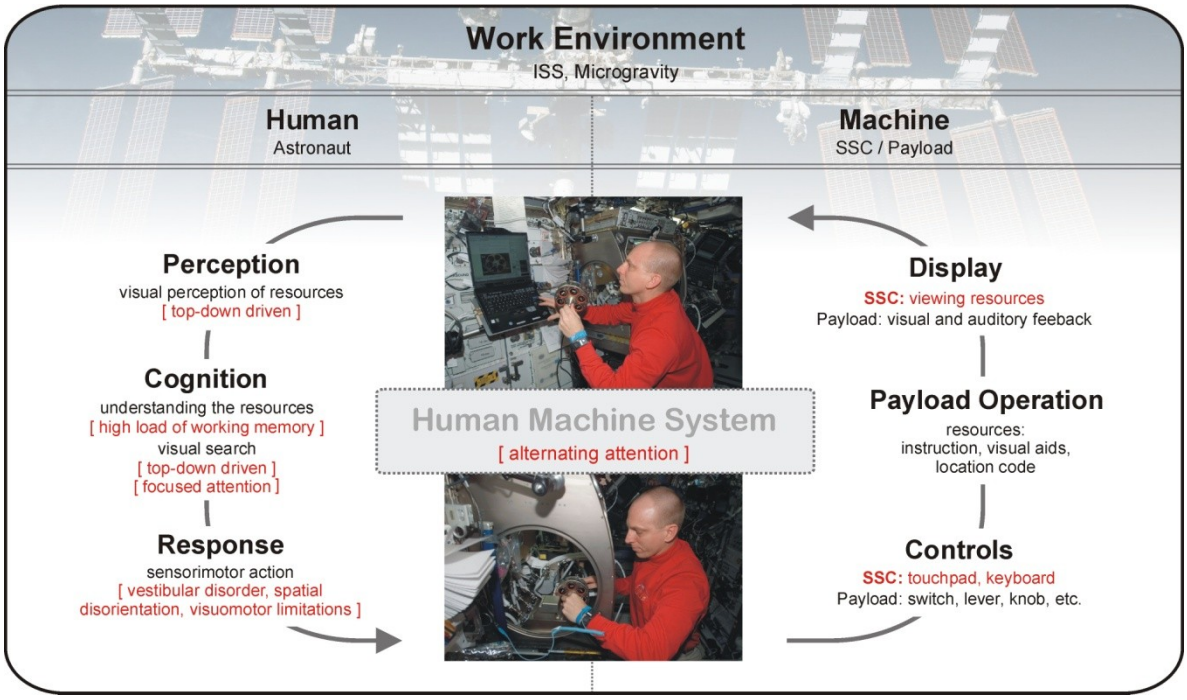


Figure 5.3: The influences of the SSC-based guidance on the underlying payload HMS with respect to astronaut's human information processing (HMS framework adapted from Proctor & Van Zandt, 2008, p. 11). The photos show NASA astronaut Clay Anderson working at the MSG (Microgravity Science Glovebox) payload in the Destiny laboratory aboard the ISS [Photos © NASA, 2007].

needs to be extracted from the visual reference sources, like drawings and pictures. Although it can be assumed that the target object is already visible in the current field of view, the effort needed to actively scan the visual scene, depends on its given features. By endogenous orienting, which is top-down driven by the knowledge of the goal held in working memory, the astronaut searches for the target object among an array of distractors. Thereby, the objects' features of the visual scene are preattentively analyzed in parallel and need to be subsequently serial combined, which can result in a costly conjunction search by scanning saccadic eye movements, allocating focused attention by multiple focusing steps (see section 3.3.3). Thereby the amount of attentional resources depends on the number of distractors, as well as on the uniqueness and connectivity of the features. Once the target object is detected, the astronauts finally respond to the requested instruction by physical interaction with the payload that is afterwards transitioning its state, and the sequence starts again by receiving the next payload instruction. Whether working at the payload or entering a command on the keyboard, the physical response is affected by the loss of the gravitational force (see *"Environment Analysis"*).

Consequently, both, the payload HMS and the SSC HCI, claim the astronaut's information processing in a separate manner, adding to a repeated switch of focus and thereby interrupting the flow of the payload work. Besides the fact that the visual search claims focused attention to conjoin features correctly, each focus change, not only calls for wasteful head and eye movements, but also requires re-adaptation to either the SSC or the payload, whereby each of them allocates alternating attention. Such changes will be more frequent when the instruction resources are complex and need to be verified in between. In conclusion, this analysis reveals that the current mode of payload operation is characterized by continuous allocation of alternating and focused attention.

Resource Analysis

Besides the location code, resources are defined as text-based instructions and visual aids. While a text-based instruction is compiled by the rules and syntax of the PODF standard (Pels et al., 2004), visual aids are intended for its complementation. Although this thesis is not intended to alter the PODF standard (see section 5.3), however, possibly effects on human information processing will be discussed. Further issues are focused on the current and future standard for visual aids, as well as the impact of their presentation.

Text-Based Instruction: Considering the way a text-based instruction is received related to the types of perceptual processing (see section 3.3.1), it is mentioned that the PODF standard already supports automaticity and unitization by using familiar perceptual representations, like embedded symbols and signs. Thereby the level of automation depends on the complexity of an instruction and the degree of expectations driven by top-down processing. Thus, the perception of a simple instruction, as shown at (a), will significantly benefit from an increased bottom-up process and can be more automatically processed if a high degree of expectations is supported, as in this case. This will change as soon as the instruction becomes more complex, as shown in (b). In such cases the features are not sufficient to trigger a rapid bottom-up processing. Hence, the visual perception of this instruction relies on astronauts' knowledge and experiences gained during their payload training. Depending on how often this instruction was practiced before, a high degree of expectations leads to top-down processing that contributes to automation, which is unlikely in this case, because of the minor strength of the material caused by low frequency or less recent usage (see section 3.3.1, *Long-Term Memory*). It rather will drive the comprehension of the sentence by word-wise perception. Thus, it can be noted that receiving a payload instruction is influenced by both bottom-up and

top-down processing, whereby the effect of the bottom-up process will be decreased by an increased complexity of the information resulting in a predominant effortful top-down process. Thereby the perception will change in comprehension while the level of expectations is low, and is particularly low in cases of off-nominal situations implying new tasks, not ground trained, where complex instructions need to be learned just-in-time. Thus, the effort astronauts make to percept, comprehend or learn a text-based instruction using the PODF standard may induce an overloading of their working memory.

"sw RACK POWER → OFF" (a)

"Screw Support Jig L and R captive screws (four per jig) to Thermal Housing frame (four per jig) to Thermal Housing frame (Ratchet Wrench 1/4", 1/4" Bit Holder 1/4" drive, BIT M6 XZN [...]" (b)

Visual Aids: The current standard for compiling a payload procedure and its presentation permits only the utilization of static two-dimensional visual aids, such as tables, diagrams, schematics or pictures. Aguzzi & Lamborelle (2012), both ESA payload instructors, analyzed and discussed the state of current standards. They found out that "[..] visual aids in procedures have been less targeted by the guidelines and standards, with no clear criteria for quality, presentation or use" (p. 1) and stated that these guidelines "[...] do not enough support the production of quality and efficient visual aids" (p. 8). Besides outlining their limitations, they have also predicted that visual resources will become more complex and dynamic in the future, whereby the application of videos, 3D content and animations is conceivable. In conclusion they stated that "visual aids are already today an important part of the procedures used on board the ISS" and that "they will still be used in the future and will become more complex with the spreading of video clips and 3D visuals" (p. 8). Based on these statements, it should be mentioned that the current procedure standard and its iPV interface are not designed to handle such resources, but may be adapted or extended with advanced user interfacing. It seems far more important that using such complex visual aids increases the risk of inappropriate presentation, which will be discussed in the following.

Presentation of Visual Aids: If a text-based instruction is complemented by one or more visual aids, an efficient understanding depends on appropriate presentation. While considering the design principles for instruction presentation suggested by the cognitive load theory (see section 3.3.4), surveying some payload procedures has revealed some adverse design-related effects. This has pointed out that using a split-attention format is common practice, whereby some efforts have been made to reduce this effect by close presentations of the instruction and its reference sources, as well as by labeled annotations inside a reference, but not reflecting the meaning of the linked instruction. It has been also proven that PODF procedures also induce the redundancy effect, which is generated by multiple redundant sources, mediating the same visual informational value. Thereby its effect size depends on the level of isolation and understanding. Both, the split-attention and the redundancy effect, can imply an inappropriate presentation of the visual material and therefore may induce a high extraneous cognitive load. Besides an increased extraneous load of the working memory, PODF procedures can also cause a high intrinsic cognitive load. Although the PODF standard provides a manageable amount of syntax and rules, which can be learned in isolation, its element interactivity can be very high, especially in case of more complex instructions, whereby the astronauts' prior knowledge and experiences is required, as has been already discussed above for text-based instructions. While the cognitive load caused by the intrinsic nature of payload procedures is fixed and not alterable, a high extraneous cognitive load unnecessarily increases the load of the limited working memory, which can be reduced by a more appropriate presentation design.

User Analysis

Payloads are exclusively operated by astronauts, who constitute an outstanding and rare population of users. From 2000 until now, only 221 people have visited the ISS (NASA, 2015a), of whom seven were tourists. Thus, 214 astronauts and cosmonauts have been eligible as potential users for payload operations, which represent a small amount, in comparison with the world's population. To meet the high demands on working space and to fit best for this outstanding job, a strict and careful astronaut selection (see section 2.2) guarantees a high level of perceptual, cognitive and psychomotor skills. Once candidates are selected, they will complete a multi-annual education program (see section 2.3.5), which includes miscellaneous trainings and practices ranging from basic to mission specific trainings. Due to their above-average capabilities and expertise, astronauts are most suited for this working environment and represent a homogeneous user group best fitted to operate ISS payload.

Training Analysis

The current payload training (see section 2.3.5) is organized in a complex process and typically carried out by experts and payload instructors. Thereby astronauts are coached in payload maintenance and mission specific tasks according to their level of qualification, and should be able to list potential hazards and the relevant payload components to ensure its safety. Such training usually takes place in front of the payload training model, while the instructor provides the lessons by means of slideshow presentations. Because it is very demanding to memorize all instructions needed for several payload tasks or remembering all components of the payloads in question, the task performance mainly relies on the ability to understand and handle the related procedures. Independent of the sufficiency of its practice, a more important fact is that the current training method cannot coach astronauts for unexpected events. Regarding this, payload instructors have critically analyzed the ESA payload training and note that “for new tasks that were not specifically ground trained, some on-orbit practice may be required” (Salmen et al., 2011, p. 3). Beyond the training to assist predefined tasks, it can be assumed that there is a significant need to prepare astronauts for off-nominal situations on-orbit, similar to just-in-time training. Rather this lack of preparation for unforeseeable payload tasks seems to be a shortcoming for the existing guidance system.

Environment Analysis

When carrying out function and task analysis of payload operations, it is important to consider that the complete HMS is embedded within the larger context of the working environment onboard the ISS (see section 2.1). Besides limited space to live and work, this environment is located in Low Earth Orbit (LEO), which implies reduced gravity. Thus, it is needed to emphasize how weightlessness is affecting astronauts' performance, especially those effects, which cause changes in the human sensorimotor system (see section 2.4). When gravity is lost as reference, the vestibular system (see section 2.4.1) of an astronaut is in disorder. This vestibular disorder (see section 2.4.2) is caused by a misinterpretation of the otolith signals that no longer stimulates the vestibular sense of the vertical. This leads to conflicting sensations of the vestibular, visual and proprioceptive signals that drastically affects astronauts' well-being and interferes with their sensorimotor performance. As a result, astronauts' work in space is challenged by spatial disorientation (see section 2.4.3) and disturbed visuomotor coordination (see section 2.4.4). It is proven that the maintenance of spatial orientation have to rely more on visual and tactile sensations, because such perceptional cues are able to overwrite the otolith or participating proprioceptive signals, and to restore a faulty spatial map. Resulting consequences in visuomotor coordination can degrade the performance, for example, of hand-eye coordination by slowing down movements and inaccurate pointing. While taking these

environment-driven effects into account, the design of new interfaces needs to be optimized in their interaction modalities accordingly to ensure adequate visuomotor coordination, as required, for example, for control and input tasks.

5.1.2 Benefitting from AR

The analyses presented in the last section expose some drawbacks that could have adverse effects on astronauts' performance of payload operations. It can be summarized that the current guidance mode dominantly induces effortful top-down processing caused by the complexity of the instructional information and the task localization. It also implies a high load of working memory, which is not only caused by the intrinsic nature of text-based instructions, but also by an inappropriate presentation of the extraneous material to support the understanding of an instruction. While the intrinsic load is fixed and considered as given, the extraneous cognitive load constitutes an unnecessary load of the working memory, which is naturally limited in time and its capacity (see section 3.3.1, *Working Memory*). Additionally, perceiving and comprehending an instruction, as well as the subsequent visual search need to allocate processing resources, also referred to as attention. In general, attentional resources are limited and as emphasized by Wickens and Hollands (2000) such limitations "[...] represent one of the most formidable bottlenecks in human information processing" (p. 69) and "if adequate performance of a task demands more resources from the operator than are available, performance will break down [...]" (p. 459). Besides the fact that the visual search claims highly focused attention to combine features correctly, most attentional resources are spent on constant task switching between receiving and executing the instructions, continuously allocating alternating attention. In conclusion, it can be stated that the current interface to guide payload operations may critically increase the cost of human information processing, which is primarily caused by inappropriate resource presentation, repeated task switching, as well as by challenging task localization. Consequently, it bears the risk of failure of attention and working memory, leading to increased workload and therefore to astronaut errors that may have serious consequence causing extremely high costs and efforts. Although it can be assumed that the user population in question is best fitted by its selection and qualification, there is no warranty for the utmost efficient performance, especially not during complex and longer tasks, and even less for untrained tasks.

Hence, the results revealed from analyzing the problem space, affirm that there is a need to design an advanced user interface to guide payload operations that consider human factors. Before it is argued, why payload operations can benefit from Augmented Reality, possible countermeasures will be introduced and discussed to mitigate the drawbacks appearing from the resource presentation, task switching, as well as by task localization:

Resource Presentation: Visual resources are intended to complement text-based instructions. To reduce the extraneous cognitive load caused by an inappropriate presentation of such resources, it is required to reduce the effect size of the split-attention and the redundancy format, both driving the current design of payload procedures. To avoid the split-attention effect, a solution is given by using *integrated instructions*[55] (Ayres & Sweller, 2005), or at least by increasing the level of physical integration, for example, by closely linking the resources, as also suggested by the contiguity design principle for multimedia presentation (Mayer, 2003). While the split-attention effect applies to those sources of information, which are unintelligible in isolation, the redundancy effect occurs if multiple

[55] Integrated instructions are "instructions in which multiple sources of information are physically integrated so that working memory resources do not need to be used for mental integration" (Ayres and Sweller, 2005, p. 146)

sources are showing the same content, but intelligible in isolation. In principle, using such redundant sources of information should be avoided (Sweller, 2008). If such sources are still being used, they should at least complement each other. At this point it should be remembered that it is planned in future to use more complex and different visual sources of information, such as 3D visuals and videos. This not only runs the risk to reinforce a redundancy format, it also will become challenging to decrease the split-attention effect.

Task Switching: During payload operations, most attentional resources are allocated by the repeated focus change between the SSC-based user interface (receiving the instruction), and the payload (executing the instruction). Such focus changes do not only call for unnecessary body and head movements, they also require to rapidly shift or split attention from one task to another. Besides disturbing concentration, alternating between the SSC and the payload can imply high switching costs, significantly downgrading the performance. As for the inappropriateness of resource presentation, the problem given by procedure-induced task switching is also addressed to a split-attention issue. Consequently, it can counteract in the same way, which requires that the instructional information need to be physically integrated into the astronaut's field of view. This means that the spatial proximity of the SSC-based user interface needs to be reduced, so that the guidance interface is entirely transparent to the user, whereby a full integration is best supported by superimposed views. Such spatial proximity enables not only parallel information processing by fast bottom-up activation, it also aims to reduce alternating attention (Wickens & Hollands, 2000).

Task Localization: To detect the position of the target object at the payload, focused attention is required for task localization, which is predominantly driven by knowledge-based top-down processing caused by an effortful serial conjunction search. Often discussed not only in human factors engineering (Wickens et al., 2003), but also in cognitive psychology (e.g., Treisman & Schmidt, 1982) and neuroscience (e.g., Hopf et al., 2012), localization by serial search is a critical aspect of human performance, especially if attention is already overloaded or distracted. Besides such endogenous orientation, selective visual attention can also be directed by spatial cueing (Posner, 1980), whereby the attentional spotlight is guided by an external stimulus. Following such exogenous cues results not only in a reflexive saccade, implying less eye movements, it also evokes high bottom-up processing. The more the physical traits of the cued stimulus differ from the distractors' features, the more the preattentive feature search will be promoted. Hence, to reduce or avoid a serial conjunction search, and thus to minimize the allocation of focused attention, the localization process at the payload can be supported by target cuing, which will trigger the bottom-up-driven salience effect, also referred to as the "pop-out" effect. To benefit fully from such salience effect and to optimally support parallel processing, it is required to bring the cue as close as possible to the target object, and thus to reduce their spatial proximity, as was suggested for task switching.

By considering the effects of integrated instructions, spatial proximity and spatial cueing, it can be summarized that the perceptual-cognitive drawbacks of payload operations can be counteracted by an *integrated user interface*. Following its definition (see section 1.3), it is obvious that Augmented Reality, as an advanced user interface technology, is capable to meet these demands. Interactive in real-time, AR enriches the real world by virtual information, which is registered in 3D in the real world, placing digital objects at the right position they are assigned to. By overlaying digital information on top of the real world view, AR is not only capable to change the way instructional information is perceived, reducing the cognitive load, it also enables to guide user's attention in an exogenous manner, decreasing the allocation of focused attention. Because AR provides a high level of integration, the processing costs requested by task switching, can be significantly reduced.

Thereby the level of integration can be determined by the level of registration, ranging from presenting a text-based instruction, statically registered to the display, over showing information semantically, but not yet spatially, registered to the surrounding, to finally labeling or highlighting real payload geometry, which complies with a full registration in 3D. Therewith, AR has the greatest potential to improve the work of astronauts during intravehicular payload operations in a transparent way. As a natural complement to the common guidance interface, AR possibly enables the increase of astronaut's perception of a payload environment, and thus, more or less, approaching a perceptual user interface.

To complete the design of an AR-based guidance system, supporting payload operations, it is also necessary to compensate the loss of traditional input devices, so that the astronaut will be able to perform procedure-related actions, such as entering command-based instruction, without using the keyboard of the SSC laptop. As revealed from the environmental analysis presented in the last section, the vestibular disorder caused by weightlessness is responsible for spatial disorientation and limitations in visuomotor coordination, like hand-eye coordination, as is requested by discrete input tasks. Thus, it is required to consider such limitations while designing an AR input device, which can be efficiently operated under microgravity conditions.

5.1.3 Competitor Analysis: Lack of Human Factors

Being aware that there where previous attempts to use or integrate AR into space flight, this section intends to identify the competition of this thesis work. Therefore, all projects or studies carried out to support intravehicular spaceflight operations by AR, or at least by advanced user interfacing, constitute the actual body of works that needs to be considered here. As presented in section 4.2, surveying relevant projects revealed only a limited number of such works, especially such applying AR. Although in 1997 the NASA has already drafted the first idea for a wearable AR-like prototype in the scope of the WARP project (see section 4.2.1), only projects initiated by the ESA made it happen that AR in its usual sense was practically realized. Not intended to support payload tasks, but also an interesting and challenging work was undertaken in the scope of the CAMDASS project (see section 4.2.3), in which the design and development of an AR system was focused on the medical assistance for prolonged spaceflight missions. Except the presentation of a demonstrator, supporting cranial sonography by AR, the prototype was not sufficiently tested to verify that a non-expert can correctly handle a ultrasound probe by following visual guidance cues. Not considering AR, but linked to payload tasks were the CRUISE (see section 4.2) and mobiPV projects (see section 4.2.5). While the CRUISE experiment investigated control conditions of the iPV, whereby the procedure content was still presented on the SSC laptop, the mobiPV project was intended to enhance the display modality of the viewer, which should provide a better presentation of instructions and related multimedia resources on a set of mobile devices, such as a wrist mounted smart phone, a Google Glass and an iPad. The prototypes of both systems were field-tested onboard the ISS, and additionally the mobiPV in the NEEMO. The tests onboard the ISS corresponded to a one-man test. While the features of the CRUISE experiment could be tested by its full potential, the test of mobiPV was limited to the backup configuration, using only the smart phone and the iPad, but was successfully tested during NEEMO by multiple astronauts using the full system configuration. Although conducted with enormous efforts, the tests were predominantly focused on the system features, but not verifying its usability, for example, by summative evaluation. Most related to the problem space in question tackled in this thesis, is the WEAR project (see section 4.2.2). It was directed towards the design and implementation of a lightweight, modular and wearable AR system to support ISS

operations, like payload tasks. Onboard tests conducted by only one astronaut partially verified the operational functionality, but lacked reasonable usability tests, quantifying and qualifying the system's benefit compared to the traditional guidance interface. Another project initiated by ESA was also intended to improve astronauts' task performance, especially during maintenance operations, but was focused on shared experiences by remote collaboration provided by a virtual co-located assistance, using a shared view in AR space (see section 4.2.4). Its usability test complied with a ground-based field test, using the Columbus mockup at ESTEC in Noordwijk and a real ISS procedure. Thus, in contrast to the other projects, the test of the virtual co-located assistance revealed the highest quality of studying usability, but was disadvantaged by technical problems and lacked quantitative assessing the performance and workload level.

Consequently, it can be stated that these projects were predominantly driven by technological aspects and were not subjected to Human Factors directives, which in general call for careful analyzing the user needs and considering the concept of usability and relevant design rules (see section 3.2), especially guidelines derived from theories of human cognition, like those presented in section 3.3. Such an analysis, for example, would have revealed that the system design of mobiPV probably raises more costs of human information processing than the traditional interface, because using various devices to view the resources of an instruction, does not just mean that each device implies its own HCI loop needed to operate, but also an unnecessary load of the working memory induced by inappropriate presentation of the instructional information, even increasing the split-attention effect, especially for tasks that need to be learned just-in-time. This is not to say that to focus on the equipment is not part of Human Factors engineering, it is quite the contrary. But the operators, who must deal with the physical devices, are still needed to be considered to match their capabilities and limitations better. Because it is difficult or not feasible to cover all usability problems for the application in question in one go, usability engineering is featured by an iterative process, including heuristic design evaluation to identify possible gaps, trade-offs or countermeasures, whereby each evaluation results in a prototype, ideally providing a high level of fidelity, which can be used for subsequent usability testing. Although the objectives of the CRUISE and the mobiPV projects have complemented each other and thus the approach of an iterative design process is recognizable, the design of both, CRUISE and mobiPV, were not indicating a user-centered process. The same can be applied to the effort spent on the CAMDASS and WEAR project, while the virtual co-located assistance more closely approached the concept of usability engineering. But all projects lacked summative evaluation, verifying that the concept of AR can contribute to the task in question compared to the conventional method. Especially the WEAR project that is closest linked to the thesis's problem space, neither considered real task conditions comparing real procedure content, nor consequences of replacing the SSC-based payload interface by AR, which, for example, would also imply to compensate the loss of input modalities, usually provided by the keyboard and touchpad of the SSC laptop.

Things look rather different beyond the spaceflight sector. Reviewing related works, as was done in section 4.1.1, revealed that instructional tasks can more or less benefit from Augmented Reality, although several system designs were not motivated by practical real-world tasks or were also driven by technological aspects, which is no less important, but not sufficient to ensure a certain level of usability and acceptance. As was already discussed at the end of section 4.1.1, there was less effort made to find out which criteria were responsible for a profitable application. Apart from technical issues (e.g., display, 3D pose tracking), such criteria could be related to the type of subtasks, the level of task difficulty, the task duration, and the amount of body and head movements. Such a set of

criteria could be meaningful for developing generic approaches, not considering a specific application domain (e.g., Schwald & Laval, 2003; Pang et al., 2006). In contrast, works directed towards a practical real-world problem, as requested by this thesis, are rare and usually only possible if a specific application domain shows special interest, triggering research on it. Only then it will be possible to induce a user-centered design process, identifying drawbacks and verifying that AR contributes decisively to improvements. Although a large body of the reviewed studies strived to incorporate a high level of controlled usability testing, it lacked mostly in the degree of realism with respect to the task and environmental conditions. In ratio to the amount of challenges for instructional operations, possibly benefitting from AR, it can be said that only a small number of works studied a representative task of a practical real-world problem, such as maintenance and repair tasks in the aviation industry (Caudell & Mizell, 1992; Curtis et al., 1999; Ockerman & Pritchett, 1998; De Crescenzio et al., 2011), in the automotive industry (Reiners et al., 1999; Wiedenmaier et al. 2003; Platonov et al., 2006; Gay-Bellile et al., 2012), in the beverage packaging industry (Webel et al., 2011a, 2011b), in the oil and gas industry (Espindola et al., 2013), in the military sector (Henderson & Feiner, 2009, 2011a, 2011b), for nuclear power plants (Ishii, et al., 2007) or for picking processes in industrial warehouses (Tümler et al., 2008; Schwerdtfeger et al, 2009; Grubert et al., 2010). Other works investigated assembling tasks, such as for furniture (Zauner et al., 2003), a computer motherboard (Baird & Barfield, 1999), a water pump (Boud et al., 1999), or a ball valve (Mura et al., 2012), but mostly by using toy blocks or puzzles if a real-world task was not available (e.g., Tang et al., 2003; Robertson et al., 2008; Salonen & Sääski, 2008; Mura et al., 2012; Khuong et al., 2014; Syberfeldt et al., 2015). Although most of these works conducted corresponding usability tests, only Henderson and Feiner approached an iterative user-centered design process, investigating various aspects of problems in question, endeavored to customize the input modality, and performing appropriate field and laboratory tests (2008, 2009, 2010, 2011a, 2011b). However, and it is likely to be neglected in general, as also discussed in the end of section 4.1.1, Henderson and Feiner assessed the quality of usability only by performance indicators and have not envisaged further measures to determine the workload level.

5.2 Gaps: Challenges for AR-Assisted Space Payload Tasks

The analysis of competitive works, presented in the last section, has revealed, on the one hand, that efforts made to apply AR to spaceflight operations were primarily focused on the technology and its features, but lacked considerations of human beings and their interaction, especially with facilities, procedures and the environment as required for payload tasks. On the other hand, it has shown that research on instructional AR beyond the spaceflight sector is favoring user-centered design approaches, but often neglect sufficient attention spent on understanding the system and users' needs by iterative design processing and adequately assessing the workload level, whereby both could provide explanations why or why not AR contributed to a specific task.

In view of the problem space in question (see section 5.1.1) and the identified lack of considering Human Factors revealed from the competitor analysis (see section 5.1.3), first of all it is basically necessary to prove the concept of applying AR to payload task in an adequate manner, especially to verify whether the traditional payload interface, currently supported by the SSC laptop, can be replaced or supplemented by AR. In the case of adopting AR for ISS payload tasks, two fundamental questions were identified from heuristic evaluation. These questions are related to the visual search needed for task localization at the payload and the device placement of interfaces used for control and

symbolic input tasks. Thus, each question implies its own Human Factors challenge in respect to the task and environmental conditions, both not considered so far for AR guided space payload operations.

In the process, the research on task localization investigates in a more detailed way the spatial orientation of target cues provided by the AR interface, with the aim to promote the most efficient support during the visual search at the payload, which is in accordance to the task analysis currently disadvantaged by conjunction search, implying effortful top-down processing. Research on task localization in AR, as presented in section 4.1.2, deals with directing users' attention towards the position of off-screen targets and their subsequent detection that is closely linked to users' ability to process spatial information about their environment and orientation within, which can be classified by the used reference frame in which objects are coded. Reviewing this research revealed that there is no clear guideline which reference frame contributes more to an efficient search and it seems that this varies depending on the task conditions. Thus, the research on task localization is initially conducted under normogravity, but adapted for payload tasks.

In contrast, the research on device placement assumes that, taking AR into account to support payload tasks, a system design would exclude the SSC laptop and thus would also not allow its use for input purposes. As a result, it is required to close this gap in an adequate way, complying with the requirements demanded by AR and not using additional physical devices, interfering with a HMD, like the wrist-worn device as was used by Feiner & Henderson (2009). Presuming direct pointing to operate such devices, research on device placement in AR (see section 4.1.3) can be effectively investigated as a function of the underlying pointing movement, resulting from visuomotor coordination. As revealed from the environmental analysis, also presented in section 5.1.1, astronauts' working environment implies the absence of gravity, which drastically causes changes in the human sensorimotor system, leading to spatial disorientation and visuomotor limitations. To counteract adverse effects caused by the absence of gravity, the research on device placement must be kept in consideration of the environmental condition. Thus, this research investigates the effect of altered gravity on the placement of interfaces supporting control and symbolic input tasks.

As a consequence, the following challenging questions were identified to ensure an adequate level of task performance while applying AR to space payload tasks:

- Proof of Concept: Does it work and is it accepted from domain experts?
- More detailed: Where target cues are displayed most efficiently?
- Environment: How influential is gravity for the placement of control AR interfaces?

With focus on Human Factors, these challenges are taken up in the scope of this thesis and will be discussed more detailed in the subsequent sections to corroborate their needs and link them their related thesis chapters. Besides these research questions that were derived from heuristic evaluation, other research has been arisen as a result of the research on device placement under altered gravity. In the scope of this research two experiments were induced to compensate environmental defects in direct AR target selection by means of sensorimotoric load. These countermeasure studies are provided in Chapter 9.

5.2.1 Proof-Of-Concept: Does it Work and is it Accepted from Domain Experts?

This research is intended to close the fundamental gap left by related space flight project (see section 5.1.3), which first should clarify whether the concept of using AR can contribute to payload task by satisfying the demand of Human Factors directives. The key criteria for doing so are, on the one hand, to demonstrate its feasibility by fulfilling real task conditions within a real payload surrounding, and, on the other hand, to evaluate that AR has the potential to replace the traditional payload guidance interface in a summative way, which enables a result-driven comparison, as was not considered so far. Such evaluation needs to be underpinned by adequate feedback, verifying a certain level of acceptance, optimally rated by the intended user population. Consequently, a controlled field study is requested to prove the application of AR. Ideally, such experimentation uses real payload tasks conducted by astronauts aboard the ISS, but this requires, besides a sufficient body of pilot studies, elaborate preparation and implementation, which would go far beyond the frame of the work for this thesis. To mitigate this challenge and maintain an expert view, representative users are needed to be recruited, who are involved in astronauts' training and have expertise with ISS payloads. The same level of realism is needed to be applied to the task-related requirements, implying the usage of a real payload and its assigned tasks. Before an application of AR can be verified at all, an adequate prototype has to be designed and developed first. This prototype should not only enable to author an AR procedure for payload facilities and ODF procedures, it should also provide an AR viewer, representing an integrated-perceptual user interface that counteracts the drawbacks of the current guidance interface (see section 5.1.2). Because it is intended to assess the level of acceptance, it is required that the prototype supports a high level of fidelity to ensure that the look and feel of the system used for summative usability testing, mediates best the AR meaning. Besides feedback from the expert audience, it is important to assess the usability in an objective way. Thereby the focus lies on the performance quality, instead of showing only that AR supports faster operation, although it should also not slow down. Rather it is more important to verify that AR can better ensure the task sequence and result in a more homogeneous workflow to sustain a constant level of mental workload, which in the best case is reduced by AR support. As a result, to initially prove the concept of AR to support payload tasks and verify an improved performance quality the following requirements are needed to be fulfilled:

- high fidelity prototype to execute AR ODF procedures
- summative evaluation by field testing with high level of control and realism
- performance and workload analysis
- workflow monitoring
- preference analysis

The associated work and research is presented in Chapter 6. The prototype, providing Mobile AR for Space Operations (MARSOP), comprises an authoring and a viewing component, which are introduced in section 6.1. The design of the proof-of-concept study and its results are following in section 6.2.

5.2.2 More Detailed: Where Task Cues are Displayed Most Efficiently?

Before executing an instruction, its position at the payload is needed to be localized. Therefore the current ODF standard provides a location code, a textual information assigned to each instruction, which indicates the task area. In contrast, the task information regarding the subsequent detection of the target object and its related operational performance must be deduced from the instruction text or from image resources, where the point or action in question is marked, labeled or drawn in (see Figure 5.4). As a result, performing a payload instruction in its usual way implies the visual search by effortful top-down processing to match the point of interest and memorizing the visual hints for task performance, while AR directly cues the position of the target object by highlighted geometry and provides in-situ 3D objects to indicate, for example, the direction of an operation. Such features were already realized as a function of the initial AR prototype used for usability testing. However, this prototype is not yet supporting the off-screen localization of the task area. It only provides the location code textually at top of the HMD view, which is still claiming user's knowledge, resulting in effortful top-down processing caused by serial conjunction search (see section 3.3.3). As already presented in section 4.1.2, research on task localization in AR deals with directing users' attention towards the position of off-screen targets and their subsequent detection. This is closely linked to the users' ability to process spatial information about their environment and orientation within.

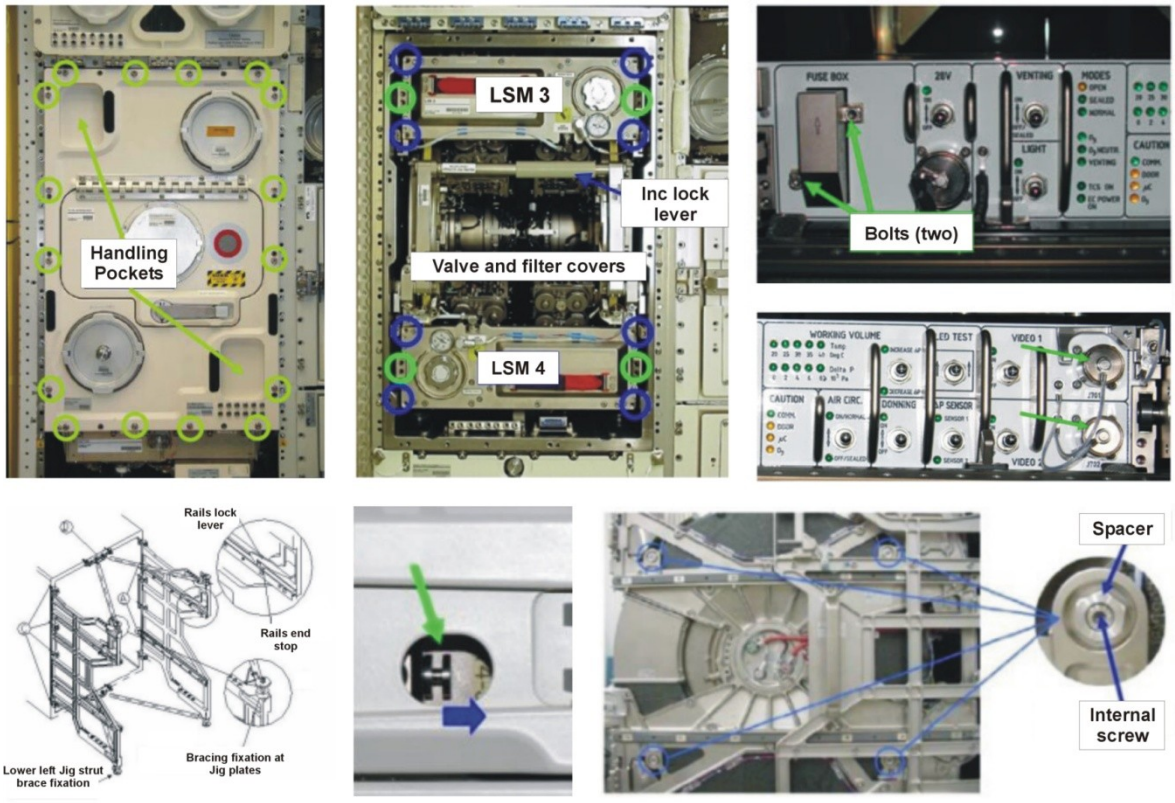


Figure 5.4: Labeled images and drawings used in ODF procedures of the ESA Biolab payload, indicating the points of interest.

Humans localize objects in their visual surrounding by extracting spatial information from their surrounding and continuously build up spatial representations of objects' positions. Thereby the visuomotor control is affected by the locations of the objects that can be defined by two different reference frames, egocentric and exocentric (also called allocentric), whereby each reference frame

has its own coordinate system (e.g., O'Keefe & Nadel, 1978; Paillard., 1991; Klatzky, 1998). While the egocentric frame of reference defines the spatial position using the body as a constant point of reference, the exocentric frame of reference provides spatial relations between the objects and depends on the external environment. Changes in human position and orientation cause spatial updating of the encoded representation that generally depends on the used reference system. Using egocentric coding requires updating of the position of each object relative to the observer and is limited by the number of objects (Wang et al., 2006). Conversely, exocentric coding only involves spatial updating of the actor's position and orientation related to this reference frame. Thus, spatial updating within an exocentric reference frame depends on the complexity of the movement, but is independent from the number of targets, even though it would increase the level of difficulty for initial learning and the cost of retrieving the object information.

As a result, the current interface to support payload tasks induces spatial updating in an exocentric way to localize the task in question, which on the one hand results in effortful matching processes and continuous head movements, interrupting the workflow, but, on the other hand, only requires spatial updating to the reference frame. Alternative approaches to display information for task localization during payload tasks may intend to use an egocentric way keeping the natural visual perception, like provided by a HMD. While wearing the display in front of the eyes, such egocentric displays maintains user's posture of the head and viewing direction, but can also vary in the spatial orientation of the information visualization as a function of the degree of registration to the captured environment, ranging from statically unregistered to full 3D registered visualizations. Displays that provide static informational content, such as a non-contact analogue head-up display (HUD), usually show task cues on an orthogonal area map, like a bird's-eye view. Although such displays avoid head movements, they force the user to change the point of view and, thus, to change the focus. In contrast, AR in its real meaning provides interactive superimposing of virtual information onto a specific spatial location in the real world keeping human's viewing direction and, thus, forces the user to an egocentric spatial navigation behavior that could maintain focused attention during payload tasks, but spatial updating is required for each object in the actual visual display, if the viewpoint has changed. While the traditional external display and an equivalent AR display can be explicitly assigned to exocentric and egocentric spatial updating, it is not clear which updating the HUD is inducing. Related research on AR target cueing for task localization (see section 4.1.2) neither answered this question nor led to a clear preference of using one specific reference frame. Although some research was able to prove that egocentric AR visualization contributes to improved target localization (Tönnis & Klinker, 2006; Robertson et al., 2008) or navigation (Medenica et al., 2011), other research has in turn revealed that exocentric task cueing, like an area map (Tönnis et al., 2005) or side-by-side visualization (Khuong et al., 2014), promotes such localization processes better. It seems that such ambiguities arise into dependence of the visualization condition or task requirements. Thus, the trade-off between egocentric and exocentric task cuing with additional consideration of HUD visualization is needed to be studied for payload tasks. Thereby the main challenge of this research is addressed to the way the efficiency is assessed. Usability testing in AR relies mainly on analyzing performance indicators and often neglects to adequately assess the workload level, which offers the opportunity to evaluate the amount of attentional resources claimed by human information processing (see section 3.3.2). If the efficiency is nevertheless considered as a function of the workload, it is usually restricted to self-reported experiences, like the rating of the NASA TLX, as was reviewed for research on instructional AR (see section 4.1.1) and on task localization (see section 4.1.2). Although sometimes the workload level is also determined by cardiovascular responses, like the heart rate variability (Tümler et al., 2008; Schwerdtfeger et al.; 2009; Grubert et al.; 2010), it is vital important

to raise the awareness that workload is a multidimensional concept (see section 3.4.3). Thus, the trade-off study claimed to show the relevance of assessing the workload in its full dimension, which covers the analysis of primary and secondary task indicators, subjective ratings as well as cardiovascular feedback.

Another aspect, which can also provide rational reason for ambiguous effects on AR task localization, deals with individual predisposition in spatial navigation behavior. While solving of spatial tasks, humans are generally able to use both reference frames, ego- and exocentric, but findings in biopsychology showed that human spatial behavior is also affected by a natural predisposition (Gramann, 2013), and proved the existence of discrete strategies in spatial navigation in the overall population (Goeke et al., 2013). Resulting from this research, humans can be classified in three groups, whereby 33.1 % can be assigned to the egocentric reference frame and 46.5 % to the exocentric reference frame, while 9.2 % are so-called switcher, which means that such humans "[...] consistently used an egocentric reference frame in the yaw plane but an allocentric reference frame in the pitch plane [..]" (Goeke et al., 2013, p. 1). The authors suggested that such distinctions in navigation strategy should be considered in studies investigating spatial navigation. This will be of special importance for AR applications where users are confronted with spatial cueing tasks and, thus, it is additionally taken into account in the scope of related research for payload tasks. The related research on task localization is presented in Chapter 7.

5.2.3 Environment: How Influential is Gravity for the Placement of AR Input Devices?

The input devices or interfaces are of special importance to handling digital content. They should be adapted to the requirements of an AR surrounding, in an adequate way. As revealed from the function and task analysis (see section 5.1.1), astronauts need to interact with the touchpad and keyboard of the SSC laptop to operate and navigate through a procedure, as well as to enter command-based instructions. To continue the use of such common input devices is not only cumbersome, it is also not in the meaning of AR. Therefore system-control techniques have to be replaced by adequate user interfaces to "avoid disturbing the flow of action of an interaction task" and to "prevent unnecessary changes of the focus of attention" (Bowman et al., 2004, p. 281). Although the spaceflight domain has strong an interest to support astronaut's work by AR (see section 5.1.3), there was no effort spent on compensating the lack of input devices so far, neither by the CAMDASS nor by the WEAR system. While the MARSOP system, introduced in this thesis (see Chapter 6), already provides simple voice commands to support procedure navigation, the question of discrete input modalities, as requested from symbolic input tasks, remains open. In accordance with guidelines for 3D interaction, it should not be assumed "[...] that speech will always be the best technique" (Bowman et al., 2004, p. 309) to support symbolic-input, which generally calls for techniques that "don't neglect user comfort" (p. 309). With respect to AR beyond spaceflight (see section 4.1.3), various design approaches were introduced to use virtual keyboards (e.g., Kölsch & Turk, 2002; Roeber et al., 2003; Tomasi et al., 2003; Harrison et al., 2011; Wang et al., 2015) or to use interactive panels to operate virtual control widgets (Poupyrev et al. 1998; Coquillart & Wesche, 1999; Lindeman, et al., 1999; Schmalstieg et al. 1999; Kohli & Whitton, 2005; Henderson & Feiner, 2008, 2010).

But beyond the question of how control or symbolic-input task can be designed and presented, the question of where such interfaces should be placed or aligned is just as important as fundamental. To ensure that such an interfaces can be easily accessed and optimally operated, it is recommended to

"use an appropriate spatial reference frame" (Bowman et al., 2004, p. 281) for its placement, such as given by nature, using the reference frame of the world or its objects, the user's head or body. Thereby world-referenced interfaces are fixed placed in the physical world using surrounding walls or objects (Fails & Olsen, 2002; Henderson & Feiner, 2008, 2010; Porter et al., 2010; Harrison et al., 2011; Marner et al., 2014). With respect to the target task conditions of payload operations, this means that such interfaces would be placed within the working environment (e.g., walls of the Columbus module) or to a planar surface of the payload (e.g., Biolab), optimally directed towards the user's viewing direction, whereby, in cases of changes, the interface needs to be replaced, taking the present task area into account. In contrast, interfaces that are referenced to a human's body or head, are detached from the surrounding environment allowing easy access any time. As also surveyed in section 4.1.3, several works designed and investigated virtual control interfaces attached to the user's body, or more precisely to their hands, whereby such interfaces can be operated either unimanually by pinching the fingers with the thumb of the same hand (Bowman & Wingrave, 2001; Piekarski & Thomas, 2003; Prätorius et al., 2014) or asymmetric bimanually by both hand, using the non-dominant hand directly as surface (Harrison et al., 2011; Dezfuli et al., 2012; Gustafson et al., 2013; Wang et al., 2015) or to hold a panel, providing the surface (Poupyrev et al. 1998; Coquillart & Wesche, 1999; Lindeman, et al., 1999; Schmalstieg et al. 1999; Kohli & Whitton, 2005; Harrison et al., 2011). Because it promotes more comfortable and faster handling (Bowman & Wingrave, 2001), it is subsequently presumed that hand-referenced interfaces are bimanually handled, whereby the interface is aligned to the non-dominant hand, while the dominant hand is the operating one. However, head-referenced interfaces are floating in the air and are positioned in front of human's head, which means that such interfaces are attached to the head-mounted display but provide sufficient space for interaction. While world- and hand-referenced interfaces commonly provide haptic feedback, head-referenced interfaces exclude haptic cues caused by the absence of a physical surface.

Interfaces for control and symbolic-input task always include the selection of an element that, in its simplest form, is conducted by direct object selection. In accordance with the guideline "use pointing techniques for selection [...]" (Bowman et al., 2004, p. 180), the research on device placement in the scope of this thesis presumes that such interfaces are operated by direct object selection accomplished by goal-directed pointing movements. But what does the interaction, resulting from operating a world-, hand- and head-referenced interface, imply for human sensorimotor coordination and thus for astronaut's performance? Although the interaction with world- or object-referenced interfaces could benefit from haptic feedback by pointing towards a physical surface, the related pointing movement could be handicapped by motoric and spatial transformations between the human body frame of reference and the world's or object's reference system. This means that operating such outside-coded interfaces requires a spatial coordinate transformation between the intrinsic and extrinsic coordinate system that can be cognitively demanding, especially during the initial phase of orientation towards the interface. In contrast, the interaction with hand- or head referenced interfaces only requires motoric and spatial embodiment inside the human body frame of reference, which did not require any spatial transformations, but nonetheless both reference types imply more physical effort than world-referenced interfaces. Thereby the absence of haptic feedback while frequently pointing towards head-referenced interfaces will be stressful and tiring, causing an increase in the physical effort. Thus, haptic cues are vital important to ensure a certain quality level for direct pointing (Lindeman et al., 1999), in particular for symbolic-input tasks via virtual keyboard, where "haptic feedback is an important component of keyboard use, so use keyboards with haptical feedback with physical buttons if practical" (Bowman et al., 2004, p. 308). However, pointing towards

hand-referenced interfaces, like hand-held panels, requires goal-synchronized bimanual handling conducted by asymmetric interlimb coordination that implies an additional postural function of the non-dominant hand in holding the interface, while the index finger of the dominant hand points towards this interface. This imposes spatial and temporal coordination of two hands that could also increase the physical effort during frequently pointing, although body-referenced interfaces are benefiting in general from the proprioceptive sense, because they are attached to limbs (Mine et. al., 1997). At all, that means that interfaces coded outside of human body frame of reference may cause more mental fatigue caused by spatial coordinate transformation, while interfaces coded inside the frame may be more impacted by physical fatigue caused by the asymmetric bimanually handling (hand-referenced) or the absence of haptic feedback (head-referenced).

Besides the fact that AR interfaces should be adapted to human needs, they also need to be optimized for the environmental conditions, especially for selection tasks, for which it is generally advised to "consider the tradeoff between technique design and environment design" (Bowman et al., 2004, p. 181). Up until now, the discussion on device placement has largely neglected the fact that payload operations are performed within the given environment of the ISS, which implies the absence of gravity. Under earth like conditions, any kind of movement is influenced by the gravitational force and thus is optimally calibrated to its frame of reference. As revealed from the environment analysis (see section 5.1.1), astronauts' work is affected by a vestibular disorder that drastically impacts their sensorimotor system, especially their visuomotor coordination. Related microgravity research (see section 2.4.4) has shown that the absence of gravity is deteriorating aimed pointing, resulting in slower movements and inaccurate pointing (Whiteside, 1961; Bock et al., 1992; Berger, et al., 1993; Smetanin and Popov, 1997; Watt, 1997; Bock et al., 2001; Mechtcheriakov et al., 2002; Bock et al., 2003; Fowler et al., 2008). But it should be mentioned that most of these works investigated the effect of altered gravity on motor coordination with respect to self-perception and, thus, visual corrections were excluded to reinforce the proprioceptive sensation. This means that such studies were performed without visual stimuli by closing the eyes or blocking the vision of the arm to avoid the effect of visual dominance overriding the otilith-mediated signals of the vestibular and proprioceptive system. However, to investigate real-world task conditions, like the operation of interactive AR interfaces, it is important to consider visuomotor performance by visual guided hand-eye coordination. Nevertheless, some studies were conducted that were intended to investigate the effect of microgravity on visual controlled hand-eye coordination (Bock et al., 2003; Fowler et al., 2008) or to compare visual conditions (Berger et al., 1993; Mechtcheriakov et al., 2002). This research has shown that such pointing movements are also affected under microgravity, but can be better adapted than not visually corrected movements. Besides excluding the visual perception channel, another characteristic, common in such research, is indicated by world-referenced pointing tasks using outside-coded targets, which in addition calls for a spatial transformation between intrinsic and extrinsic coordinates that can be the primary reason for spatial disorientation, and thus for deteriorated pointing performance. Research on task performance inside the human body frame of reference is mostly geared to investigate issues on spatial disorientation. Such research has shown that an egocentric reference system promotes normal sensorimotor performance in the absence of a gravity (Gurfinkel et al., 1993; Lipshits et al., 2005). Thus, it can be assumed that interactive AR interfaces for control and symbolic input tasks benefit more from inside-coded targets, as provided by a hand- and head-reference, than from outside-coded once, as resulted from a world-reference. On the other hand and besides that using two hands for operation will generally demand more physical effort than one-handed pointing, it has proven that goal synchronization of bimanual skills, as requested by a hand-referenced interface, is depending on the proprioceptive sense under normogravity

(Kazennikov & Wiesendanger, 2005). Microgravity research on limb coordination, considering vestibular and spinal reflexes, reported that limb proprioception is affected by weightlessness (Lackner & Graybiel, 1981; Reschke et al., 1984; Money & Cheung, 1991; Lackner & DiZio, 1992; Lackner et al., 1992; Roll et al., 1992). Consequently, it can be expected that the visuomotor performance resulted from hand-referenced AR interfaces, which are bimanually handled, could be deteriorated by the disorganization of the biomechanical chain caused by a lack of appropriate limb proprioception that takes an important role in the control of goal-directed movements (Park et al., 1999). However, the lack of proprioception can be mitigated by the fact that the associated pointing is conducted in the body-related frame of reference, which is less affected in the absence of gravity (Fisk et al., 1993; Watt, 1997), and the fact that proprioceptive restoration of movement disorder is supported by tactile sensation (Lackner et al., 2000), which would be provided by haptic cues. In addition, hand-held interfaces will reduce head movements, which are required for orienting towards the interface, although head movements in turn can contribute to pointing accuracy (Fogt et al., 2002), which in turn is beneficial for operating outside coded interface, like the world-referenced one. Besides bimanually hand-referenced AR interface, head-referenced interfaces are also detached from the surrounding and, thus, will also benefit from pointing towards inside-coded targets, but is additionally advantaged by their unimanual handling mode. However, the lack of tactile sensation will reverse the advantages, because tactile feedback provides short-term recovery and supports proprioceptive restoring of movement disorder (Lackner et al., 2000).

In conclusion, it can be noted that *spatial coordinate transformation*, *proprioceptive limb coordination* and *tactile sensation* are indicators of the sensorimotor demand that can be used to specify the placement condition of AR control and symbolic input interfaces (see Figure 5.5). As was shown by related microgravity research, it is not clear which indicator promotes better visual guided hand-eye coordination in the absence of gravity. Thus, their resulted trade-offs between world-, body- and head-referenced interfaces are needed to be studied under altered gravity conditions. Although bimanually handled, it can be excepted that body-referenced interfaces contribute most to a reliable and fast hand-eye coordination, because such interfaces may benefit, on the one hand, from inside-coded targets, and on the other hand from tactile sensation in the absence of gravity. Related experiments in the scope of this research are presented in Chapter 8.

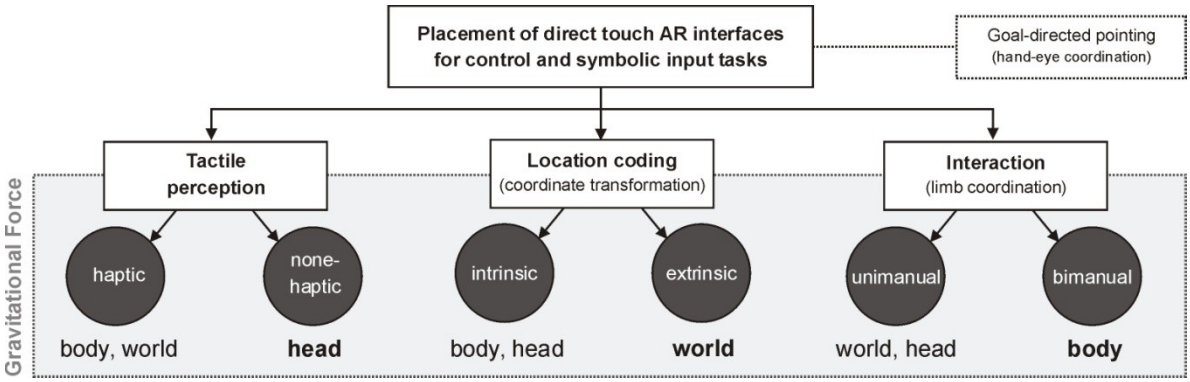


Figure 5.5: Identified sensorimotor indicators for spatial placing of AR direct touch interfaces, needed to study the trade-off between world-, body- and head-referenced placement conditions.

5.3 Restrictions and Delimitations

As emphasized repeatedly, the research conducted in this thesis aspired to a user-driven design process, iteratively considering various aspects to support payload tasks by AR, taking related Human Factors directives into account. Thus, it was not primarily geared towards technical aspects, such as high-quality HMDs and elaborated methods used for 3D pose tracking, although for the latter one, a related master theses was called and supervised. Furthermore, the authoring of AR procedures was not intended to change the ODF procedure standard, which means that related sources of information, like instructions, location codes and figures, were considered as predefined and used in its given form. Also related to ODF procedures is the restriction of only using instructions where astronauts are operating autonomously, not considering procedures or instruction, which imply ground-based communications with control centers. Other instructions, demanding input via the keyboard or touchpad, were neither considered for AR authoring nor for proof-of-concept studying, but motivated the research on device placement. Finally it should be mentioned that the research on device placement under altered gravity is limited to short-term conditions, as provided by parabolic flight (see section 8.3.1). Thus, possible sensorimotoric adaption to weightlessness under long-term conditions remain disregarded. Nonetheless, research under short-term microgravity is the best way to first prove initial effects and is the best precondition to perform subsequent ISS experimentation.

Chapter 6

Proof-Of-Concept: Prototyping and Field Testing of Mobile AR for Space Operations (MARSOP Study)

As discussed in the previous chapter, there is a demand for an extended human machine interface to assist astronauts during space payload operations. Analyzing the problem space in question revealed that the perceptual-cognitive drawbacks currently claimed on payload tasks, can be apparently counteracted by an integrated user interface using Augmented Reality. Outstanding projects previously conducted in the space flight section, were predominantly driven by technological aspects and did not consider the users' needs. Taking up this challenge, a controlled field study is requested to prove the concept of AR, verifying its capability to replace the present payload guidance interface and its general acceptance. Therefore, a summative usability test was conducted at the European Astronaut Center, which is presented in this chapter. However, this requires the prior development of a high-fidelity prototype, providing content creation and viewing of AR payload procedures in an adequate manner, whereby all aspects of a full-fledged AR system must be considered. Thus, before presenting the controlled field test, the design of the prototype of Mobile AR for Space Operations (MARSOP) is presented in the first part of this chapter, whereby related background information not explained up to here, will be clarified. Parts of this chapter were previously published in Markov-Vetter & Staadt (2013a) and Markov-Vetter et al. (2009, 2013d).

6.1 The MARSOP Prototype

To verify the concept of AR for intra-vehicular space payload operations, a high-fidelity prototype is required, satisfying the task and environmental requirements. For this purpose, it is necessary to engineer a guidance system, which provides AR experience in an appropriate manner. Besides making design decisions related to 3D pose tracking and the output device, it is desirable that the content creation is provided in a separate manner, decoupled from the actual guidance system. In the context of payload operations, content creation means to extend an ODF procedure by complementary AR resources assigned to the payload module in question, while taking the technical expertise of the end user into account. Hence, the purpose of the prototype is twofold: On the one hand, an authoring component should enable non-programmers to create an AR ODF procedure, while, on the other hand, a viewer component should provide the AR support during payload tasks, replacing the current iPV-interface. Besides appropriate visualization and registration of the AR resources, the viewer should also be able to provide interactive user interfaces for procedure navigation or to selectively explore a payload rack. Considering the aspects of a full-fledged AR system, this section gives a description of the constituent parts of a high-fidelity prototype, supporting the authoring and viewing of Mobile AR for Space Operations (MARSOP). The prototype not only supplements the work conducted by competitors (see section 5.1.3), but it also provides an

environment to perform summative usability testing (see section 6.2), investigating whether AR has the potential to improve intra-vehicular payload task, which was not considered so far.

6.1.1 3D Pose Tracking

To register and overlay the instructional content onto payload components it is necessary to continuously track the user's 3D pose[56] as accurately and stable as possible in real-time, which is a "[...] a critical component of most AR applications" (Lepetit and Fua, 2005, p. 3), and thus is one of the key technologies in AR. Such tracking generally results in a coordinate transformation from the user's position and orientation into the coordinate system of the virtual world, representing the physical surrounding that is intended to enrich. Although not accurate enough to support detailed tasks, most outdoor AR applications estimate the user's position by GPS[57] (Global Positioning System), while the orientation is obtained by sensors using an inertial measurement unit, a gyroscope, or a magnetometer, as was used, for example, by Feiner and colleagues (1997). In contrast, to realize indoor AR and provide a high accuracy, optical tracking has been established from the early beginning. Other than acoustic (e.g., InterSense IS-900) and magnetic tracking (e.g., Polhemus FASTRAK®), optical tracking aims to estimate the 3D pose of objects from a video stream relative to the image capturing device. Although optical tracking varies in its quality, depending on the given lighting conditions and benefit most from an even illumination, it is the more popular method for tracking purposes in AR. Because it is stable against varying lighting conditions, most commercial optical tracking systems, such as from ART[58] and VICON[59], uses infrared lighting, whereby, for example, at least two infrared camera are needed to detect a passive marker setup equipped with a retro-reflective surface. Such infrared tracking can attain precision in the millimetre range. Generally, vision-based approaches can be distinguished between marker-based and markerless tracking, both are provided by the MARSOP prototype and introduced in the subsequent sections. Besides the tracking algorithm itself, it is necessary to choose the right strategy to place a tracking sensor, as presented and described in Figure 6.1. Because the ISS represents a fix environment, like a research laboratory on the ground, it would be appropriate to use an outside-in strategy. On the other hand, this would cause high costs and an unacceptable amount of effort needed to install and get such a system running aboard the ISS. To be smart and flexible, the MARSOP system is intended to use an inside-out strategy, capturing the surrounding, while a camera setup is connected to the user's head mounted display. Finally, it should be mentioned that the fulfilment of the real-time aspect needs to consider that "human perception requires a rate of 25 frames per second to view a continuous movement" (Mühlhäuser and Gurevych, 2008, p.459). This means that the 3D pose needs to be optimally tracked in 40 ms. However, in AR practice, a conservative minimum for interactive systems is given by 15 frames per second (Lee and Höllerer, 2007).

---

[56] The 3D pose of an object is defined by its position and orientation in 3D space. Thus, to estimate the pose of a rigid object it is needed to cover the determination of six degrees of freedom (DOF), the number of its independent motion parameters within three DOF for translation and for rotation respectively.

[57] Tracking a 3D pose via the Global Positioning System (GPS) normally requires the view of four satellites, which is not suitable for indoor AR when using conventional receivers. Adding to this is the problem of precision. While the common GPS attains a precision of five meters (Carmigniani et al., 2011), differential GPS enhances the accuracy to one meter. Nonetheless, such discrepancies are not acceptable for AR applications, supporting instructional task at complex facilities.

[58] Advanced Realtime Tracking (ART). URL: http://www.ar-tracking.com/home/, last visit: 28.09.2016

[59] VICON Motion Capture Systems. URL: https://www.vicon.com/, last visit: 28.09.2016

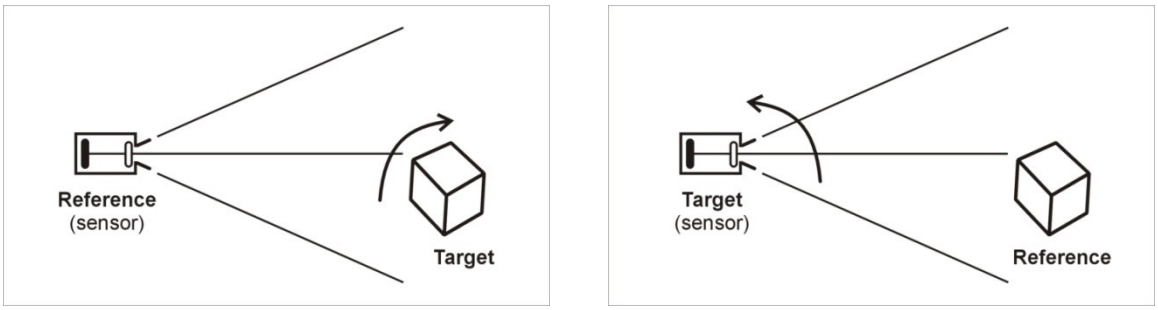


Figure 6.1: Strategies to place a tracking sensor. *(left)* **Outside-in:** The reference (e.g., optical sensor) is fix installed in the environment, while the target is attached to an object, intended to be tracked. Thus, most conventional systems, like from ART and VICON, provide outside-in tracking, which is useful for AR laboratories or VR environments. *(right)* **Inside-out:** The tracked object itself serves as reference and determines its 3D pose by detecting the target reference, that is fix installed or given in the environment. Thus, an optical sensor in an AR environment can be mobile and placed, for example, on top of the HMD.

6.1.1.1 Model-Based Markerless Tracking

Current and future work in vision-based 3D pose tracking in AR is focused on markerless approaches. Instead of using fiducial markers, markerless tracking relies on extracting, detecting and tracking natural features in the video stream, like edges, corners, optical flows or textures. One possible method of markerless tracking, as provided by the MARSOP system, is based on the idea to use the knowledge about the 3D objects in question, which are already serving the purpose of augmenting the AR scene. A further argument in favor of model-based tracking is the fact that each ISS payload, such as the Biolab, is usually delivered with a CAD model (see Figure 6.2), so that the effort to create an appropriate digital copy is not required or at least reduced.

Figure 6.2: Using a CAD model for tracking and superimposing different Biolab models.

The model-based tracking provided by the MARSOP system, is the result of a supervised master project and thesis (Millberg, 2011, 2012). This tracking algorithm was developed in line with approaches introduced by Lepetit et al. (2003), Platonov et al. (2006) and Park et al. (2011). Therefore, a keyframe database needs to be generated in a previous offline process, which contains different perspectives of the payload module and information about the characteristics of 2D features and their related 3D correspondences. To generate this information, each keyframe needs to be superimposed exactly with the CAD model, either manually or by fiducial marker tracking. After registration, a GPU-based version of the SURF[60] algorithm detects the 2D features, which are used for subsequent intersection tests with the CAD model to determine the respective 2D/3D correspondence. Finally, all collected data, such as the characteristics of the images, information about their feature descriptors and 3D correspondences, are stored based on the Extensible Mark-up Language (XML) technology (see Appendix B). However, the actual online tracking is realized by

[60] Speeded Up SURF. URL: http://asrl.utias.utoronto.ca/code/gpusurf/, last visit: 28.09.2016

approaches using tracking by detection and recursive tracking. As shown in Figure 6.3, tracking by detection is used to determine the initial camera pose, which is facilitated by the stored keyframe information. As for the keyframe generation, features are detected by the SURF algorithm and matched to the features of the stored keyframes. Then the 2D/3D correspondences of the best match are used to determine the 3D pose by applying the non-iterative Efficient Perspective-n-Point (EP*n*P) algorithm, introduced by Lepetit et al. (2009), which generally needs at least four 3D correspondences to estimate the camera pose exactly. Because false 3D correspondences are falsifying the resulted 3D pose substantially, a prior elimination of potential outliers, for example, by the RANSAC method (Fischler and Bolles, 1981), is required. The initialization process is completed by refining the camera pose using the Levenberg-Marquardt iteration (Levenberg, 1944; Marquardt, 1963), which aims to reduce reprojection errors[61] by a non-linear least square minimization. Due to a high error rate resulted from least squares methods, the square function of the residual as the projection discrepancy is replaced by a more robust Huber M-estimator (Huber, 1981), a maximum-likelihood estimator used to mitigate the effect of outliers. Only when the initialization was successful, a recursive frame-to-frame tracking, using the Kanade-Lucas-Tomasi (KLT) method (Lucas and Kanade, 1981) is initiated (see Figure 6.3). In preparation for doing so, the Shi-Tomasi operator (Shi and Tomasi, 1994) is used to detect bad features, which are determined by a feature selection criterion for monitoring feature dissimilarity. This enables to use only good features for intersection testing to find the corresponding 3D coordinates, which are then passed to the KLT tracker. Based on optical flow estimation, the KLT feature tracker is continuing to pursue all valid continuously updated 3D correspondences from frame-to-frame. Thereby the 3D pose of the previous camera frame is refined by Levenberg-Marquardt minimization using the Huber M-estimator as is done for initialization purposes. The KLT tracker continues until an abort criterion is reached (e.g., the threshold of the minimum number of correspondences is exceeded, or the reprojection error is too big). Otherwise the tracking is lost and needs to be recovered by reinitialization. An evaluation showed that the resulting online tracking algorithm is robust against occlusions and environmental changes. The frame-to-frame KLT tracking reaches an average frame rate of 18.5 fps, but may vary depending on the scene complexity and the frequency of reinitialization.

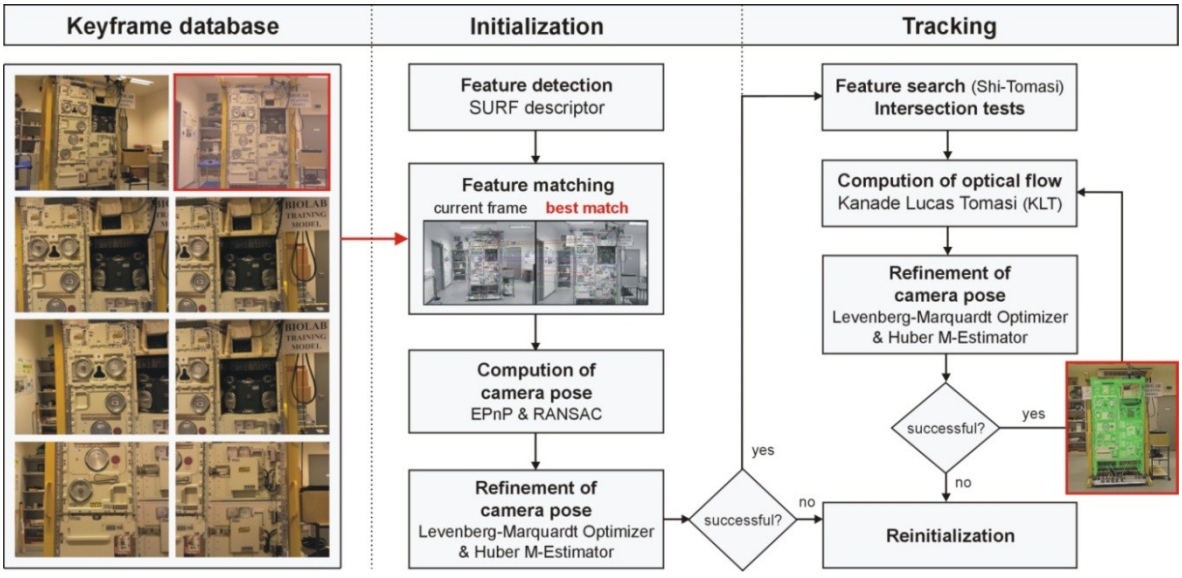


Figure 6.3: The model-based tracking algorithm used by the MARSOP system (adapted from Millberg, 2013).

[61] A reprojection error is the squared Euclidean distances between a 2D feature point and its reprojected 3D point, which is calculated by the projection matrix provided by the intrinsic and extrinsic camera parameters. This distance or error not only depends on the quality of the identified 2D point, but also on the quality of the camera calibration.

6.1.1.2 Marker Tracking

Besides markerless tracking presented in the last section, the MARSOP prototype also supports fiducial marker-based tracking for rapid prototyping as well as providing interactive user interfaces. Due to its easy use, stable pose estimation and low consumption of computing capacity, using fiducial markers is the most common tracking method used in AR and has already been established in the late nineties by the pioneering work from Jun Rekimoto (1998). Generally, marker tracking is based on the idea to use artificial squared 2D markers, bordered by a distinct frame, enclosing a black and white pattern, such as the square shaped barcode as was used by Rekimoto, although using squared shapes is not a requirement. Systems, like the commercial IS-1200 VisTracker from InterSense, for example, are using round shapes for marker constellations. Following the approach of Rekimoto (1998), Kato and Billinghurst released ARToolkit[62] in 1999, which has become the most popular free licensed tracking library until today and is also used by the MARSOP system, but utilizing osgART[63], an OpenSceneGraph[64] integration of the ARToolkit library (Looser et al., 2006). ARToolkit provides 6 DOF pose estimation by using squared planar markers, either by defining a single-marker interface or by defining an interface using a multi-marker configuration, which is useful to cover complex objects and large areas, or to simply counteract tracking losses caused by occlusions. To estimate the 3D camera pose by ARToolkit, its tracking algorithm first identifies the marker border by simple contour and corner detection processed on the captured image, which is prior converted to a binary image by thresholding operation. Then, the inner region extracted by the marker edges and vertices is normalized by the intrinsic camera parameters obtained by camera calibration needed to conduct in a preliminary offline stage. The marker's orientation is determined by the orientation of the marker pattern, which is identified by template matching. This requires that the pattern is known, either given by the system or individually trained. Based on the pattern orientation, each corner is assigned to its correct location within the marker coordinate system. By Direct Linear Transformation (DLT) a homography is given that maps the four coplanar 2D vertices from the marker coordinate system into the camera screen space. Then, by using the intrinsic parameters of the camera, its 3D pose can be recovered, which needs to be optimized by its refinement afterwards. Instead of using standard approaches, such as the Levenberg-Marquardt method, the 3D pose is refined by simple rotational alterations of the pose, whereby the angle is iteratively decreased.

6.1.2 Augmented Reality Displays

Besides choosing the appropriate tracking technique, another basic design criterion follows the question how the user visually experience the AR content. This requires to choose the right display technique for interfacing AR, which will be depending on the nature of application. In case of supporting instructional task it is important to use an AR display, which ensures a high level of task integration, but also handsfree operation. Thus, instead of interfacing AR by handheld devices (Wagner, 2007) or spatial AR displays (Bimber & Raskar, 2005), it is useful to provide AR by a head-mounted display (HMD), a device worn on the head equipped with one (monocular) or two (binocular) optic displays in front of the eyes. Generally, there are two types of HMD technologies, which can be distinguished between video see-through (VST) and optical see-through (OST), whereby each of them implicates its own advantages and disadvantages. Using a video-based approach (see Figure 6.4, left) implies that the reality is captured by one or two cameras, integrated in

[62] ARToolkit library. URL: http://www.hitl.washington.edu/artoolkit/, last visit: 28.09.2016
[63] osgART library. URL: http://www.osgart.org, last visit: 28.09.2016
[64] OpenSceneGraph. URL: http://www.openscenegraph.org/, last visit: 28.09.2016

or mounted to the display. While keeping up the user's view to create the impression to see the surrounding, the video stream is combined with digital content, which is created by the scene generator, taking the head pose into account. Besides high costs, the drawbacks of using VST HMDs include a reduced resolution of the real world view and usually a restricted peripheral vision. Nonetheless, using a VST HMD carries the decisive advantage of warranting that a real object can be fully occluded by a virtual one, which is not feasible when using an optical see-through interface. The functioning of an OST HMD (see Figure 6.4, middle) is similar to the one provided by a VST HMD, but waives the utilization of a video stream. Thus, it maintains the real world view and the natural peripheral field of vision, which is enabled by using an optical combiner, such as a semi-transparent LCD display in front of the eye, showing the superimposed virtual objects on top of the real world. Because such overlaying usually corresponds most with additive blending, the virtual objects appear semi-transparent and thus a full occlusion of real objects is not possible, which is however needed for depth perception. Whether by the video or optical see-through display, a general problem is given by latencies, which can be caused by calculating the 3D pose and the scene generation, whereby both display types are affected in different ways. While using an OST HMD can imply a delayed view of the virtual objects, a video-based interface induces a delay between the real world view and the video. Regardless how small this latency is, at the latest when the user sees a part his/her body (e.g., the hand), the proprioceptive perception is no longer consistent with the visual one. Finally it should be noted that both approaches require the calculation of an offset. Assuming that the camera setup is rigid installed within the same frame as the display, the VST only requires, if any, the calculation of an offset between the camera and display view, while the OST HMD requires an individual eye-sensor calibration (Kato and Billinghurst, 1999), which is optimally conducted each time the user puts on the HMD.

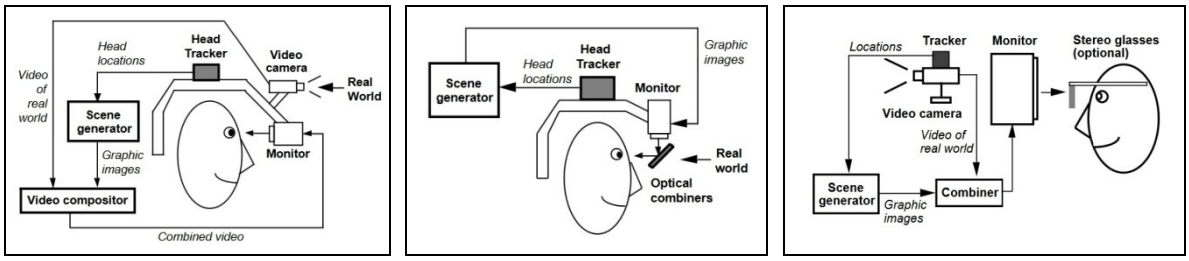


Figure 6.4: Different interface concepts to display AR content (from Azuma, 1997): *(left)* video see-through (VST) AR *(middle)* optical see-through (OST) AR, *(right)* monitor-based AR.

6.1.3 The MARSOP Constructor

To provide AR assistance supporting payload tasks, it is necessary to create prior its related content, which is also referred to as AR authoring. This requires not only the spatial alignment of a virtual payload model to the real one, but also to extend an existing ODF procedure by AR resources assigned to each procedure step. If not using a specific framework or tool, the process of generating an AR surrounding can be elaborate and time-consuming, because it requires programming skills, but also fundamental knowledge of computer graphic and interaction design. Thus, it is necessary to offer an opportunity to create AR content in an easy way, not targeting programmers, but rather application domain experts. But before describing how the MARSOP system provides authoring of AR ODF procedures, a brief overview on AR authoring is given.

6.1.3.1 Augmented Reality Authoring

Research on AR Authoring generally tackles the challenge of providing platforms for AR content creation, which vary in their complexity, calling for different levels of expert knowledge and programming skills. Thereby a very high level of knowledge is requested by using low-level libraries, such as ARToolKit (Kato and Billinghurts, 1999), which ease the development of AR application, but only empowers programmers to interface with tracking and graphic libraries. More efficient AR authoring is enabled by using frameworks, such as Studierstube[65] (Schmalstieg et al., 2002) or the MORGAN framework (Ohlenburg et al., 2004), which support, amongst others, easy access to various tracking, input and output devices. Although such frameworks still require a high level of expert knowledge, they provide a reusable software structure offering a generic functionality in an abstract manner that can individually adapted to a specific AR application. Other approaches are built upon such frameworks, like the APRIL[66] scripting language (Ledermann and Schmalstieg, 2005), but focused on increasing the level of abstractions for experts, instead of supporting application domain users. More pragmatic approaches, targeting non-programmers, provide generic tools either by utilizing commercial toolkits, like Macromedia Director[67] used by the DART[68] system (MacIntyre et al., 2003) or Microsoft PowerPoint[69] used by PowerSpace (Haringer and Regenbrecht, 2002), or provide specialized visual editors to deploy AR tutorials or ubicomp experiences, using building blocks, like the GEM collection proposed by the AMIRE[70] project (Grimm et al., 2002), or property components as utilized by the ECT interface (Hampshire et al., 2006). Visual editors are also used by immersive tools that enable in-situ content creation, for example, within a desktop 3D environment to author situated documentaries (Güven and Feiner, 2003) or directly within a desktop AR environment to provide intuitive authoring of tangible AR atop marker-based registration (Lee et al., 2004; Seichter et al., 2008). Similar approaches were proposed to generically create AR content for sequential procedures, as requested by assembly, maintenance and repair tasks. So, for example, Knöpfle et al. (2005) provided AR authoring by arranging templates within a 3D environment, whereby a template needs to be defined prior to each operation (e.g., screwing), which still requires high programming skills. Other work introduced the GUI-based SUGAR tool (Gimeno et al, 2012) that was intended for content creation based on marker-equipped photos assignable to procedure steps, which can however entails great effort spent on authoring of longer working procedures.

Regardless of the level of complexity and expertise, the approaches mentioned before have been targeted at a high level of application abstraction, although it can be expected that each application domain with its specific application scenarios has its own requirements, like the prescribed nature of ODF procedures (see section 2.3.4) and the structure of payload modules (see section 2.3.1), as the Biolab payload (see section 2.3.2). Hence, there is a need for a more customized approach to create AR content for payload tasks. Without knowledge in programming and AR backgrounds, users, such as usually designing ODF procedures (see Section 2.3.3), should be easily able to extend an existing procedure with AR content on a desktop system to subsequently provide appropriate AR elements in a payload environment interfaced by a HMD.

---

65 The Studierstube project. URL: http://www.icg.tugraz.at/project/studierstube/, last visit: 28.09.2016
66 The APRIL scripting language. URL: http://studierstube.icg.tugraz.at/april/, last visit: 28.09.2016
67 Macromedia Director (now Adobe Director). URL: http://www.adobe.com/products/director.html, last visit: 28.09.2016
68 The Designer's Augmented Reality Toolkit (DART). URL: http://ael.gatech.edu/dart/, last visit: 28.09.2016
69 Microsoft PowerPoint. URL: https://products.office.com/en/powerpoint, last visit: 28.09.2016
70 The AMIRE project: URL: http://mi-lab.org/projects/amire/, last visit: 28.09.2016

6.1.3.2 Authoring AR ODF Procedures

To be smart and avoid the deployment of capacity-intensive frameworks and commercial toolkits, as well as to focus on customized end user support, it was decided to engineer an individual solution to provide authoring of AR supported ODF procedures in a desktop-based environment. Therewith, a non-programmer should be able to easily create AR content in-situ, supplementing an existing ODF procedure in an immersive way. Besides loading a procedure and the virtual payload model, the user should be able to interactively make use of the geometry of a payload model to assign helpful descriptions, to position 3D registered resources as well as for highlighting purposes. As both payload modules and ODF procedures, are subjected to standards, a generic approach within the application domain of intra-vehicular space payload operations was envisaged. While taking these requirements into account, I developed the MARSOP Constructor, a GUI-based application running on Windows and written in C++ using Qt[71] for the 2D graphical user interface, OpenSceneGraph (OSG) to render 3D content and osgART to provide marker tracking based on the ARToolkit library (see section 6.1.1.2). Besides marker tracking, the MARSOP Constructor also provides model-based KLT tracking as described in section 6.1.1.1. To offer in-situ authoring in an immersive environment, the constructor supports different viewing modes of the payload scene (see Figure 6.5), which is presented within an integrated OSG widget. Depending on whether a camera is connected, the widget provides either a monitor-based AR environment or a 3D viewer, ready to load the payload model. While the 3D view enables authoring everywhere, without the presence of the physical payload model, the AR view offers the ultimate environment in which the MARSOP Viewer is finally working. This can be useful, for example, to label small payload components or to verify the congruence of the 3D model with the physical one. However, the constructor's functionality is provided by a stack of tabbed widgets, offering three editors to itemize the 3D payload model and the procedure in question, as well as to specify the tracking method. Besides loading, naming and describing a virtual payload model (e.g., CAD model), the *3D Model Editor* also enables to pick the model's geometry by left mouse clicking. Once geometry is selected, it can be individually named and described, but also grouped, which enables to specify different levels of a context-sensitive geometry resolution.

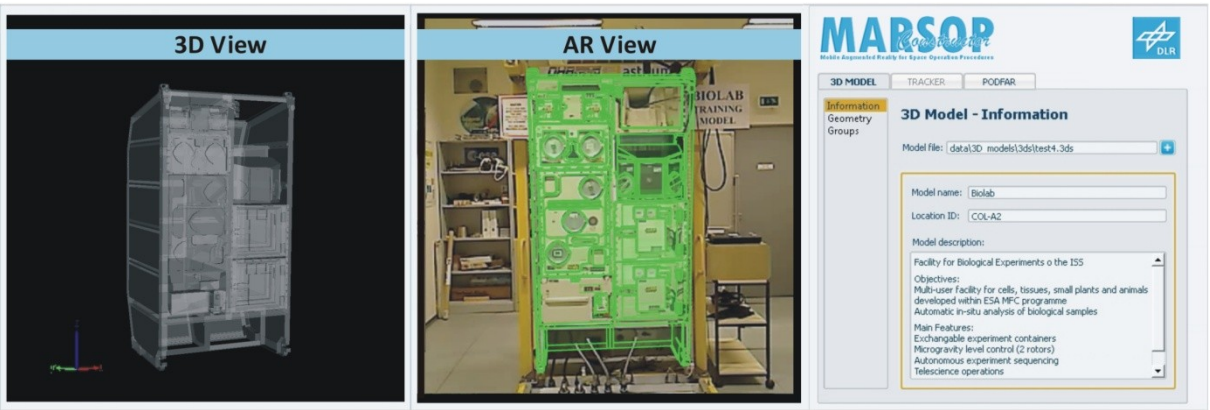


Figure 6.5: The MARSOP constructor to create AR ODF procedures. For in-situ authoring, different viewing modes are supported (3D view, AR view), whereby a payload model can be viewed in different model states (the original state, a wireframe and a transparent state). The functionality is provided by a right-justified tabbed section, comprising three editors to itemize the payload model, the 3D pose tracking and the payload procedure.

[71] Ot. URL: https://www.qt.io, last visit: 28.09.2016

Using the AR view requires to specify the tracking method needed to register the virtual payload model to the physical one. This can be done by the *Tracker Editor*, which offers the choice between using model-based KLT tracking and fiducial marker tracking, whereby both are calling for specifying the intrinsic camera parameters as well as for setting the video resolution. Using the KLT tracker additionally requires a reference to the corresponding keyframe configuration, which needs to be generated in a prior offline process (see section 6.1.1.1). It is also enabled to set optional parameters, such as the maximal number of iterations and reprojection errors, adjusting the Levenberg-Marquardt algorithm for 3D pose refinement. In contrast, the marker tracking requires the loading of a single or multi pattern configuration, as well as it offers the possibility to specify an offset matrix with which the 3D pose of the virtual model can be corrected.

The core feature of the MARSOP Constructor is the *Procedure Editor* (see Figure 6.6), which is capable to read a conventional ODF procedure in XML format resulted from the PAT tool (see section 2.3.4). Once a XML procedure is loaded, it will be parsed to break down its structure and mapped to an associated feature list, which is divided in following categories: *"General"* (procedure number, procedure title, description, assessed time period), *"Tools"* (ISS provided items), *"Items"* (special items) and *"Steps"* (sequential instructions), whereby a step will be hierarchically organized since it can hold a number of substeps. Besides its title, number and instructional text, each step is automatically tagged with a standardized location code, indicating the task area at the payload module. If a step or substep is selected, a Resource Manager, as presented in Figure 6.6, is automatically available, which provides the loading of resources as images, videos and 3D objects, as well as enables to generate labels and highlight geometry. After loading a file-based resource and activating the viewing feature, a separate OSG widget is previewing this content in 3D, whereby an image and video resource will be texturally mapped onto a 3D plane. No matter which viewing mode is active, a 3D related resource can be interactively assigned to the payload model. Like with 3D Model Editor, geometry can be highlighted by left mouse clicking, whereby the IDs of the selected geometries will be linked to the invoked instruction. Also by using the left mouse button, a textual label can be set by selecting the desired position at the virtual payload model. Afterwards the label can be lettered in the Resource Manager. Both the lettering and the 3D position of the label will be

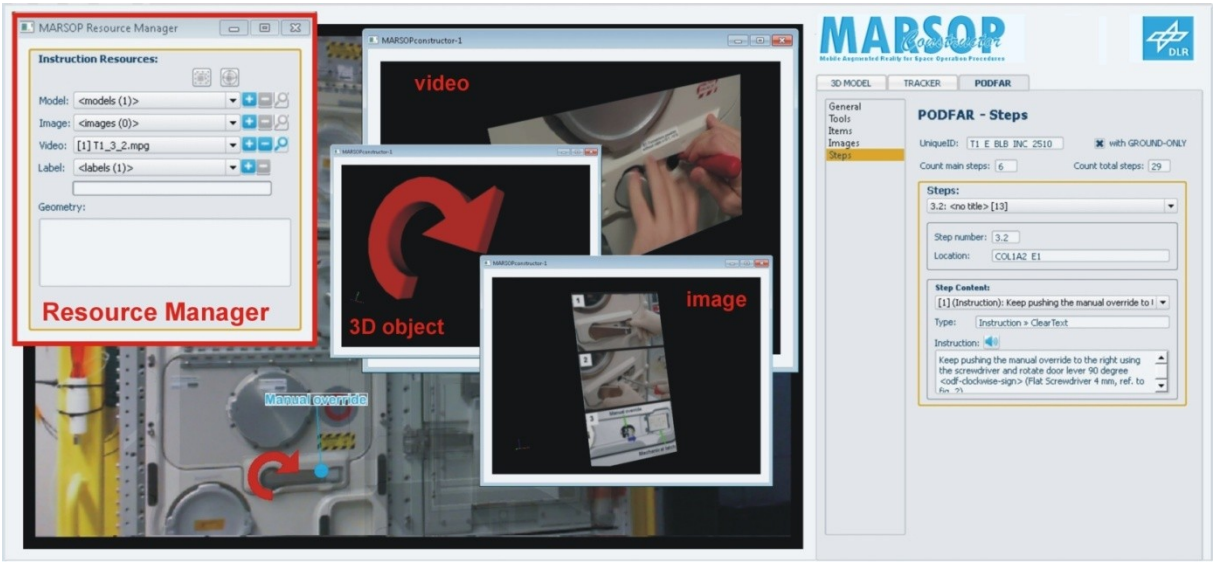


Figure 6.6: The Resource Manager provided by the Procedure Editor of the MARSOP Constructor, which enables to insert 3D objects and labels, to highlight geometry and to load an image or a video source.

linked as belonging to the instruction. In contrast, after loading a 3D object (e.g., arrow), it will be positioned fix in front of the payload model, whereby its final size and pose can be manipulated by interactive scaling, translation and rotation, using the mouse. The information linked to the instruction includes the file name and the resulting transformation matrix.

According to the data required and specified by the editors as outlined in Figure 6.7, the resulted model description, tracking configuration and AR procedure needs to be published in a proper format, capable to hold all information, which can be subsequently loaded by the viewer component. To provide an easy integration mechanism to exchange digital content between the MARSOP Constructor and the Viewer, an XML-based format was specified, which provides an AR Procedure File (ARPF). Along their attributes, main core elements, such as model, geometry, group, tracker and podfAR, are used to structure an ARPF. An example ARPF as well as an overview of the XML schema, showing the core elements, the belonging attributes and cardinalities, is presented in Appendix B.

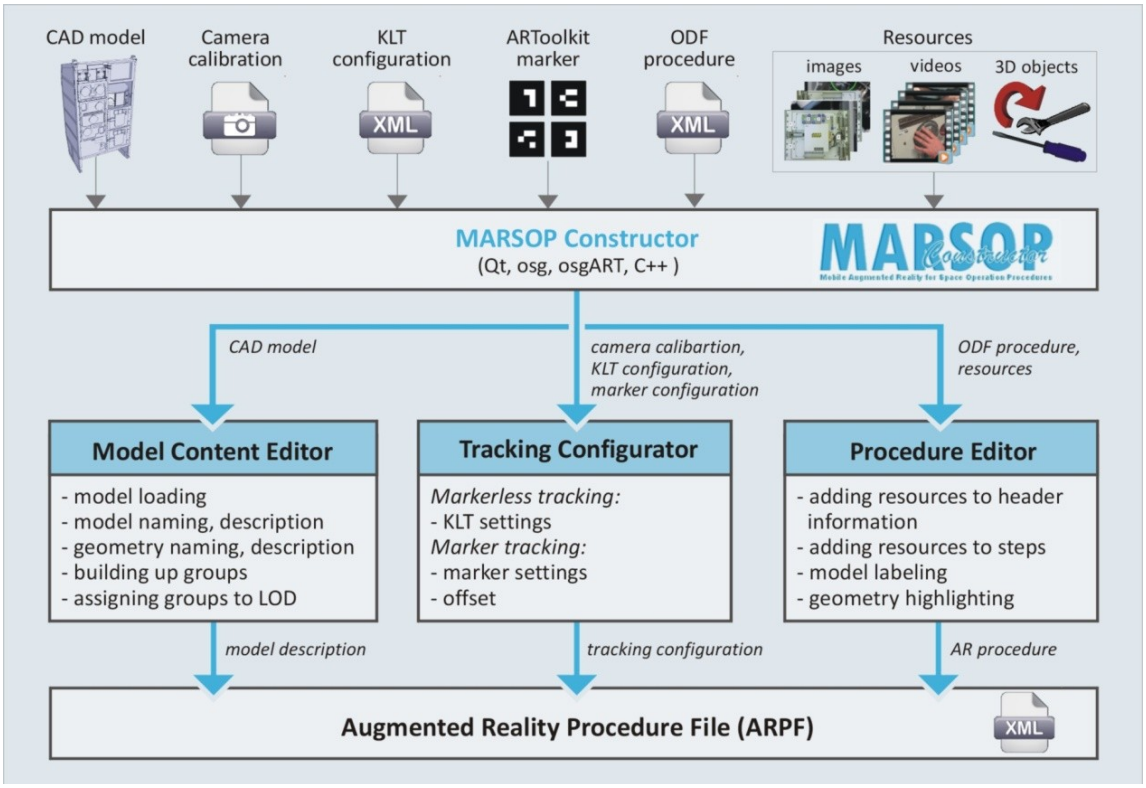


Figure 6.7: The data of the MARSOP Constructor required to create an XML-based AR ODF procedure.

6.1.4 The MARSOP Viewer

While the MARSOP Constructor, presented in the last section, enables the authoring of AR ODF procedures, a viewer is needed for presenting such AR procedures to support payload task during operations. Therefore, the MARSOP Viewer was developed in line with the constructor, providing a lightweight Windows application[72], capable to read and execute an AR ODF procedure. After loading

[72] Meanwhile, announced in May 2013, nearly all laptops available aboard the ISS have been changed from Windows to Debian 6 "Squeeze", a stable Linux distribution released in 2011. Because Qt, OSG and osgART support cross-platform software engineering, it should be possible to easily migrate between the platforms.

an ARPF, the viewer supports different display modes to interface AR, providing monitor-based and HMD-based AR simultaneously. As already mentioned in section 6.1.2, the MARSOP system enables the use of an optical see-through, but also of a video see-through HMD, whereby the required eye-sensor calibration or camera-display offset can be loaded while the viewer is starting. Once the viewer is running, the payload module in question will be tracked according to the specified pose tracker, either by model-based KLT tracking or by marker tracking. But regardless of which tracking was chosen, the viewer always enables marker tracking to provide auxiliary AR interfaces, such as the Resource Pad. Besides the matter how the real and the virtual payload are combined, and which output device is suited best, the most important feature of the viewer deals with the visualization and registration of the procedure content, fulfilling the requirements of an integrated user interface concept. In addition to visualizing and registering the content, the MARSOP Viewer also needs to consider that the user must be guided through the procedure. Therefore, this section additionally suggests a way the procedure progress can be controlled by the user, without using conventional input devices. Finally, this section ends with outlining the features not covered by the viewer, presenting its preliminary restrictions (see section 6.1.5).

6.1.4.1 Content Visualization

According to the claim of an integrated design, as discussed in section 5.1.2, AR is characterized by providing a natural user interface, capable to seamlessly embed the information required for payload operations in the real world view. Although utilizing head-mounted AR ensures the necessary spatial proximity by superimposed views to reduce or even avoid task switching, it does not yet indicate how the content of an AR ODF procedure is styled and presented, and even not just where it will be positioned, which is dedicated to the extent of the split attention effect and, thus, determines the amount of extraneous load needed to be allocated by the working memory. In the context of AR supported payload tasks, content visualization deals with the display of the specified AR resources relevant to a procedure step, whereby its level of integration is determined by its level of spatial registration in the real payload surrounding. Using the terms from Güven and Feiner (2003), the MARSOP Viewer distinguishes between *screen-stabilized* and *world-stabilized* content visualization. Assuming that each content relates to the present work scenario, then screen-stabilized content is linked to information that is semantically, but not spatially registered, while world-stabilized information is specified in the coordinate system of the physical world and, thus, meets both semantic and spatial registration, being context-aware in 3D, which is also referred to as *situated visualization* (White and Feiner, 2009). While semantic registration requires a content-related connection to the real world scenario, spatial registration describes the way a virtual object is positioned relative to its geometric registration. Therewith, world-stabilized or situated visualizations offer a close proximity in the real world coordinate system and, thus, counteract best the split-attention effect. Besides geometric registration, the presentation style of the content determines its visual appearance, whereby in case of regular AR it is striven to be in coherence with the real ambient conditions, like the illumination situation reproducible by, for example, image-based lighting (Debevec, 1998). Such photometric registration is, however, not suitable to provide instructional AR. Quite the contrary, it would annul the pop-out effect, which is necessary to trigger fast bottom-up processing, facilitating the detection of visual additives. For the same reason, the rendering of shadows was omitted, although it is an indicator for depth cueing, but promotes immersive AR experiences. The screen- and world-stabilized visualizations provided by the MARSOP Viewer are shown in Figure 6.8 and are described briefly hereafter.

Screen-Stabilized Visualization

Also referred to as head-up content, screen-stabilized visualizations are referenced to the display and are independent of the viewing direction. They are provided by adding an orthographic projection over the 3D view, which offers space for descriptions, but also for state indicators and environmental maps. Although such visualizations have a close spatial proximity to the viewed real world scenario, they are not located in the same world coordinate system and, thus, require to be refocused by eye movements but not by head nor by body movements. Furthermore, it is noted that the MARSOP Viewer solely shows head-up content, which is coherent with semantic relatedness among either the procedure content presented by textual information, or the state of the 3D tracker indicated by a tracking frame. Thereby, textual information complies with text-based pieces of information corresponding to the present procedure content and includes the location code of the related payload component, the step title and the instruction text. This information is always provided by the viewer, irrespective of whether or not the tracking system supplies valid data, which is indicated by a *tracking frame*. This frame is a distinct border, enclosing the display's field of view. The state of the tracking system is shown by coloring the frame, whereby by a green frame stand for successful pose tracking, while a red colored one is shown when the tracking is lost. This enables to provide continuous visual feedback, so that the user is aware, that the tracking behavior has a large influence on the presence of world-stabilized visualizations, which are registered to the payload model.

Figure 6.8: Screen- and world stabilized content visualization provided by the MARSOP Viewer.

World-Stabilized Visualization

Corresponding to the resources, which were specified during the content creation, the MARSOP Viewer provides world-stabilized visualizations by textual labels, highlighted geometry and 3D objects, but also by a panel used to presenting additional resources, the so-called Resource Pad. In contrast to screen-stabilized information, world-stabilized content is referenced to the physical world in the view of the display and, thus, is depending on the user's viewing direction. In the context of MARSOP, world-stabilized visualizations are not only geometrically coherent with the payload model, but also semantically with the present procedure content and, thus, they can be considered as situated visualizations. However, this applies only to textual labels, highlighted geometry and 3D objects, but not to the Resource Pad, which is subjected to a special concept of AR interfacing and,

therefore, will be separately described hereafter. *Textual labels*, such as provided by the MARSOP Viewer, belong to the class of annotations, which enables the assignment of naming and notes, or in general of descriptive metadata, used to complement the screen-stabilized instruction text. Such labels will be automatically anchored with the prior defined 3D point, which is located in the payload's coordinate system. Thereby a label will be statically repositioned relative to the referred object and linked via a straight connection line. Therewith, both the label and the referred object will have a close proximity within the real world scenario. However, the closest proximity is given by *highlighted geometry*, which, in the broadest sense, also belongs to the class of annotations. According to the geometry IDs previously assigned to a procedure step, the related components of the payload model pop out, which will trigger the salience effect, best catching the user's attention to identify the point of interest. In contrast to annotations, 3D auxiliary objects offer the opportunity to provide supplementary visualizations, supporting the meaning of the present instruction by, for example, presenting a 3D arrow to indicate the direction where an instruction must be performed. But independent of which type of situated visualization is applied, such resources need to be carefully utilized during content creation, because it carries the risk to overload the user's field of vision by a cluttered scene, which in turns reduces the salience effect striven by situated visualizations.

The Resource Pad

Besides situated visualization, as presented in the last section, the MARSOP Viewer is challenged to sustain all sources of information given by the original ODF procedure (see section 5.3). Therefore the presentation of the step title, location code and the instruction text is realized by screen-stabilized visualization, as described before. In addition to these sources, ODF procedures also provide figures, such as images, diagrams and drawings. Furthermore it is predicted that in the future visual aids will be supplemented by videos (Aguzzi & Lamborelle, 2012). Because AR output devices, such as a HMD, only provide limited space for presenting additives, the further display of figures and videos will impede the field of view. To nevertheless ensure a close spatial proximity to the payload model and to not utilize different devices, as was done by the mobiPV system (see section 4.2.5), the *Resource Pad* is introduced, which is a world-stabilized AR panel solely intended to be an output device to show such resources. As seen in Figure 6.9, the Resource Pad is equipped with a preset pattern configuration for parallel marker tracking. Depending on the actual procedure content, it shows the figures and videos available for an instruction step, but also related images for tools and items to be used. Although world-stabilized and context-aware, the Resource Pad does not correspond to the concept of situated visualization, because it is not registered to the payload model and, thus, the Resource Pad exhibits no geometric coherence with the payload model. By natural interaction, this enables that the pad can be physically positioned anywhere and relocated by the user, which implies that its spatial proximity can be determined by the user himself/herself. By this characteristic, the Resource Pad would belong to the class of tangible user interfaces (Fitzmaurice et al., 1995; Ishii and Ullmer, 1997). However, because a physical manipulation of the pad is not affecting its virtual content, it cannot really be considered as tangible AR.

Figure 6.9: The Resource Pad to view multimedia resources (figures and videos) linked to an instruction.

6.1.4.2 Procedure Navigation

Following the sequence of instructions provided by an ODF procedure means, in its simplest form, to go one step forward or, in case of mistakes or repetitions, one or several steps backward. For this purpose the current iPV interface requests input entered by the touchpad or keyboard of the SSC laptop, which would not be appropriate under AR conditions. Hence, the MARSOP Viewer needs to provide an alternative solution that at best should avoid head movements and focus changes, but also ensures handsfree operation. An obvious approach is the utilization of voice control, enabling procedure navigation by vocal commands while staying in front of the payload module, performing related tasks. To provide such voice commanding, the translation of spoken language into text is required, which is generally enabled by speech recognition. Speech recognition is widely practiced by using Hidden Markov Models (HMMs), which are stochastic models generated "[...] from known utterances and compares the probability that the unknown utterance was generated by each model" (Paul, 1990, p. 41). Because they can be trained automatically and used in a simple way, the most popular speech recognition systems are based on HMMs (Gales and Young, 2007). Instead of using built-in approaches, as provided by Windows, it was intended to equip the MARSOP Viewer with a customized speech recognizer, making simple commands, like "next" and "back", available. Therefore, the ATK[73] real-time API was used, which is a C++ layer on top of the libraries provided by the Hidden Markov Model Toolkit (HTK[74]), a portable toolkit to build, train and analyze HMMs. By using ATK, the speech recognizer provided by the MARSOP Viewer includes the need for a dictionary, a grammar and a HMM-set for the acoustic models, which needs to be prepared in an off-line process using HTK. The dictionary contains the pronunciations for each command, while the grammar defines the command loop, representing the used commands and their connections (see Figure 6.10). By doing so, the MARSOP Viewer is capable of executing voice-activated commands for procedure navigation.

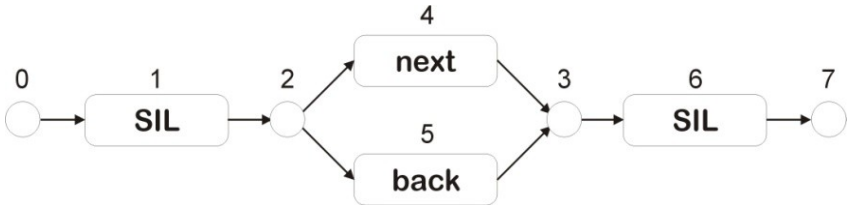


Figure 6.10: The ATK grammar used for voice commanding to enable procedure navigation.

6.1.4.3 Preliminary Shortcomings

The MARSOP prototype, as presented here, implies two major shortcomings, not closing the gaps identified in section 5.2. These shortcomings are related, on the one hand, to the lack of adequate localization support and, on the other hand, to missing input functionality. Considering the first shortcoming, localizing the target object at a payload module is currently supported by situated visualization, but only if the task area in question is in the display's field of view. However, the search of the off-screen task area remains unsupported and must still be derived from the screen-stabilized location code or from images and videos presented by the world-stabilized Resource Pad. Thus, the shortcoming of adequate task localization needs to be remedied, which is challenged by more detailed questioning where task cues can be displayed most efficiently, which in turn requests to solve the

[73] The ATK API for HTK. URL: http://mi.eng.cam.ac.uk/research/dialogue/atk_home.html, last visit: 28.09.2016
[74] The Hidden Markov Model Toolkit (HTK). URL: http://htk.eng.cam.ac.uk/, last visit: 28.09.2016

trade-off between egocentric and exocentric target cueing (see section 5.2.2). The second shortcoming, which is addressed as missing input functionality, results from the fact that the input modalities of the SCC laptop are no longer utilized by the MARSOP system. Besides procedure navigation, it is conceivable that voice commanding can also be applied to control tasks, but nonetheless there is no way to enable symbolic input, which is however needed to enter command-based instructions or make settings of discrete values. As has been already discussed in section 5.2.3, before designing an appropriate AR interface for symbolic input tasks, its spatial placement is a more fundamental issue needed to be previously solved. Assuming direct touch selection, such an interface is operated by hand-eye coordination, which can be influenced by its placement. Because weightlessness is, amongst others, affecting human visuomotor coordination, potential reference systems conceivable for interface placement need to be investigated in consideration of modified conditions of the gravity. Hence, besides field testing the preliminary prototype, as presented in next section, both task localization and input devices placement will be covered by related research in the scope of this thesis.

6.2 Proof-of-Concept Study

Before starting any specific research on issues coming up by reviewing related works (see sections 5.1.3, 5.2), the fundamental question is whether intravehicular payload tasks can benefit from utilizing AR at all. Therefore, this section presents the first research activity in the scope of this thesis, contributing to human computer interaction in the space flight sector. By utilizing the preliminary MARSOP prototype introduced in the first part of this chapter, the design and results of adequate usability testing are presented here.

6.2.1 Research Objective

To verify that the MARSOP system can replace the common payload guidance iPV interface, it is necessary to demonstrate its feasibility within a real payload surrounding and therefore requested for a controlled field study. Besides feasibility, the level of acceptance is equally important, optimally classified by feedback obtained from the intended user population. As already discussed in section 5.2.1, it would go far beyond the frame of work for this thesis to bring such experimentation aboard the ISS, which would constitute the most adequate study environment. In order to nevertheless maintain an expert's view, representative users were recruited, who corresponded either with payload instructors or those responsible for developing and carrying out payload task. To ensure a similar level of realism for the task-related requirements, it was intended to use a real payload and task associated with it. Upon a request, it was possible to conduct the study in the training facilities of the European Astronaut Center (EAC). Besides mockups of the Columbus module and the ATV, this facility provides training models of the ESA owned payloads, like, the EPM, EDR, FSL and the Biolab rack (see section 2.3.2), whereby the latter one was used in the scope of this study (see Figure 6.11). The corresponding CATIA model of the Biolab, used for tracking and visualization purposes, has been kindly provided by EADS Astrium. Consequently, instead of performing usability tests onboard the ISS using astronauts, a controlled field study in the EAC training hall was conducted, which was enriched by the advice of application domain experts and the Biolab training model. While comparing the performance of real payload tasks between the MARSOP system and the traditional guiding method, the study should provide the beneficial value and future prospect of the MARSOP design, as well as was intended to identify the user preferences and requirements (see section 3.4.1,

Preference Analysis). To verify that AR promotes a faster performance of payload tasks is beyond the scope of this study, although it should not slow down the performance. AR support is rather intended to ensure the task sequence and relieve astronauts' work by decreasing the workload. Thus the main focus of this study is on performance quality.

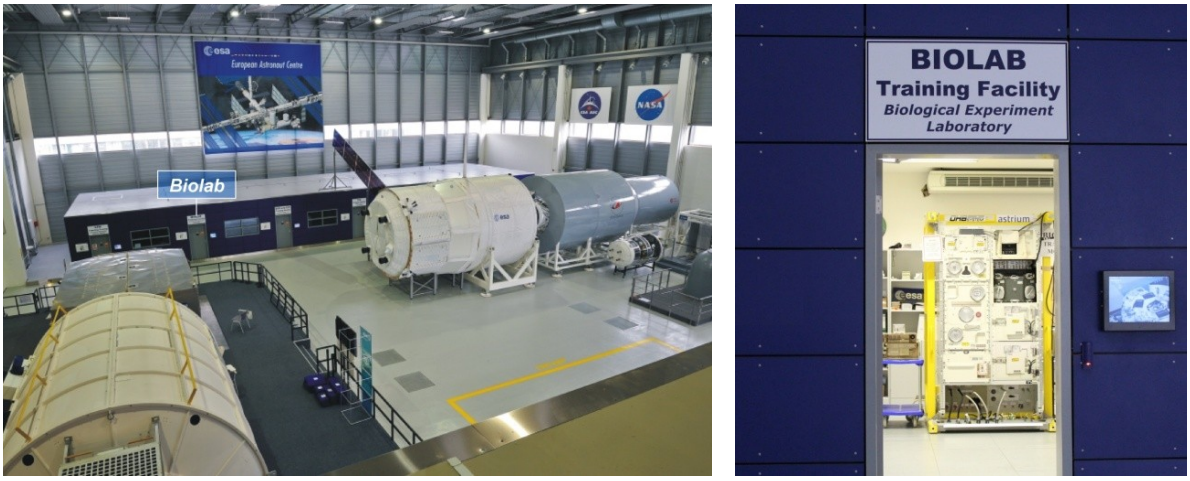


Figure 6.11: The Biolab training model provided in the training hall of the European Training Centre.

6.2.2 Apparatus

The experiment setup consisted of a data processing unit (Sony Vaio VPCF 11Z1E, NVIDEA GeForce GT 330M, Intel® Core™ i7 CPU processor, 8GB RAM, 64-bit Windows 7), a binocular video-see-through HMD, using its integrated cameras for capturing the environment, for 3D pose tracking as well as for tracking the marker-based resource pad used as tangible AR interface to display and handle the informational resources. The pad was equipped with Velcro®, so that the participant could relocate the pad easily. The HMD was the WRAP 920AR from Vuzix (see Appendix B). It was foreseen to use the integrated camera setup for KLT tracking as described in section 6.1.1.1. After performing comprehensive pre-tests to analyze the tracking quality using the KLT tracking, it was established that the field of view of the camera setup was too small to always

Figure 6.12: The experimental apparatus of the MASROP study: *(left)* The WRAP920^{TM} VST HMD from Vuzix. *(middle)* A study participant wearing the HMD. *(right)* The CATIA model of the Biolab payload.

cover the area of the Biolab and to provide sufficient features during task performance at a distance of less than 30 cm. This resulted in frequent initialization phases with 4-6 frames per second that were not acceptable for the study performance. In future work cameras with a broader field of view or the application of a fisheye lens should be tested.

Considering this limitation, it was decided to use marker-based tracking (see section 6.1.1.2) for this study. Therefore a multi-pattern board (see Figure 6.13) was attached above the Biolab rack. For the tracking of this multi-pattern a third camera was installed on top of the HMD, which was a common web camera from Logitech (Quickcam Pro) and expanded the MARSOP system to support a second tracking camera. This camera was looking up toward the pattern. To ensure that the pattern was always visible, the participants were requested to keep their head as upright as possible. Consequently, both the capture camera and the tracking camera had to be calibrated to each other. Converting the transformation matrix of the coordinate system of the tracking camera M_1 into the coordinate system of the capture camera M_2 was obtained from $M_1 = M_2 M_x$, with M_x as the calibration matrix, which in turn was resulting from $M_x = M_1^{-1} M_2$. To perform the calibration, a big multi-pattern was installed (see Figure 6.13, right), so that both cameras could capture the same pattern simultaneously.

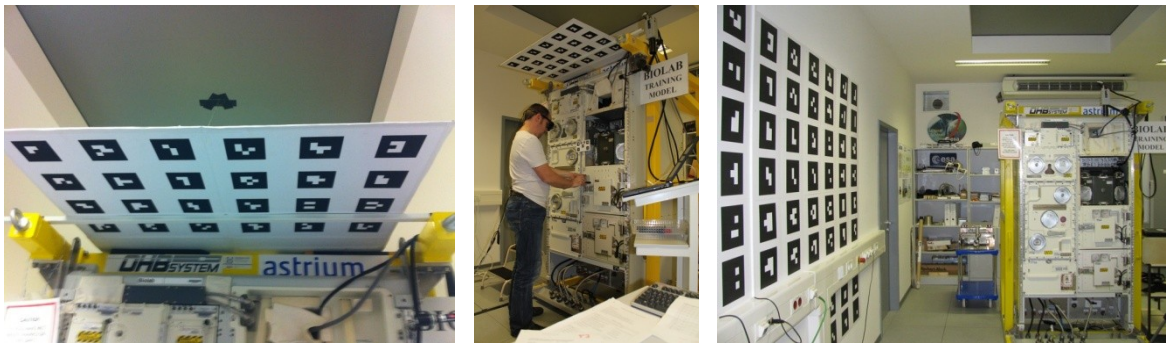


Figure 6.13: Marker tracking for performing the MARSOP study. (*left*) The multi-pattern that was used for the study. *(middle)* Particpants wokring under multi-pattern. (*right*) The calibration pattern.

6.2.3 Participants

The experiment was performed by 10 participants (3 female,7 male) with an average age of 38 (SD = 9.18). Six participants are payload instructors at the EAC and usually responsible for the payload training, such as the Biolab. Thereby astronauts are trained in accordance with their assigned levels of qualification (see section 2.3.5). Three participants are employed at the DLR MUSC, which is responsible for preparing, controlling and monitoring of maintenance, repair and experimental task for the Biolab payload (see section 2.3.3). They are also involved in the design and development of the related ODF procedures. Thus, nine participants had experiences in working at payload racks and handling of ODF procedures. It was proposed that the tenth participant should also be an ESA payload instructor, but had problems wearing the HMD. Instead, a scientific employee from the German Aerospace Center (DLR) stepped in. This participant had no experience with procedure guided operation tasks at payload racks, but was trained before. An overview of the demographic characteristics of the participants is presented in Table 6.1. It shows, among others, that seven participants have specific expertise with the Biolab rack, but differing in the numbers of related procedures, which they have already attended.

Participant	Gender	Age	Height (cm)	Dominant eye	Glasses/lens (used)	HMD exp.	ODF exp.	Biolab exp.	# Biolab ODF
S1	male	34	183	right	yes (yes)	novice	yes	yes	600
S2	male	40	160	right	no (no)	novice	yes	yes	200
S3	female	26	170	left	no (no)	beginner	no	no	0
S4	female	30	172	right	no (no)	expert	yes	no	0
S5	male	45	172	right	yes (no)	expert	yes	yes	800
S6	male	36	196	right	yes (no)	beginner	yes	yes	10
S7	male	32	177	right	yes (yes)	novice	yes	yes	500
S8	male	46	191	right	no (no)	beginner	yes	yes	500
S9	male	57	168	right	yes (no)	novice	yes	yes	25
S10	female	34	173	right	no (no)	novice	yes	no	0

Table 6.1: Demographic characteristics of participants.

6.2.4 Procedural Experiment Tasks

To satisfy the task-related requirements by using real ODF procedures (see section 2.3.4), two different procedural tasks were designed that were representative for the Biolab payload. Both tasks as were viewed by the participants are presented in Appendix C. Thereby each task starts with its title and objective description, followed by the required tools and items, whose corresponding images were displayed on the resource pad. Both tasks involved several steps and sub-steps and varied in their complexity. The first experiment task (T1) reflected the installation of an experiment container into the incubator and consisted of five major steps and 23 sub-steps. Thereby the participant should get an experiment container (EC) from the top Temperature Control Unit (TCU) and install it at the incubator's (INC) right rotor. During the second task (T2) the participants were requested to disengage the launch brackets of the TCU. This procedure involves three major steps and five sub-steps. The participant had to disengage and engage three launch brackets from the launch location to the stowage location at the top TCU. Thereby two brackets were built with four captive screws and one bracket was built with three captive screws. Because the screws differed in size and location, it was important that the procedure was performed precisely. While the experiment task T1 was more complex because of a high number of sub-tasks, the task T2 was more exhausting because of long screwing tasks. Appendix C provides a table, summarizing the selected sub-tasks of both experiment tasks. It also itemizes the used resources that were provided by the MARSOP system, showing the amount of used images and video resources, as well as the amount of labels, 3D objects and highlighted geometry that were superimposed on the Biolab payload in a 3D-registered manner. The table also displays the playing times in seconds for the video resources. The task T1 was supported by eleven videos with a cumulative playing time of 84 seconds, while four videos with a cumulative playing time of 209 seconds were provided for task T2. All videos and images were presented on the mobile resource pad.

6.2.5 Study Design

In order to evaluate the benefits, gained by the use of the MARSOP system, a summative usability test was conducted, comparing the common method with AR guidance during the performance of the payload tasks T1 and T2, both described in last subsection. In a within-subject design, each participant performed the tasks under the two different guiding conditions (PDF, AR). Thus, the experiment was framed by two independent variables. While performing under the PDF condition, the participant used the common guidance method provided by a laptop computer that was installed at

the right side of the experimentation area, just like the SSC laptop onboard the ISS. For the AR-guided condition, the participant used the MARSOP system with the video see-through HMD. To eliminate learning effects, the guidance methods and the task order were systematically randomized across all participants. To mitigate fatigue effects, each procedural task has taken no longer than 15 minutes and the experiment was divided into two sessions per participant (morning, afternoon). Before starting with the experiment, each participant performed a short demo session to train the handling of the MARSOP system. There were no time constraints for completing the tasks, but the participants were asked to try their best while performing the tasks. To ensure a certain quality level of usability testing, but nevertheless to comply with a controlled field test (see Figure 6.14, left), the study was carried out by an operator, who was responsible for technical aspects, and a neutral moderator to schedule the study, observing the participants and taking notes about participants' workflow and comments (see Figure 6.14, middle). While the operator was always the same, the study was guided and observed by two alternating moderators, both employees of the DLR. Furthermore, care was taken to keep the environment free of interruptions (see Figure 6.14, right). For a subsequent reconstruction of participants' performance, all experimentation sessions were recorded with a Sony HD Handycam®.

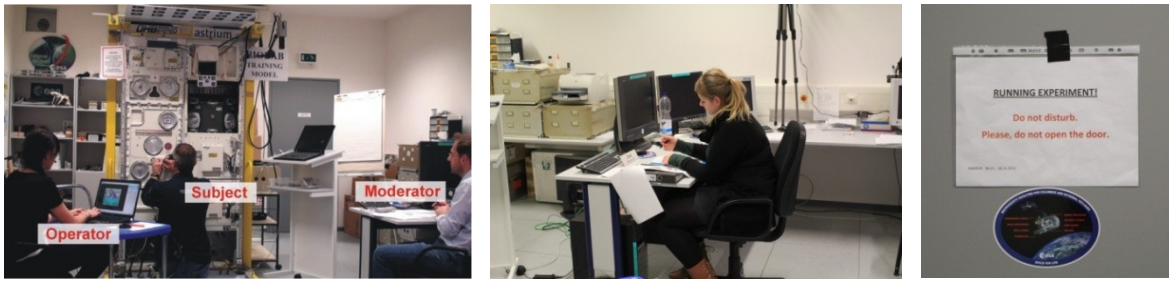


Figure 6.14: The controlled MASROP field study conducted at the Biolab training model in the EAC training hall. *(left)* The experiment was accompanied by an operator and a neutral moderator. (*middle*) A moderator, observing the participant and taking notes. (*right*) Ensuring an interruption-free zone of experimentation.

6.2.6 Measurements and Data Analysis

The usability test was intended to assess the effectiveness, efficiency and satisfaction of the MARSOP system, investigating whether the independent variables (PDF, AR) affected the dependent variables identified for this experiment (see Table 6.2). Effectiveness was evaluated by the time needed to complete a task. Under the PDF condition the completion time was measured with a stopwatch handled by the moderator. Measuring the completion time under the AR condition was done by the MARSOP system that was also extended for stepwise data acquisition. During the experimentation sessions, the moderator observed the participants and took notes about their handling and comments. However, subjective measurements were used to evaluate the efficiency and satisfaction. Therefore the participants needed, on the one hand, to rate their perceived workload by the NASA RTLX and, on the other hand, to fill in a self-metric questionnaire, which additionally included five items of the System Usability Scale (SUS) (Brooke, 1996). The self-metric part consisted of usability factors related to the HMD and tracking, as well as the visual feedback and the Resource Pad. The questionnaire's items were rated on a 5-point Likert scale ranging from *"strongly disagree"* to *"strongly agree"*. Finally the participants should indicate the preferred guiding condition and note likes and dislikes, as well as impressions. While the NASA RTLX should be rated after completing a task (T1, T2) performed under one guidance condition (PDF, AR), the self-metric questionnaire had to be filled in after completing all test series.

Dependent Variable	Description
Performance	
Completion time [s]	Time to complete a task (T1, T2), from reading the first instruction to termination of the last instruction.
Subjective Workload	
Loading Factors	NASA RTLX rating scale (see Appendix A).
Subjective Opinion and Experience	
Usability factors	Self-metric questionnaire (see Appendix C).

Table 6.2: Performance and workload measures as dependent variables of the POC study.

The data are graphically presented by distributions using box plots and by mean values plots that are marked with a confidence interval (CI) of 95 %. The rating scores of the self-metric questionnaire are presented by simple bar charts. For the task completion time, I present additionally descriptive statistics including the mean value, the standard deviation and the standard error, respectively for each guiding condition across the tasks and per task. Considering only the AR condition, the completion time of the tasks' major sub-steps is also shown by the same measures of central tendency and variability. Statistical analysis of the variance for time measures was done with a linear mixed model using SAS® PROC MIXED with Simulate adjustment, whereby all tests were rated with a significance level α of .05. Thereby the guidance conditions (PDF, AR) were compared across both tasks and per task. To identify effects between participants who are familiar with the Biolab (S1, S2, S5, S6, S7, S8, S9) with those without Biolab experiences (S3, S4, S10), I additionally compared the completion times related to this differentiation. Knowing that the low number of participants without Biolab experiences can decrease the variability, nonetheless I found it important to compare both groups, particularly because one participant of the non-familiar group did not have any experiences in reading and conducting ODF procedures and was not familiar with any ISS payloads. Furthermore, to assess whether the mean ranks related to the NASA RTLX rating scores have differed between the two samples (PDF, AR), I conducted a non-parametric test SAS® PROC NPAR1WAY with the WILCOXON option, which in case of two sample populations induces a Wilcoxon rank sum test.

6.2.7 Restrictions of the Proof-of-Concept Study

The AR condition was conducted with the video-based WRAP920™AR HMD that is associated with a number of issues. First, the HMD is limited by its resolution and the small field of view of the cameras, which seriously impaired spatial orientation and awareness of the participants. Second, the wearing comfort of the HMD was not optimal, because the participants needed to continuously readjust it, creating loss of attention and increased workload. And third, the HMD provided insufficient space between the eye and the displays to allow participants to wear their own spectacles. Thus, future work requires the consideration of optical see-through displays. An additional restriction was the non-uniform data acquisition of the completion time. While the task completion time under the PDF guided condition was manually measured with a stop watch, for the AR guided condition it was automatically done by the MASROP system, and even measured for each sub-step. Thus, no comparison of the completion time could be made between the PDF and AR procedure steps within one task. A potential solution for a uniform data acquisition could be the representation of the PDF condition in the frame of the MARSOP system. This would also allow the distinguishing between getting the task information and the operation itself. I believe that the experimental characteristics, although restricted, were appropriate for the initial verification of the concept.

6.2.8 Results

The MARSOP system worked with an average of 17.5 frames per second (SD = 4 fps). All ten participants were able to perform both experimental tasks (T1, T2) under both guiding conditions (PDF, AR). Thus, the expected number of data sets (N = 40) could be evaluated. Related to the resource pad, the participants immediately understood and accepted its meaning and if an image or video was available it was continuously used. Because all ten participants used the video resources via the resource pad, it has constantly influenced the completion time of the AR condition. Thus, a third condition (AR*) was introduced, reflecting the completion time of the AR condition, but was reduced by the video playing times (see Table 6.2). Consequently, AR* was reduced by a cumulative playing time of 84 seconds provided by eleven video resources for T1, while for T2 by 209 seconds was obtained from four video resources.

6.2.8.1 Task Performance

First of all it should be mentioned that the PDF and AR guiding condition induced entire different ways of performance respectively. Using the PDF condition (see Figure 6.15), it was noted that nearly all participants read out loud the instruction while viewing it at the laptop computer. Thereby, some participants read several steps at once or got an overview of the document initially and then executed it, which generally stood against the intention of sequential step-by-step instructions. In between checks with the PDF document forced the participants to interrupt their work to change the position

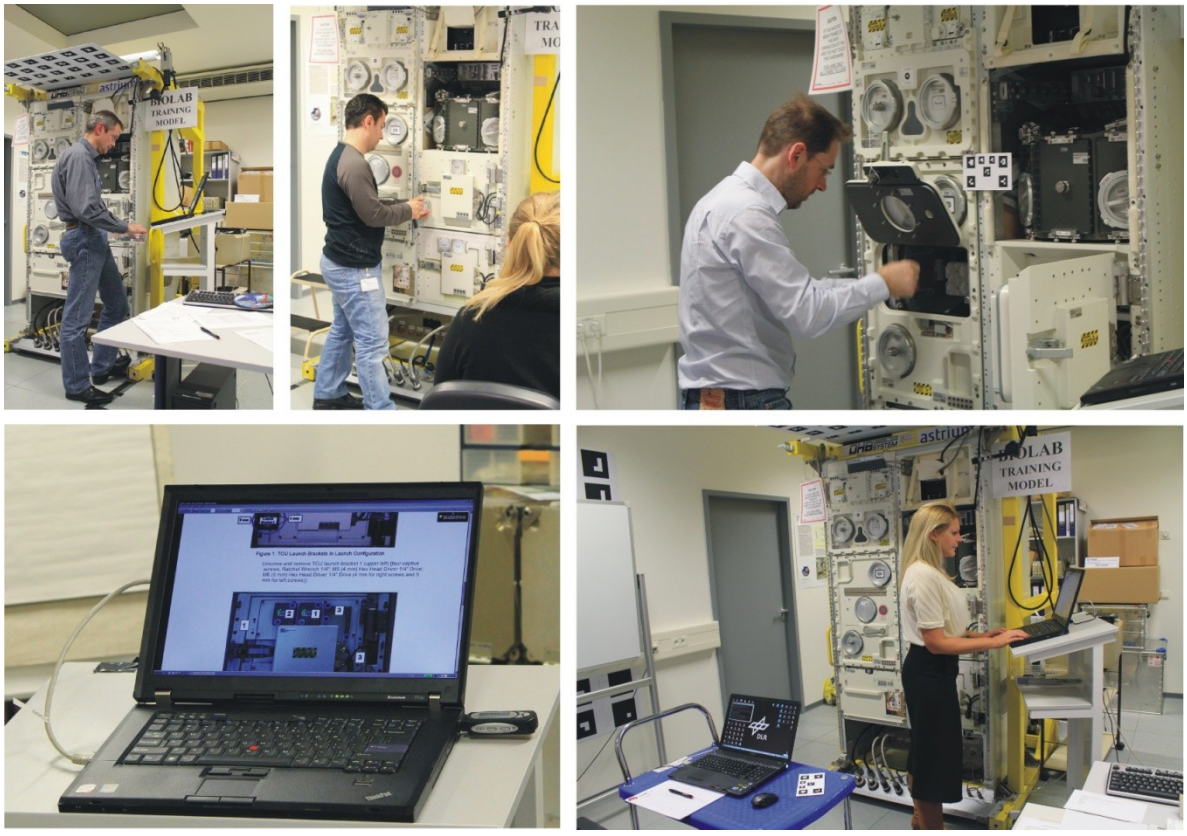


Figure 6.15: Participants performing the experimental tasks under the PDF condition using a laptop provided on the right side of the experiment area.

and gaze. There were also long searching processes for payload components and difficulties to understand an instruction while fulfilling task T1 (e.g., *"Keep pushing the manual override to the right using the screwdriver and rotate door lever 90° ↻"*). It turned out that the PDF resources, like images or drawings, were partially not sufficient or accurate. Although easier to understand, critical mistakes were done while performing the task T2. Some participants unscrewed the launch brackets in the wrong order or fixed the brackets to the wrong final position. After realizing their mistakes, it was corrected, but they were irritated and annoyed. One participant (not familiar with Biolab, but familiar with ODF procedures) did not realize the mistake and thus the procedure performance resulted in a faulty operation. As a result, it can be stated that while using the PDF condition, all participants completed the task T1 successful, but only nine participants completed T2 successfully.

In contrast to the PDF condition, under the AR condition using the MARSOP system the tasks were more fluently performed by the participants (see Figure 6.16). Using the 3D information, the participants were able to localize the target components and positions in a fast manner. Especially while performing the task T2 the participants used the 3D information and could fix the brackets at the correct position. Thereby only one participant did not use the 3D information that resulted in the same sequence error than using the PDF condition, but nevertheless the procedure was completed successfully. Hence, using the MASROP system has better supported the sequential mode of operation, decreasing the number of sequence errors. It was noted that the resource pad was a very helpful tool, and it was used by all participants. Handling and using the pad was adopted in a natural manner and first the participants played around with it, like taking in their hands, moving closer and further away or rotating it. Commonly the participant looks for the best position of the pad before starting the procedure and during the performance they relocated the pad if necessary. Especially the

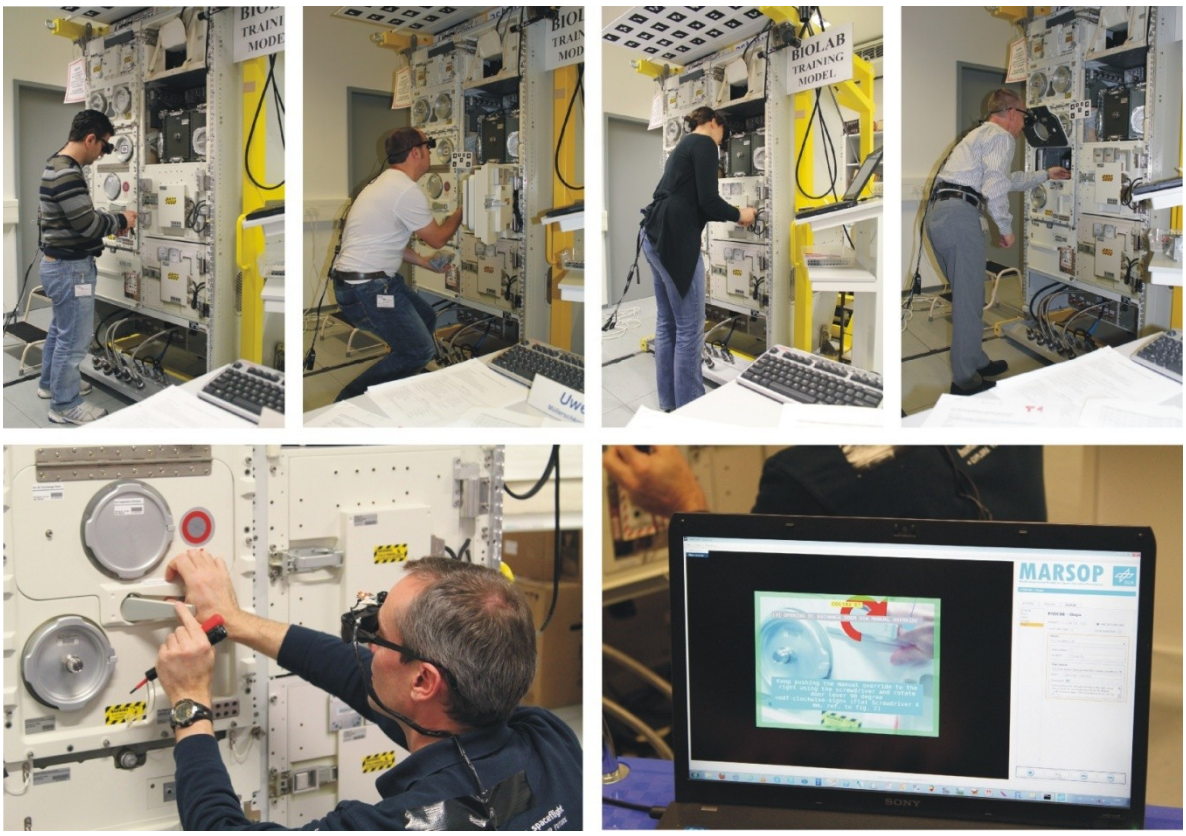


Figure 6.16: Participants performing the experimental tasks under the AR condition using the MARSOP system.

video resources were important for the participants. Whenever a video was available, the participants watched it until the end and then performed the instruction. The performance of the tasks under the AR condition resulted in nearly the same workflow: reading the instruction (HUD information), orienting at the payload (highlighted geometry, labels), watching videos, or looking at images, and finally performing the instruction. Related to the HMD, all participants needed to readjust it frequently and complained about its bad video resolution. Despite this complaint, all participants have tried to look through the HMD all the time, except during long screwing actions, whereby some have tried to look underneath the HMD. Wearing the HMD and following the request to keep their head upright, the participants were sometime forced into an unnatural posture, like, for example, instead of bending down, their preferred to kneel down.

Task Completion Time

Although it was beyond the scope of the MARSOP system to speed up the performance of payload task, it is meaningful to verify that the application of AR does not significantly interfere with the general performance. Therefore, the performance was assessed by comparing the task completion times of both guiding methods. The completion time is the time that the participant needed to perform one experimental task (T1, T2) under one guiding condition (PDF, AR). Nevertheless, it should be mentioned that the level of comparability is that high, because of the heterogeneous nature of data acquisition and of the work flow resulting from both methods. Figure 6.17 shows the distributions of the completion times for the guiding conditions across both tasks (left) and for each task (right). Because the distribution of the variance exhibited positive skewness (see Appendix C), it was adjusted by a natural logarithmic transformation. The corresponding descriptive statistics are presented in Table 6.3, showing that the mean completion times across both tasks (T1, T2) were 594.29 s (PDF), 656.05 s (AR) and 509.55 s (AR*), and thus the PDF condition was 9 % faster than the AR condition, but 17 % slower than the AR* condition. This starts to vary when distinguishing between the tasks. Considering only the task T1, the participants performed 27 % faster with PDF as with AR, but however 12 % faster compared to the reduced AR condition (AR*). In contrast, the completion time of T2 has benefited more from AR. Thereby the PDF condition was 2 % slower than AR, and even 37 % slower than AR*. But regardless of the specific task, statistical analysis revealed no significant differences between the guiding conditions (see Table 6.4), and thus no disadvantages for the use of AR.

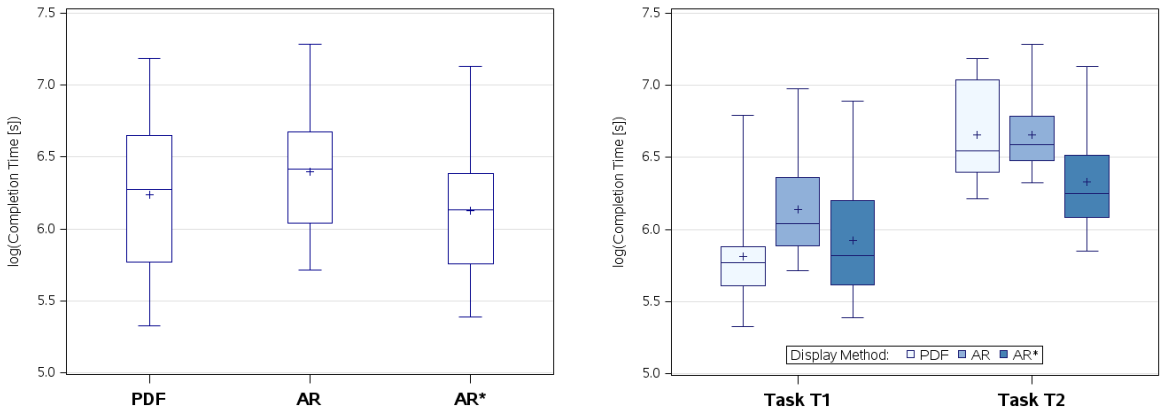


Figure 6.17: Distribution of log(completion time) of the guiding method, *(left)* across both task, and *(right)* grouped by the experiment tasks (T1, T2).

Task	N	**PDF** [s]			**AR** [s]			**AR*** [s]		
		M	SD	SE	M	SD	SE	M	SD	SE
T1	10	363.10	191.08	60.42	499.84	228.69	72.32	415.84	228.69	72.32
T2	10	825.48	305.63	96.65	812.25	269.46	85.21	603.25	269.46	85.21
total	20	594.29	343.22	76.75	656.05	291.29	65.14	509.55	261.55	58.49

Table 6.3: Measures of the completion time per guiding condition across all participants.

Task	Guidance		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*
-	AR	PDF	57	1.07	0.2908	0.4643
-	AR*	PDF	57	-0.71	0.4817	0.7023
T1	AR	PDF	27	1.80	0.0832	0.1460
	AR*	PDF	27	0.62	0.5407	0.7623
T2	AR	PDF	27	-0.00	0.9977	1.0000
	AR*	PDF	27	-2.13	0.0423	0.0764

Table 6.4: Comparison of log(completion time) between the guiding conditions across both experiment tasks and per task using SAS® PROC MIXED (pdiff=control('PDF'), adjust=simulate).

Task Completion Time by Biolab Experience

Taking into account the distinction between the participants who are familiar with the Biolab rack (S1, S2, S5, S6, S7, S8, S9) and participants who are not (S3, S4, S10), Figure 6.18 shows the corresponding distributions of the log transformed completion times for the guiding conditions across both tasks and for each task. As also shown in Table 6.5, and contrary to my expectations, the results indicate that non-familiar participants always performed slightly faster than participants who were familiar with the Biolab rack. Thereby the experienced participants performed 7 % slower than the non-familiar under the PDF condition and even 25 % slower under the AR condition. Considering only one level of knowledge, this means that the experienced participants performed 13 % faster with the PDF guidance than with the MARSOP system, while the experienced participants performs 2 % slower with PDF than with AR. But multiple pairwise comparisons of the log transformed completion times did not show any significant differences between participants' level of knowledge across both and per guiding conditions, and between the guiding condition on same stage of Biolab experience (see Appendix C). Revisiting the descriptive statistics presented in Table 6.5, it is remarkable that

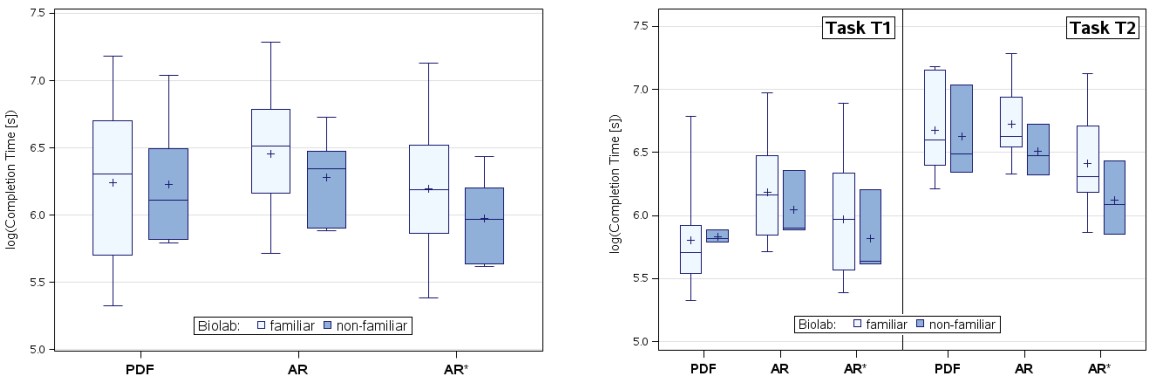


Figure 6.18: Distribution of log(completion time [s]) of the guiding conditions by Biolab experiences, *(left)* across both task, and *(right)* per experiment task.

participants who are familiar with the Biolab reveal higher variations in the standard deviation than non-familiar participants, which is indicative of a more diverse population, although the selection of users with a high level of expertise was intended to select the most representative for such tasks.

Biolab experience	Task	N	**PDF** [s]			**AR**			**AR***		
			M	SD	SE	M	SD	SE	M	SD	SE
familiar	T1	7	372.57	233.08	088.10	527.87	264.98	100.15	443.87	264.98	100.15
	T2	7	840.83	328.37	124.11	869.14	299.59	113.24	660.14	299.59	113.24
	total	14	606.70	365.89	97.79	698.50	324.33	86.68	552.00	293.98	78.57
non-familiar	T1	3	341.00	16.09	9.29	434.44	124.68	71.98	350.44	124.67	71.98
	T2	3	789.67	306.79	177.13	679.52	140.41	81.07	470.52	140.41	81.07
	total	6	565.33	313.28	127.90	556.98	179.23	73.17	410.48	135.76	55.42

Table 6.5: Measures of the completion time per guiding conditions by Biolab experiences.

6.2.8.2 Subjective Workload

The participants rated their perceived workload level each time after completing a task under the PDF and AR condition using the NASA RTLX. As shown in Figure 6.19 (left) the scores across both tasks indicate slightly higher ratings for the PDF condition, except for the physical demand and effort, which were slightly higher rated for the AR condition. But considering the loading factors per task revealed that their ratings differed (see Figure 6.19, right*)*. While for the task T1 all loading factors were rated lower for the AR condition, the rating scores for the task T2 did not show prominent variations between the PDF and AR conditions, although the mental and physical demands, as well as the effort were even rated higher for using the MARSOP system.

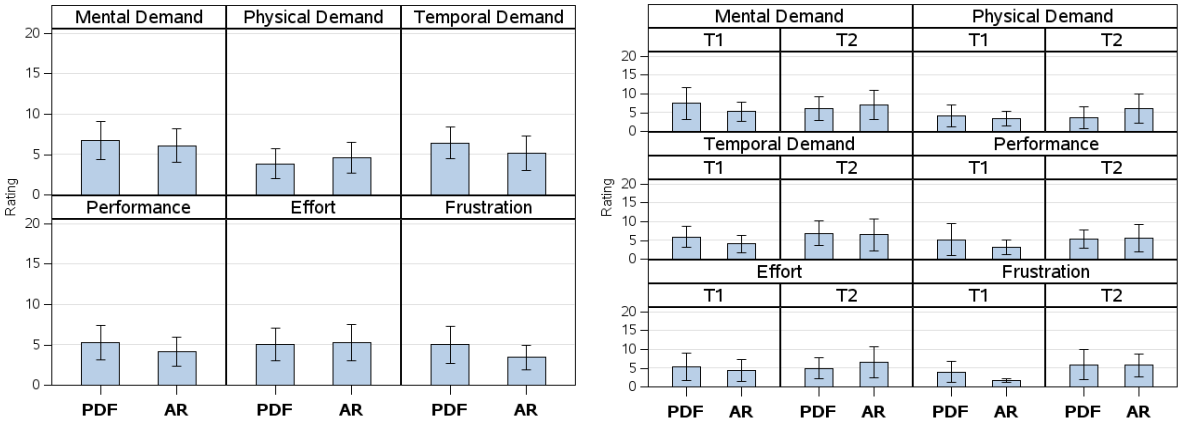


Figure 6.19: Mean ratings (with CI=95%) of the NASA RTLX scale by the guiding conditions: *(left)* paneled by the ratings items, and *(right)* paneled by the ratings items per experiment task.

Figure 6.20 shows the distributions of the Wilcoxon scores across all rated items for both tasks joined and for each single task. The rank sum test performed across both experiment tasks showed a higher mean score for PDF (M=125.49) than for AR (M=115.99), but did not indicate a significant difference, Z = 1.067, $p = 0.2858$, r = 0.0689. Same test but only performed for task T1 revealed that the workload was significantly increased under the PDF condition (M=70.95) than with using AR (M=57.91), Z = 2.016, $p = 0.0438$, r = 0.1796. In contrast, the Wilcoxon rank sum test performed for

task T2 showed a higher mean score for AR (M=60.33) than with PDF (M=54.95), but did not show a significant difference, Z = 0.871, $p = 0.3837$, r = 0.0816.

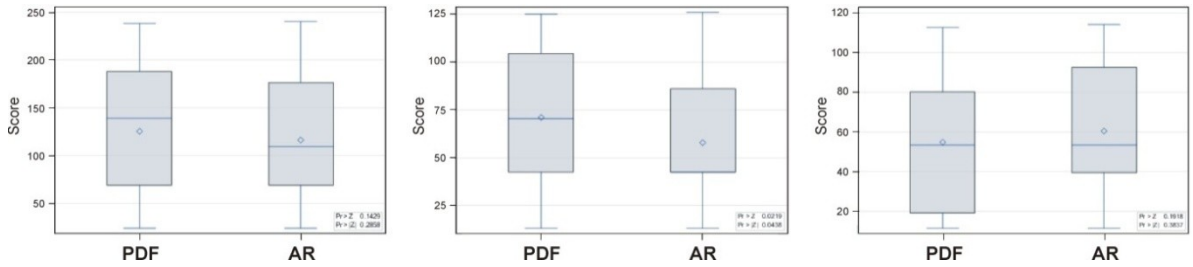


Figure 6.20: Distribution of the Wilcoxon scores by the guiding conditions of the NASA RTLX across all ratings items: *(left)* for both tasks, *(middle)* for task T1, and *(right)* for Task T2.

6.2.8.3 Subjective Experiences

The final post-questionnaire (see Appendix C) should reflect participants' opinion related to five categories, as there are HMD, tracking, system's visual feedback, provided resources and controlling the system. Thereby each category provided subordinate usability factors. Figure 6.21 shows the participants' rating scores grouped by the five categories and their factors. It is evident that more than half of the participants were not satisfied with the wearing comfort of the VST HMD, which was confirmed by the moderator's observations. It was noted that nearly all participants have commented that using the VST HMD is not convenient and that the peripheral field of vision was partially missing. Furthermore, it was noticed that if the MASROP system could not contribute with beneficial information, the participant looked underneath the HMD (e.g., during long torquing of screws), which was reasoned because of the poor video resolution. The same level of agreement has confirmed that more than half of the participants were confused by seeing the environment as video images, which could be also be explained with the bad video resolution. However, the ratings of the questions related to the category for visual feedback were evenly distributed with a high level of agreement for using the 3D information provided by the MARSOP system, while the tracking frame was mostly ignored, indicating that there is no need for additional tracking feedback. Equally clear ratings were resulting from the feedback related to the resource pad, which was helpful for nearly all participants and could be handled in an easy way. Not that strong, but also clear rated were the questions in respect to the overall system and its control level. Nine participants opined that the MARSOP system was easy to use and not too complex. Seven participants agreed that the functions were very well integrated and eight participants felt very confident while working with the system in a none cumbersome way.

After rating the usability categories, the participants were requested to rate their preference, whereby eight participants preferred the MARSOP procedure guidance compared to the conventional method. Additionally the participants were requested to describe what they liked and disliked while performing the tasks with the AR system. The original comments are provided in Appendix C, omitting redundant ones. Nearly all dislike-comments were related to the wearing comfort of the VST HMD and its video quality. In contrast to that, the participants have given more feedback about what they liked, such as the superimposed information, especially the 3D information and labels, which supported easy search for locations and items. Furthermore, all subjects were amazed by the resource pad and its handling. They were equally impressed by the provided video resources, describing a task step, which was always used by all subjects. Additional comments were assigned to the mode of

navigation by voice commands and rated it easier than using the laptop. Finally the participants could write down general comments about the MARSOP system, which should have reflected their impressions and suggestions (see Appendix C). The comments related to the impression corresponded with the above-mentioned dislikes and likes. On the one hand it was confirmed that such technology is very promising to support astronauts' work, while on the other hand they repeatedly emphasized the improperness of the VST HMD and suggested to use an optical see-through HMD, which should be adjustable in distance. Other suggestions were made to introduce a hint, indicating that a resource is available, and using ear plugs to protect from loud noises.

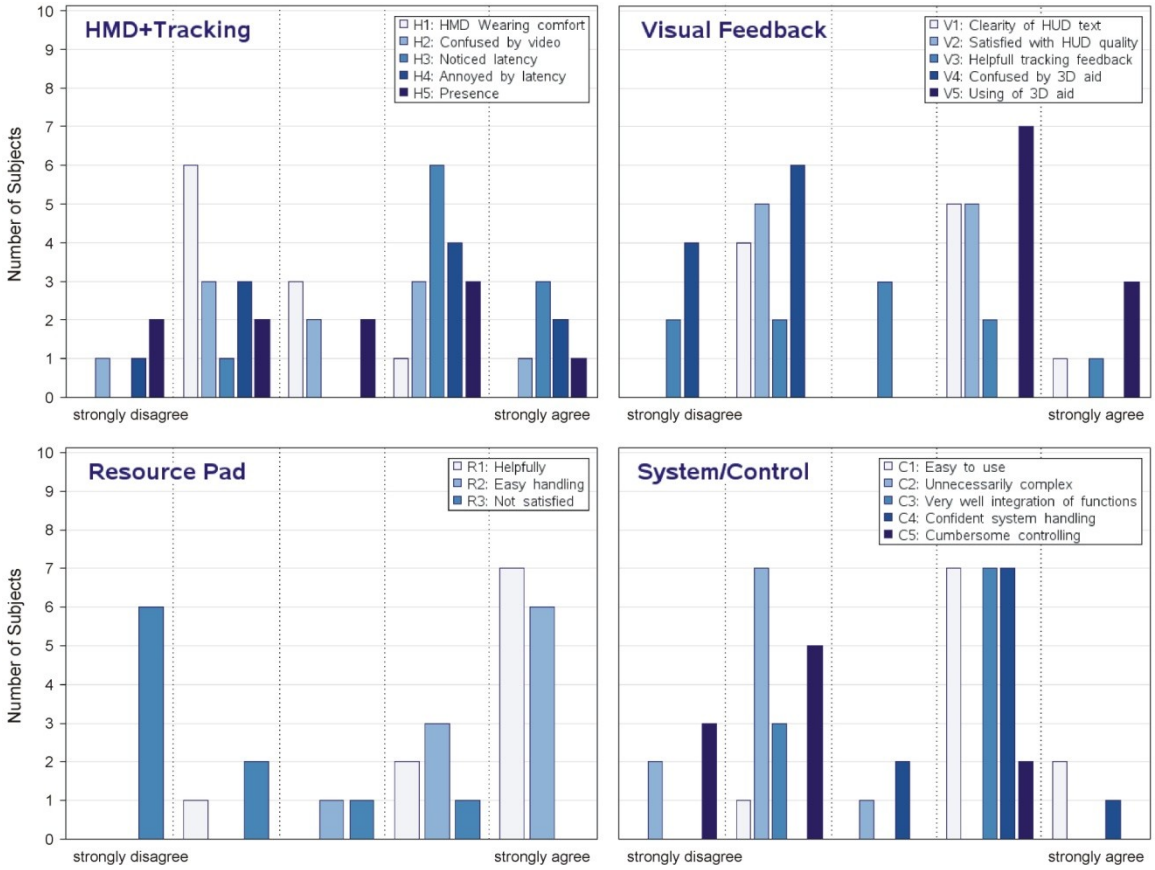


Figure 6.21: Participants' rating score of the post-questionnaire.

6.2.9 Discussion

The results showed that using the MARSOP system compared to the conventional method was applicable and accepted by the audience of application domain experts. Most of the participants were familiar with the ODF procedures and the Biolab payload and they were used to use the conventional method. While using the conventional PDF method, one experienced participant did not complete a task successfully. In contrast, the MARSOP system has contributed to succeed in task performances, while the performance using the PDF method often runs into sequence errors. The permanent change between the effective task performance and reading of the instructions also led to steady interruptions in the workflow by using the PDF. However, the performance with the MASROP system was in general more fluent, implying a higher level of attention control and supporting the maintenance of focussed attention. Thus the number of sequence errors could be reduced and, as critically important, the sequential mode of operation could be ensured.

Although barely comparable, quantitative measurements of the completion time have shown that both systems did not significantly differ from each other (see section 6.2.8.1). Considering the cumulative effect of the video resources, it should be mentioned that the task T2 was even better performed under the AR condition. On the one hand, this gain was due to the fact that this task has requested more accuracy to screw the brackets in the right order and thus this task could benefit more from the AR support, although the HMD became annoying during with long screwing tasks. On the other hand, this tasks were supported more by additional sources of information that were integrated into the participants' field of view, and thus the effect size of a split-attention format was decreased, especially by highlighted geometry and labeled annotations. By comparing participants, who were familiar with the Biolab rack with participants who were not, I found out that the completion times were nearly the same or even faster completed by the non-familiar participants, as was similar observed by Schwertdtfeger et al. (2009). This novice-expert difference may be linked to the worked example effect and the expertise reversal effect provided by the cognitive load theory on instructional design (see section 3.3.4). Thereby, the worked example effect states that learners using a worked example for the first time perform better in problem solving than learners who are experienced with the given problem. Because the MASROP system provided redundant instruction information, like the HUD instruction, the video source and the 3D-registered information (highlighted geometry, labels, 3D arrow), the lower performance of the Biolab experienced participants may also be reasoned by the expertise reversal effect. This effect indicates that learners with a high level of expertise need to allocate more working memory resources, if an instruction is presented by redundant sources of information, because such learners will mentally integrate each of them. As also suggested by Kalyuga and colleagues (2003) "an instructional format without redundant guidance is likely to be the best instructional format for more experienced learners [...]" (p. 29). Thus, it can be assumed that participants unfamiliar with the Biolab benefitted more from the redundancies, because they only integrate the sources they understood in isolation according to the redundancy effect, while the experienced participants have integrated all information, whether they wanted to or not. The same differentiation can be made related to the modality effect, which suggests that using a dual-modality presentation (HUD instruction and video source) can reduce the extraneous load of the working memory, but only for novice learners. However, when dual-modality instructions are presented to experienced learners, Kalyuga and colleagues (1999) proposed that the instructions "[...] should be presented in auditory rather than written form [...]" (p. 369). They also suggested that the different level of knowledge and experiences should be considered by a proper selection of a user-adapted instructional design.

In contrast to discussing the performance based on the completion time, the participants also rated their mental workload by the NASA RTLX to gain knowledge about the perceived effort spent on task performance using the MARSOP system in comparison to those perceived by the traditional guiding method (see section 6.2.8.2). Thereby the loading factors reflecting the mental and temporal demand, the performance and the frustration were slightly lower rated for the MARSOP system, but not significantly differed from the PDF condition. But there were different results for a task-wise consideration. Although the MARSOP system could contribute to a slightly better performance during the task T2, the participants rated their related workload higher (mental demand, physical demand, effort) or at least equal (temporal demand, performance, frustration) compared to the PDF condition, but the differences were not significantly. The higher ratings may be reasoned by the stress caused by wearing the HMD during repetitive and monotonous tasks, like the long screwing processes, where the added value, provided by the HMD, was not needed. However, the rating scores of the workload perceived during the performance of the task T1 differed significantly between the

guiding methods, whereby all loading factors were lower rated in favour of the AR condition, which indicates that fulfilling tasks with varying elements, where constant learning is involved, seemed to profit more from the support of the MARSOP system.

Besides evaluating quantitative measures, the users' opinions and preferences were identified through analyzing the post-questionnaire (see section 6.2.8.3). Overall it can be stated that the MARSOP system was accepted, reasonable to handle and predominantly preferred. Whether the superimposed information, the labeled annotations, 3D objects, the text-based HUD information or the images and video streams, all provided resources were accepted and approved, as well as have contributed to satisfying the participants. Furthermore, it became apparent that the resource pad was a strong tool to provide additional sources of information in a natural manner, holding or placing the pad where and when it was needed. It was easily adopted by all participants and was in the same way as the superimposed information or viewing a video guidance gladly accepted and appreciated. In contrast to the favored opinions, the only negative aspect was related to the wearing comfort and quality. It was repeatedly stated that the used VST HMD and its image quality as well as its limited field of view is not suitable to support payload task, and as mentioned before, especially not for those which are repetitive and monotonous, like long screwing tasks. Instead and also suggested by the participants, using an optical see-through HMD not only would maintain the natural peripheral field of view, but would also increase the degree of immersion and thus more closing the gap between the real and the virtual world, although such displays would require an additional eye-sensor calibration.

In conclusion, it can be stated that the MARSOP system is a promising approach to support astronauts work at payload racks and thus has the potential to outperform the traditional method, especially for task not trained in advance. Projects previously conducted by ESA, have already considered the application of AR, to either guide astronauts through ultrasound examinations (see section 4.2.3), or to support ISS operation in a collaborative (see section 4.2.4) or non-collaborative way (see section 4.2.2), whereby the latter one is most comparable to the work presented here. Although the AR system resulted from this project was tested aboard the ISS, it was focused on compiling and testing the technical setup (e.g., HMD and 3D pose tracking) and was neither directed towards the inherent task nor towards the user's needs, and thus was predominantly technology-driven. Complementary to this, the design of the MARSOP system resulted from a prior front-end analysis, whereby the study to proof its concept was the initial step to identify task and user related requirements. It shows that the MARSOP system is doing the same to fulfill payload tasks than the traditional method, but is capable of reducing the mental workload. By avoiding the common split-attention effect, the application of AR not only consumes less attentional resources that is needed for task switching, it also brings the required information into one field of view, supporting an improved presentation of the instructional information and thus decreases the extraneous cognitive load. While this contributes to maintain focused attention on task performance, however, focused attention is reduced by highlighted geometry and labeled annotations, cueing the visual attention during task localization. Thus, AR also prevents the allocation of additional resources otherwise needed for step-wise conjunction search and maximize the bottom-up process by the pop-out effect, which promotes the pre-attentive feature search. Besides these perceptual-cognitive benefits, it has also been proven with the MASROP system that the application of AR is accepted and predominantly preferred by domain application experts.

6.3 Summary

Operating on-board systems and the performance of experiments at payloads accounts for the major part of astronauts' everyday work on the ISS and it is important to ensure successful performances. The analysis of the problem space, presented in the last chapter, exposed that the current guidance mode increases the cost of astronauts' information processing, which could be countered by applying AR. Therefore I introduced and tested the first high-fidelity prototype of a Mobile Augmented Reality for Space Operations (MARSOP) that enables easy authoring and application of standardized payload procedures in a mobile AR environment using head-mounted displays. Besides providing several inside-out tracking methods (model-based KLT and marker-based) and using a specific exchange file format, the system supports the sequential mode of operation by text-based HUD information, showing the instruction and the location code, by 3D-registered information to highlight and label payload components, as well as to supplement them with 3D auxiliary objects to indicate, for example, the direction during rotating movements by 3D arrows. A strong tool is provided by the resource pad, a mobile tangible AR interface to display images or video streams related to the presented instruction. To be hands-free while operating the MARSOP system and using no mouse or keyboard, simple voice commanding was realized to sequentially navigate through a procedure.

To identify the users' needs and acceptance the prototype should be tested in a controlled field study. Therefore I conducted a proof-of-concept study at the European Astronaut Centre, using the Biolab training model and related ODF tasks to comply with the task-related requirements. To satisfy the user-related requirements of field testing, application domain experts were recruited, who were payload instructors, introducing and training astronauts in handling the ISS on-board systems. Using a within-subject designed, the traditional guiding method was compared with guidance provided by the MARSOP system. The performances were evaluated by analyzing the resulting workflow and the completion times, while the effort spent on task performance was assessed by the workload level obtained from subjective ratings using the NASA RTLX. To grade the experts' level of acceptance and preference, the ratings and comments of a self-metric usability questionnaire was analyzed. The results showed that the MARSOP system can reduce sequence errors and thus supporting the sequential operation mode better than the traditional method. No differences in completion times have proven that the MARSOP system fulfills payload tasks in the same way as usual. It was also shown that the application of AR reduces the self-reported workload level for those tasks that vary in their instructions and require constant learning. Besides not satisfied with the video see-through HMD, the feedback of the domain application experts showed a high level of satisfaction and acceptance, and revealed that most of them would prefer to use MARSOP instead of the conventional system. Hence, this study has shown that AR in the form of the MARSOP prototype has the potential to replace the common guidance payload interface and to support a better quality of task performance. But nevertheless, it is needed to be verified, optimally under long-term microgravity conditions onboard the ISS.

In the case of using the MASROP system to support ISS payload tasks, two fundamental issues, initial limiting the current prototype (see section 6.1.5), remain to be clarified. As has also been pointed out in section 5.2, these issues are related to task localization (see section 5.2.2) and the placement of devices needed for control and input tasks (see section 5.2.3), both are covered by the following research conducted in the scope of this thesis. While the former one has based its request on ambiguous guidelines related to egocentric and exocentric target cueing to localize the off-screen target area and support the in-view operation (see Chapter 7), the latter one is deduced from missing

microgravity research, studying the trade-off between world-, body- and head referenced interfaces under modified gravity conditions (see Chapter 8).

Chapter 7

Task Localization During Payload Operations: Egocentric and Exocentric Target Cueing for Visual Search and Task Operation *(ARGuide Study)*

The success of procedural payload guidance can be affected by the way how the instructions or target cues are spatially oriented. According to human information processing, spatial cues can be encoded in an exocentric or egocentric way. The traditional guidance method for space operations implies an exocentric display and thus also an exocentric visualization of the target cues. In contrast, using a head-mounted display provides an egocentric display, but can nevertheless differ in the spatial orientation of the visualized task cues. As a consequence and in the style of payload operations, this chapter reports a trade-off study, examining the effect of the spatial orientation of task cues on operator workload, assessed in its full dimension. With this experiment I aimed to inform the design of future AR applications in display settings for the instructional presentation during payload operations.

7.1 Research Objective

When astronauts carry out instructional payload tasks, a deciding factor for their performance is related to the process of the visual search that is needed to localize the target area and to subsequently detect the target object in questions to perform its related task. While in the previous chapter the feasibility and acceptance of AR guided payload operations in a domain expert field was studied, this research is more focused on the visual search to optimize the task localization and detection in consideration of human's ability to process spatial information, which can be influenced by the way the task cues are presented, either by egocentric or exocentric presentation schemes. Thereby a presentation scheme is defined by the spatial orientation of its display condition and task cues' visualization. Because of missing guidelines and ambiguous findings, as reviewed and discussed in section 5.2.2, the trade-off between egocentric and exocentric target cuing for task localization during payload tasks is needed to be studied by analyzing their effect on visuomotor coordination and workload. Resulting from this review, three different conditions for information presentation were identified (ED, AR and HUD), whereby each presentation scheme has its own characteristics caused by the reference frames used for the display screen on the one hand and on the other hand for the visualization of the task cues (see Table 7.1). Thereby the ED condition corresponds to an exocentric display, also implying an exocentric visualization of the task cues. From here this condition is also called external display. To present the target cues, the external display shows an off-screen orthogonal area map of the operation scene, providing the task instructions in a sequential order and is provided laterally at the operation area, and thus complied with the spatial orientation of the traditional SSC-based payload interface. Thereby the user is constantly forced to change the focus by head, and possibly by body, movements, as well as needs to memorize and cognitive transform the

task information to match its location. This implies, on the one hand, that the target localization using the external display calls for effortful top-down processing to serial combine the objects' features of the visual display, but, on the other hand, if spatial updating is requested, only the user's 3D pose to the underlying exocentric reference frame is needed to be updated. In contrast to the external display, presentations schemes that are provided by a head-mounted display (HMD) enable an egocentric view of the working area, but can differ in their reference frame used for the visualization of the target cues. Thereby an egocentric visualization is supported by the AR condition, considering the meaning of Azuma's definition of Augmented Reality (1997). Such AR display is characterized by two different modes for target cueing, which imply navigational support to initially localize the off-screen task area, while the target object is subsequently detected by, for example, highlighted geometry, which can be complemented by labeled annotations and 3D auxiliary objects (e.g., arrow) to support task performance. Thus, using the AR condition, the user is initially navigated to the next target location by showing the direction towards the off-screen target in an egocentric way. Once the target object appears in the HMD's field of view, the user subsequently gets 3D registered task information by superimposed hints (e.g., highlighting, arrows), as was already provided by the initial prototype used for proof-of-concept studying (see section 6.4). Consequently, AR enables to reduce the perceptual-cognitive effort during both phases. On the one hand, instead of localizing the target area by cognitive transformation and matching processes, the user only needs to follow the navigation hint. On the other hand, it maximizes bottom-up processing by the salience effect to match the target object in question, and thus reduces the demanding conjunction feature search. But following the approach of an egocentric model, the AR condition requested to update the spatial 3D pose of each object relative to the user if the viewpoint has changed, which can be challenging if the set size increases (Wang et al., 2006). Besides egocentric target cueing, using an exocentric HMD visualization will guide the user by an area map, as is shown by the ED display, but firmly presented to the display, like a static head-up display (HUD), whereby the user is also forced to change the focus, but only by eye movements, and also to use a memory task and cognitive transformation for matching the location. Because the HUD display is combining both egocentric and exocentric target cueing, it is not clear which kind of spatial updating will be triggered.

Display scheme	Reference frame of display	Reference frame of target cues	Target cue visualization	Task localization by	Spatial updating	Visual search processing	Focus change
ED	exocentric	exocentric	off-screen area map	memorizing, cognitive transformation and matching	relative 3D pose of the user	increased top-down by serial conjunction search	yes
HUD	**egocentric**	**exocentric**	in-view area map	memorizing, cognitive transformation and matching	**?**	increased top-down by serial conjunction search	yes
AR	egocentric	egocentric	in-situ conformal symbology	navigation, 3D registered target hint	relative 3D pose of each object	maximized bottom-up by salience effect	no

Table 7.1: The characteristics of the identified presentation schemes for task localization.

Hence, in this chapter I present an experiment that aimed to inform the design of future AR applications in spatial display setting for instructional presentation during payload operations. While the initial proof-of-concept study has verified the feasibility and acceptance of AR guided payload tasks, this research aimed to optimize their requested visual search, which can be influenced by the spatial orientation of target cues. Thus, the effectiveness and efficiency of exocentric and egocentric target cueing used to localize the off-screen task area and subsequent task operations was evaluated. More detailed and using the terms for location support from Robertson et al., (2014), this research is intended to investigate the trade-off between *in-situ* conformal target cueing (AR), unregistered *in-view* cuing (HUD) and *off-screen* target cueing (ED). While the effectiveness was assessed by usability measures obtained by visuomotor performance, the efficiency of the presentation schemes was evaluated by the operator workload assessed in its full dimension (see section 3.4.3), using secondary task and physiological measurements, as well as subjective ratings. This should enable to identify possible deficits in visuomotor performance and lacks of allocated attentional resources, as well as to verify that an egocentric representation of visual target cues using AR support is most suitable for payload guidance. By comparing the presentation schemes (ED, HUD, AR) related to the aspect of spatial processing during the localization of the off-screen target object and subsequent operation, led me to the following research questions:

Q1: How does the visual spatial orientation of a display influence the visuomotor performance?

- Do egocentric displays (HUD, AR) speed up the visuomotor performance?
- Do egocentric displays (HUD, AR) enhance the pointing accuracy?
- Is there a significant improvement in performance induced by egocentric AR target cueing?
- Does the egocentric AR presentation generally reduce head movements as concluded by Henderson and Feiner (2009, 2011a)?
- What kind of spatial reference frame will be triggered by egocentric displays with exocentric target visualization (HUD)?

Q2: How does visual spatial processing affect the operator workload?

- How is the spatial orientation of displays affected by dual-task costs?
- Are egocentric displays (HUD, AR) decreasing the workload?
- Is egocentric AR target cueing providing the lowest workload level?
- Does the workload caused by egocentric AR target cueing decrease over time

As already discussed in section 5.2.2, human spatial behavior is also affected by a natural predisposition. Thus, according to Goeke et al. (2013) I also considered and evaluated distinctions in individual navigation strategies of the participants to answer the following questions:

Q3: How does an individual predisposition for spatial navigation affect the performance and workload?

- Does the spatial orientation of the displays correspond with the individual navigation strategy?
- Does the performance with the egocentric AR display benefit from egocentric users?
- Does the performance with the exocentric ED display benefit from exocentric users?
- Does the individual predisposition for visual spatial processing need to be considered in display settings for payload guidance?

7.2 Methodology

To explore the effect of visual-spatial processing on performance and workload a within-subject experiment was conducted, whereby the presentation schemes (ED, HUD, AR) constituted the levels of the independent variable. Thereby the external display (ED) complied with an exocentric presentation scheme, while the AR and HUD schemes were provided by an egocentric display, using a HMD, but differed in their spatial orientation of the target cues. The participants needed to perform the experiment under all three display conditions and were requested to complete a visuomotor task, whereby visual presented instructions required motoric responses by aimed pointing movements (see section 7.2.1). To identify deficits in visual-spatial attention caused by the presentation schemes (ED, HUD, AR), I evaluated the performance and operator workload in its full dimensions. Therefore I analyzed primary task measures collected by the visuomotor task under a single-task and dual-task condition. Consequently, to quantify the effect of dual-task interferences, the participants were tested under the single-task condition and under the dual-task condition. During the dual-task condition, the primary visuomotor task was interfered with an auditory reaction time task that needed a response of a simple motoric output (see section 7.2.2). Thereby the participants were requested to allocate their attention in favor of the visuomotor task. The workload was also analyzed based on physiological and subjective indicators. The physiological workload was assessed by cardiac responses obtained from the heart rate variability (HRV) (see section 3.4.3, *Physiological Workload*). As it was done in the previous experiment, the subjective level of workload was assessed by the NASA RTLX rating scale (see section 3.4.3, *Subjective Workload*). Furthermore, to analyze differences between egocentric and exocentric participants, it was required to clarify which navigation strategy is given to the participants by nature. Therefore, each of them performed the Turning Study[75] related to Goeke et al. (2013).

7.2.1 Visuomotor Task

The primary visuomotor task was focused on localizing visual targets that needed to confirm by aimed pointing movements. Derived from ISPR operations for the Biolab payload (see section 2.4.2) I used a modified hardcopy of the upper left part of the Biolab (see Figure 7.1, left) with the dimension of 765 mm x 740 mm that was sufficient to mimic the reorientation process for spatial

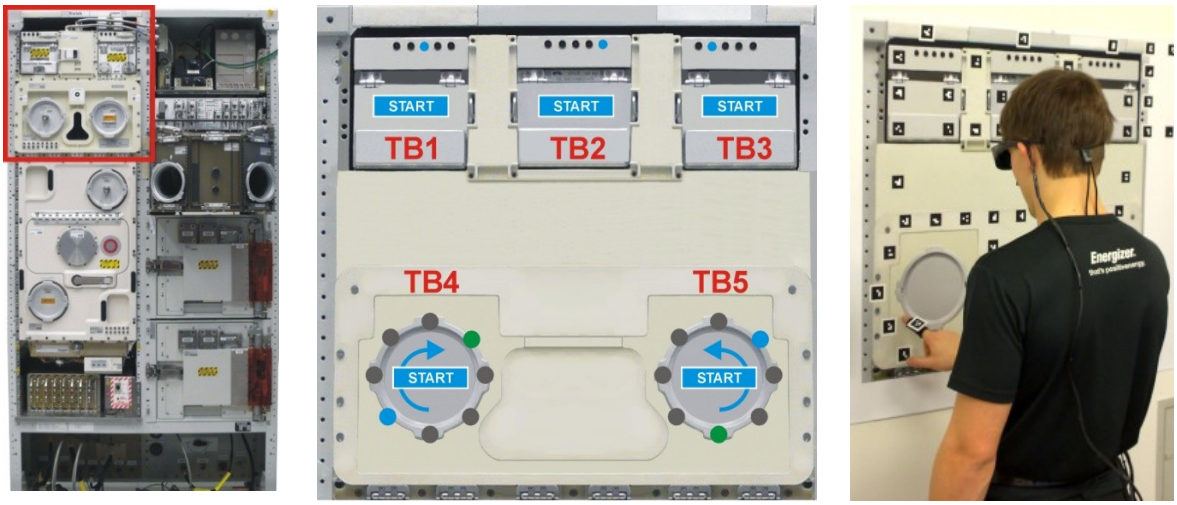


Figure 7.1: The visuomotor task using the Biolab. *(left)* The red frame indicates the used part. *(middle)* The modified task area used for the experiment with the five task batteries. *(right)* Participant performing a task.

[75] Turning Study. URL: http://www.navigationexperiments.com/TurningStudy.html, last visit: 28.09.2016

navigation. The Biolab copy was divided into five task areas for operations, each containing one task battery (see Figure 7.1, middle). Thereby two types of task batteries were defined differing in in their level of difficulty, which implied operation either by single-pointing (TB1, TB2, TB3) or by multiple directional-pointing (TB4, TB5). The participant was requested to perform twelve operation tasks (6 single- pointing, 6 multiple directional-pointing) that are displayed in a sequential order using the actual presentation scheme.

To extract the process of localizing the task area from the task completion, the participant had to confirm the actual target area of a task battery before the operation task could be executed. When the target area was localized, it had to be confirmed by pointing towards a virtual start button that was placed in the center of the task battery. Only if the start button was pressed the corresponding task battery was activated. A task battery for single-pointing consisted of five horizontal arranged buttons, whereby one button was visual highlighted and had to be confirmed by pointing towards the button. In contrast, a task battery for multiple directional-pointing consisted of eight circularly arranged buttons and the task was presented by the start and end button, as well as an arrow indicating the pointing direction. Thereby, each button between start and end must be confirmed in correct direction. While the single-pointing operation constituted an easy task with two task requirements, the operation by multiple directional-pointing implied an increased level of difficulty including four task requirements (see Table 7.2). Only if the participant had finished one task sequence in a correct way, the next task operation was initiated by the visual search process, which was intended to initial localize the next task area in question.

Task	Task Requirements	Stimulus (input)	Response (output)
Primary Visuomotor Task			
Search of task area	1. localizing task area 2. confirming task area	**ED:** visual (area map) **HUD:** visual (area map) **AR:** visual (navigation line)	motoric (single pointing)
Single-pointing	1. detecting task button 2. pointing towards task button	**ED:** visual (area map) **HUD:** visual (area map) **AR:** visual (3D registered superimposed button)	motoric (single pointing)
Multiple directional-pointing	1. detecting start button 2. detecting end button 3. identifying pointing direction 4. confirming all participated task buttons in correct order	**ED:** visual (area map) **HUD:** visual (area map) **AR:** visual (3D registered superimposed start and end button, inset centered arrow)	motoric (successive pointing)
Secondary Reaction-Time Task			
Reaction-Time	1. localizing single tone 2. confirming single tone	auditory (tone)	motoric (push button)

Table 7.2: The characteristics of the primary and secondary task related to their requirements, stimuli settings and response modalities.

All pointing tasks were performed by participants' index finger of the dominant hand (see Figure 7.1, right). While the exocentric presentation schemes (ED, HUD) provided the complete task information from the beginning, the egocentric scheme (AR) was represented by an initial navigation to the target

area before getting the task information. Using the exocentric presentation schemes (HUD, ED; see Figure 7.2, top-left and top-right), the participant localized the task area and the operation task using a drawing that showed a complete overview of the task environment. While the ED condition was provided by a virtual drawing (area map) that was firmly displayed directly to the right side of the task environment, the HUD condition showed this drawing statically fixed to the upper left part of subject's field of view. Both presentation schemes provided the complete task information from the beginning, which means that all information was already available during the localization of the target area. In contrast, the egocentric AR presentation scheme (see Figure 7.2, bottom) was represented by an initial navigation to the target area before getting the task information. Thereby the participant was navigated in an egocentric fashion by a line that was on-line rendered starting from the center of the field of view and directed toward the off-screen target area (see Figure 7.2, bottom-left). Once the target area was visible, the direction hint disappeared and the participant confirmed this area by pressing the start button. Then the operation task was visualized by 3D registered information by highlighting the single-pointing target or the start- and end button, as well as the direction hint that was centered placed in the button circle. To ensure a uniform data acquisition the ED condition was mapped in the virtual environment instead of using additional hardware, like a laptop. This implicated that the participants performed the experiment with the HMD under all display conditions. Figure 7.3 shows the studied display settings by a user's view through the HMD.

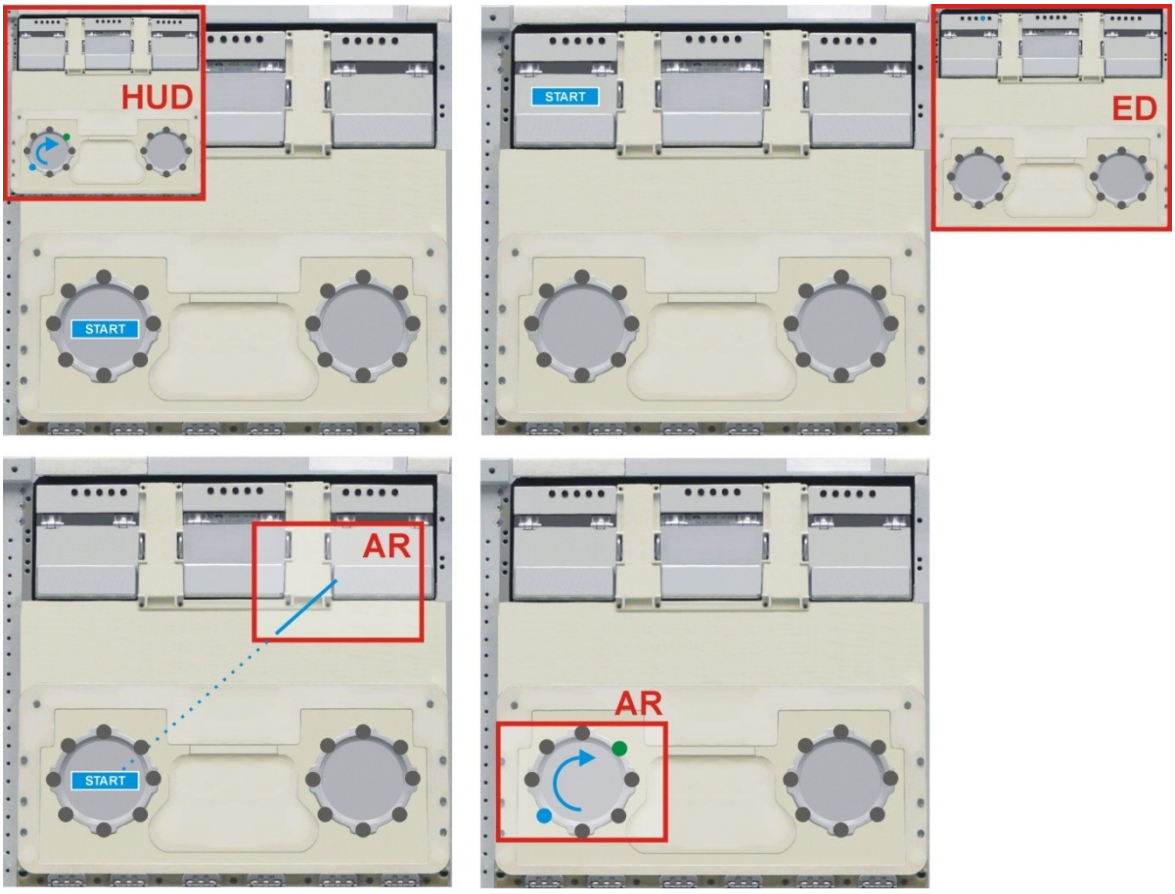


Figure 7.2: The studied display conditions with its spatial reference frames for the display and visualization*: (top-left)* HUD: exocentric visualization within and egocentric display, *(top-right)* ED: the exocentric external display, and *(bottom)* AR: egocentric visualization within egocentric display showing the localization of the off-screen target area by egocentric navigation and the superimposed 3D registered visual hints for operation.

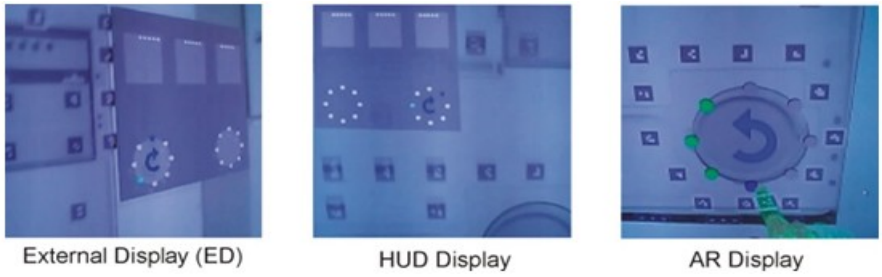


Figure 7.3: The studied display techniques showed by user's view through the HMD.

7.2.2 Secondary Task

To identify the amount of attentional resources demanded by the visuomotor task, its performance was extended by a concurrent task as index for the cognitive capacity. Therefore the primary visuomotor task was interfered by an auditory reaction-time task that required a simple motoric response by pressing a button. Thereby the participant was continuously exposed to an auditory stimulus and was requested to generate one response to one discrete signal. To avoid adaptation, the signals were randomly triggered every two to four seconds. The confirmations of the signals were performed by a single-handed controller, which the participants had to operate with the non-dominant hand. Table 7.2 summarizes the characteristics of the dual-task model. While the primary task required a motoric output by aimed pointing movements in response to a visual stimulus (visual attention), the secondary task required a simple motoric output in response to an auditory stimulus (auditory attention).

7.2.3 Experiment Procedure

Figure 7.4 shows the experiment procedure that each participant had to conduct and consisted of two HRV tests and six experiment sessions. The HRV tests served as baseline measures, whereby each participant was requested to perform a pre- and post-test, before starting and after completion of the experiment. During the tests the participant had to keep calm in standing posture for five minutes. To eliminate the assimilation of the cardiovascular system during the dislocation I only used the last 100 seconds for baseline purposes. For the experiment phase the participant was requested to try to stand on a marked point that was centered placed with 30 cm distance in front of the Biolab drawing (see Figure 7.1, right) that was fixed to the wall and individually adjusted in height for each participant.

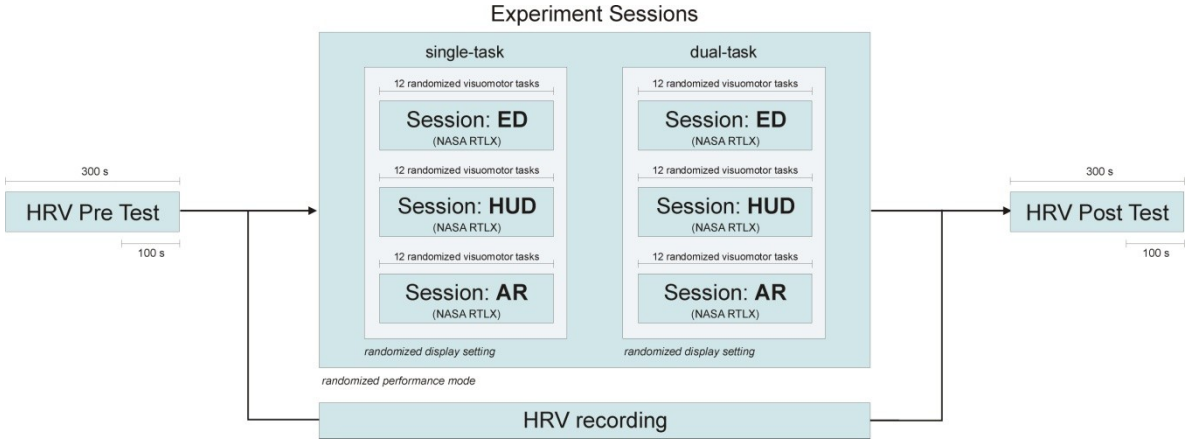


Figure 7.4: The experiment procedure performed by each participant showing the HRV tests and the experiment sessions randomized for the trial types, for the display settings and the visuomotor tasks.

In a randomized order the participant started with the single- or the dual-task condition and performed three randomized experiment sessions, each for one display condition. Within one display condition twelve randomized tasks (six single-pointing tasks, six multiple directional-pointing tasks) were presented that the participant had to conduct successively. After each session the participant was requested to fill in the NASA RTLX. During the experiment session the HRV was continuously recorded, whereby the changes of the sessions were marked by a short bow causing a shift in direction.

7.2.4 Apparatus

The experiment setup consists of a data processing unit (Lenovo ThinkPad T420s, 2.8 GHz CPU, NVIDIA Quadro NVS 4200M) with a connected binocular video-see-through HMD. The HMD was the video-based Vuzix WRAP 920AR (see Appendix B), whereby the integrated camera setup was also used for optical inside-out marker tracking to obtain the 3D pose of participant's head and the pose of the index finger for pointing purpose. As shown in Figure 7.1 (right), the working area was equipped with a multi-marker pattern and a single marker was attached to the index finger of participant's dominant hand. A single-handed controller (Nintendo Wii) was used to confirm the auditory signal. To assess the physiological workload by cardiac responses the participants were equipped with a wireless eMotion HRV sensor from Mega Electronics that recorded the heart rate variability at a sampling frequency of 1000 Hz and an accuracy of one ms. I implemented the experiment software in C++ using Qt-4.8.5 for the 2D operator interface, osg-3.0.1 for rendering the 3D content and osgARTProfessional-1.1.3 for marker tracking.

7.2.5 Participants

The study was conducted with 13 participants (3 women, 10 men) aged from 20 to 46 years (20-29 years: 10 participants, 30-39 years: 1 participants 40-49 years: 2 participants; $M = 27.23$, $SD = 8.66$). Five participants had experiences with AR interfaces in terms of participation in previous studies, seven participants were novices and one was an expert. They came from backgrounds in sports science, medical science, biology and computer science. All participants had a right-dominant arm that was used for the pointing task. Related to individual predispositions for spatial navigation, the results of the Turning Study (Goeke et al., 2013) showed that seven participants used an egocentric strategy and six participants an exocentric strategy. Each participant was previously introduced and completed a short training session to get familiar with the system and the display techniques.

7.2.6 Measurements and Data Analysis

To quantify human performance and workload, several features were identified that represented the dependent variables of the experiment (see Table 7.3). The performance of the primary visuomotor task was evaluated by time effects, whereby the session time and task completion time were considered separately. To execute a payload instruction the operator must first localize the target area before performing the task. Therefore the task completion time was additionally split into search and operation time, both separately measured and evaluated. The performance was also evaluated by the accuracy of pointing towards the task buttons. The workload was assessed multi-dimensionally (see section 3.4.3), using three indicators: secondary task performance, subjective rating and physiological measures. Therefore the secondary reaction time task was used to evaluate the attentional demand by, on the one hand, analyzing the response time and the response rate of the

auditory stimuli, and on the other hand to assess the dual-costs as the percentage shift between the single-task and dual-task performance, which was analyzed for each dependent variable of the primary tasks. The response time, measured by the reaction time task, was analyzed for the task sequence and differentiated by the localization of the target area (search) and the subsequent operation. The response time of the secondary task was also used for a resource trend analysis (time line analysis) to predict workload levels. For the subjective perceived workload, the loading factors of the NASA RTLX were analyzed. The HRV was analyzed by the frequency domain using the ratio of low-frequency power (LF, .04 to .15 Hz) to high-frequency power (HF, .15 to .40 Hz) within the respective spectrum. This LF/HF ratio reflects the interaction between the parasympathetic nervous system (relaxation, HF) und the sympathetic nervous system (stress, LF), also known as the cardiac sympathovagal balance. An increase in the LF/HF ratio correspond to an increase of the sympathetic nervous system and thus to a higher stress level indicating a higher mental workload. However, a decrease in the LF/HF state is related to an increase of the parasympathetic nervous system (LF) causing a more relaxed state, indicating a decreased level of the mental workload.

Dependent Variable	Description
Primary Visuomotor Task	
Session completion time [s]	Time to complete a session (twelve tasks), from first input to last response.
Task completion time [s]	Time to complete a task sequence covering the search and operation
Search time [s]	Time between visual stimulus onset and motoric response onset. Time that is needed to visually localize a target area within one task sequence, i.e. the time between finishing the last operation task until confirming the next start button.
Operation time [s]	Time between visual stimulus onset and motoric response onset. Time that is needed for the operation task (single-, multiple directional-pointing) and measured from pressing the start button until the task is correctly completed.
Pointing Accuracy [mm]	Euclidean distance between the center of the task button and the resulted pointing intersection within this buttons.
Secondary Reaction-Time Task	
Response rate [%]	Percentage ratio of responded to triggered auditory stimuli.
Response time [ms]	Time elapsed between auditory stimulus onset and motoric response onset.
Dual-Task cost [%]	Percentage shift between single-task and dual-task scores calculated by: $100 * (singleTask\ score - dualTask\ score) / singleTask\ score$
Subjective Workload	
Loading Factors	NASA RTLX rating scale (see Appendix A).
Physiological Workload by HRV	
LF/HF ratio [ms^2]	Frequency domain parameter representing the cardiac sympathovagal balance.

Table 7.3: Performance and workload measures as dependent variables of the ARGuide study.

The outcomes are presented by mean value plots that are marked with the confidence interval (CI) of 95% and by distribution plots. Statistical analyzing of the variance for repeated measures revealed from primary and secondary task performance was done with a linear mixed model using SAS® PROC MIXED with simulated adjustment to keep the experimentwise error rate $\alpha = 5\ \%$ (see section 3.4.2, *Data Analyzing*). Thereby the display conditions (ED, HUD, AR) were multiple

pairwise compared at the same stage of trial type (single-task, dual-task) and between the trial types. Differences between participants' navigation strategies were analyzed by multiple pairwise comparisons at the same stage of display condition. For analyzing main effects on the NASA RTLX rating scores I conducted a non-parametric test using the SAS® PROC NPAR1WAY with the WILCOXON option that for three sample populations induced a Kruskal-Wallis rank test, only reporting a main effect. For post-hoc testing the SAS® procedure PROC GENMOD was used for multiple pairwise comparisons between the display techniques across all loading factors and between the factors at the same stage of display technique. Also hereby, the experimentwise error rate was kept by the Simulate method. Furthermore, all outcomes were analyzed in the same way for differentiation of participants' strategy in spatial navigation (egocentric, exocentric). Overall, 78 data set for the test sessions, 936 data sets for the task sequences and 1082 for analyzing the response time to the auditory stimuli were analyzed. To evaluate the physiological workload by the LF/HF ratio, I could only analyze the data from twelve participants because the wireless HRV sensor did not record the data for one subject.

7.3 Results

The results are divided into six subsections, considering the analysis of the visuomotor task performance, the secondary task performance, the physiological workload, the subjective level of workload, an analysis of the resulted head movements, as well as the analysis of the visuomotor and secondary performance with respect to participants' individual spatial navigation strategy. Thereby I analyze the impact of the spatial orientation of the studied presentation schemes within each of these sections. Analyzing the secondary performance is complemented by a resource trend analysis to find out which display condition is most suitable to support a low workload level for longer tasks. Findings will be reviewed within the discussion section (see section 7.4.7). Appendix D provides the descriptive statistic for the outcomes showing measures of the central tendency and variability. For the physiological workload, it additionally presents the measures and results of the comparison of the LF/HF ratio between the display conditions.

7.3.1 Visuomotor Task Performance and Dual-Task Costs

In this section, the differences between the display conditions (ED, HUD, AR) related to the performance metrics of the visuomotor task will be evaluated and discussed. This happens by analyzing the session completion time, the task completion time, the search and operation time, as well as the pointing accuracy. The outcomes are analyzed and presented for each condition of trial type (single-task, dual-task), whereby the time effects on dual-task interference will also be considered to reflect the costs of dual-task processing.

Session and Task Completion Time

While the session completion time is the time that the participant needed for a successive series of six task sequences under one display condition, the task completion time reflects the time over one task sequence. The variations of the mean session and task completion time per display condition and trial types are presented in Figure 7.5, while Table 7.4 shows the results of the comparison of the display conditions at the same stage of trial type. Whether the session or task completion time, for both and independent from the trial type, the data revealed that the egocentric displays (AR, HUD) promote significant faster performance than the exocentric display (ED). But contrary to the tendency shown

in Figure 7.5, there is no significant effect ($p > .05$) between the egocentric visualizations (HUD, AR).

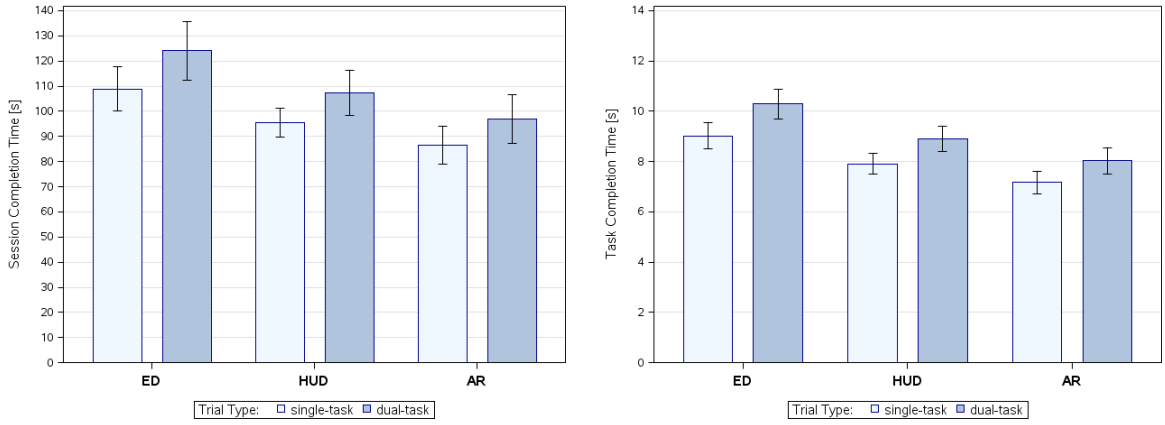


Figure 7.5: Mean completion time (with CI=95%) of the visuomotor task per display and trial type: *(left)* for session sequence covering twelve tasks, and *(right)* for task sequence.

Variable	Trial Type	Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Session Completion Time [s]	single-task	AR	ED	36	-4.59	<.0001	0.0002	*
	single-task	AR	HUD	36	-1.83	0.0755	0.1743	
	single-task	ED	HUD	36	2.76	0.0091	0.0240	*
	dual-task	AR	ED	36	-4.10	0.0002	0.0006	*
	dual-task	AR	HUD	36	-1.59	0.1205	0.2628	
	dual-task	ED	HUD	36	2.51	0.0167	0.0432	*
Task Completion Time [s]	single-task	AR	ED	465	-5.49	<.0001	<.0001	*
	single-task	AR	HUD	465	-2.20	0.0286	0.0730	
	single-task	ED	HUD	465	3.29	0.0011	0.0030	*
	dual-task	AR	ED	465	-5.83	<.0001	<.0001	*
	dual-task	AR	HUD	465	-2.25	0.0250	0.0643	
	dual-task	ED	HUD	465	3.58	0.0004	0.0011	*

Table 7.4: Comparison of the session and task completion time between the display methods at the same stage of trial type using SAS® PROC MIXED (adjust=simulate).

The comparison of the trial types for same display condition enables the assessment of performance costs related to dual-task interference processing (dual-task costs). As also shown in Figure 7.5, the dual-task processing deteriorates the performance by increased session and task completion times. For both completion times, the dual-task processing caused a percentage decrease of 14 % for ED, of 13% for HUD and of 12 % for AR. The comparison of the trial types at the same stage of display condition (see Table 7.5) shows that the dual-task interference yielded significant effects on the session completion time for the ED and HUD condition, but not significantly interfered the performance under the AR condition. With respect to the task completion time, the differences were significant for all display conditions. While the visuomotor performance under the AR condition was not significantly affected by dual-task processing over the complete session, it might be reasonably assumed that in-situ overlays contribute with decreased dual-task costs to a more efficient performance during longer tasks. But this is needed to be verified by analyzing the workload, especially by time line analysis related to the secondary task performance (see section 7.3.3.). Also

the physiological workload (see section 7.3.4) might provide guidance on whether the AR display has the potential to support longer task better.

Variable	Trial Type		Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Session	dual	single	AR	-	24	1.83	0.0801	0.0801	
Completion	dual	single	ED	-	24	2.26	0.0334	0.0334	*
Time [s]	dual	single	HUD	-	24	2.43	0.0229	0.0229	*
Task	dual	single	AR	-	310	2.45	0.0150	0.0150	*
Completion	dual	single	ED	-	310	3.17	0.0017	0.0017	*
Time [s]	dual	single	HUD	-	310	2.93	0.0037	0.0037	*

Table 7.5: Comparison of the session completion and search time between the trial types showing the effect of dual-task interference at the same stage of display method using SAS® PROC MIXED (adjust=simulate).

Search and Operation Time

A task sequence consisted of a search process to find the off-screen target area and its subsequent operation process. For further evaluation I separately considered their different times. Thereby, the search time reflects the time elapsed while localizing and confirming the presented task area. The variations of the mean search time per display condition and trial type are presented in Figure 7.6 (left). Mean values were most prominent for the external display (ED) under both task types. Compared to ED, the data showed that with HUD the target area was localized 24 % faster and with AR 33 % faster under the single-task condition, while with both HUD and AR the target area was 31 % faster localized under the dual-task condition. This may be a lot, but it is not surprising, since the search process includes the initial learning phase while using the exocentric display. The comparison of the display conditions at the same stage of trial type (see Table 7.6, top) confirmed this visual inspection and revealed significant differences in search response times between the exocentric display (ED) and the egocentric displays (HUD, AR) regardless of the trial type, but also showed that with HUD the target area was localized significantly slower than with AR, though only under the single-task condition, which may be an indicator that HUD was less interfered under the dual-task condition.

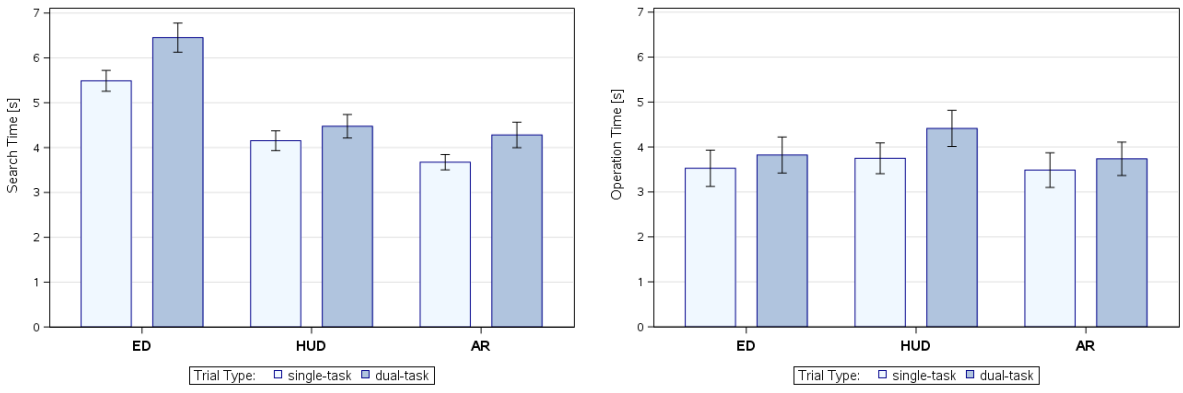


Figure 7.6: Mean search and operation time (with CI=95%) of the visuomotor task per display and trial type.

Contrary to the search time, the data for the subsequent operation process (see Figure 7.6, right) showed highest mean values for the HUD display under the dual-task condition, but revealed only significant time effects (see Table 7.6, bottom) compared to AR, while under the single-task condition the presentation schemes did not vary in their mean operation times. Otherwise, it is

noticeable that the operation performance using the external display (ED) did not differ from the egocentric display conditions (AR, HUD). This may be the case because the participants memorized the majority of the task information during the search process and if visual verifications via the off-screen drawing were required, they only needed to spatially update their position and orientation to the exocentric reference frame, while the egocentric nature of the AR condition demanded updating for each object in the visual display. Although the HUD display provided target cueing by an exocentric presentation, the fast search and slow operation do not suggest that learning of the task information was induced and required permanent visual verifications, and thus it can be assumed that an exocentric behavior was not triggered.

Variable	Trial Type	Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Search Time [s]	single-task	AR	ED	465	-12.07	<.0001	<.0001	*
	single-task	AR	HUD	465	-3.19	0.0015	0.0043	*
	single-task	ED	HUD	465	8.88	<.0001	<.0001	*
	dual-task	AR	ED	465	-10.39	<.0001	<.0001	*
	dual-task	AR	HUD	465	-0.93	0.3541	0.6231	
	dual-task	ED	HUD	465	9.47	<.0001	<.0001	*
Operation Time [s]	single-task	AR	ED	465	-0.15	0.8816	0.9878	
	single-task	AR	HUD	465	-0.97	0.3326	0.5963	
	single-task	ED	HUD	465	-0.82	0.4122	0.6903	
	dual-task	AR	ED	465	-0.30	0.7624	0.9508	
	dual-task	AR	HUD	465	-2.41	0.0163	0.0430	*
	dual-task	ED	HUD	465	-2.11	0.0355	0.0892	

Table 7.6: Comparison of the search and operation time between the display methods at the same stage of trial type using SAS® PROC MIXED (adjust=simulate).

By comparing the trial types at the same stage of display condition (see Table 7.7, top), the data showed significant time effects in dual-task costs during the initial search process under the ED and AR conditions, but not under the HUD condition, as was already assumed above. To search the task area, the dual-task processing caused a significant decrease of 17 % for ED and AR, but only 8 % for HUD. Contrary to the search time, where the ED and AR conditions were interfered but not the HUD, the dual-task processing of the operation (see Table 7.7, bottom) caused a significantly slower performance with the HUD (18 %), but did not interfere the performance under the ED (8 %) and AR (7 %) condition. Consequently, it is proven that using the in-view area map allocated least processing resources during the localization of the off-screen target area, while the same presentation scheme allocated the most for the subsequent operational process.

Variable	Trial Type		Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Search Time [s]	dual	single	AR	-	310	3.62	0.0003	0.0003	*
	dual	single	ED	-	310	4.76	<.0001	<.0001	*
	dual	single	HUD	-	310	1.86	0.0640	0.0640	
Operation [s]	dual	single	AR	-	310	0.93	0.3553	0.3553	
	dual	single	ED	-	310	1.03	0.3058	0.3058	
	dual	single	HUD	-	310	2.48	0.0135	0.0135	*

Table 7.7: Comparison of the search and operation time between the trial types showing the effect of dual-task interference at the same stage of display method using SAS® PROC MIXED (adjust=simulate).

Operation Time by Level of Task Difficulty

An operational sequence differed in its level of difficulty and task requirements, requesting for either single-pointing or multiple directional-pointing (see Table 7.2). Besides the fact that a higher level of task difficulty caused longer operation times, visual inspection of them (see Figure 7.7) shows the same mean variations of the display conditions, relative to one another for both levels of task difficulty, with slowest performance under the HUD condition as already shown for the overall operational performance (see Figure 7.6, right). But comparing the mean operation times between the display conditions revealed partially different findings (see Table 7.8). Considering only the performance under the single-task condition, this time the analysis confirmed the visual observation, showing a significantly slower operation with HUD compared to AR, but this only for the single-pointing task, while the multiple directional-pointing task was not affected by the display condition. However, the performance under the dual-task condition additionally showed significant differences between HUD and ED, while the overall operation (see Table 7.6, bottom) only differed between AR and HUD.

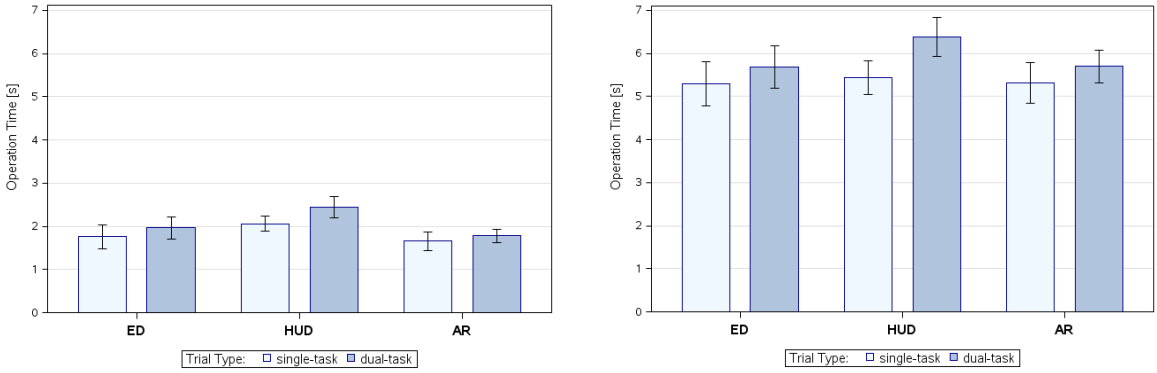


Figure 7.7: Mean operation time (with CI=95%) of the visuomotor task per display and trial type: *(left)* for the single-pointing task, and *(right)* for the multiple directional-pointing task.

Variable	Trial Type	Display	*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Operation of Single-Pointing [s]	single-task	AR ED	231	-0.60	0.5510	0.8218	
	single-task	AR HUD	231	-2.46	0.0146	0.0387	*
	single-task	ED HUD	231	-1.86	0.0638	0.1520	
	dual-task	AR ED	231	-1.16	0.2466	0.4774	
	dual-task	AR HUD	231	-4.25	<.0001	<.0001	*
	dual-task	ED HUD	231	-3.09	0.0023	0.0064	*
Operation of Multiple Directional-pointing [s]	single-task	AR ED	231	0.26	0.7942	0.9631	
	single-task	AR HUD	231	-0.52	0.6053	0.8629	
	single-task	ED HUD	231	-0.78	0.4370	0.7165	
	dual-task	AR ED	231	0.02	0.9809	0.9997	
	dual-task	AR HUD	231	-2.63	0.0092	0.0248	*
	dual-task	ED HUD	231	-2.65	0.0086	0.0232	*

Table 7.8: Comparison of the operation time between the display methods at the same stage of trial type using SAS® PROC MIXED (adjust=simulate).

The same ratios can be applied to compare the types of dual-task processing (see Table 7.9). As for the overall operation (see Table 7.7, bottom), the HUD was the one that was interfered by dual-task processing, regardless of the level of task difficulty. These results suggest that the level of task

difficulty is not really an indicator of varying time differences between the display conditions, but have verified that such operations are disadvantaged by using an in-view area map (HUD).

Variable	Trial Type		Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Operation of Single-Pointing [s]	dual	single	AR	-	154	0.89	0.3745	0.3745	
	dual	single	ED	-	154	1.10	0.2748	0.2748	
	dual	single	HUD	-	154	2.53	0.0124	0.0124	*
Operation of Directional-Poin ting [s]	dual	single	AR	-	154	1.26	0.2087	0.2087	
	dual	single	ED	-	154	1.08	0.2817	0.2817	
	dual	single	HUD	-	154	3.14	0.0020	0.0020	*

Table 7.9: Comparison of the operation time between the trial types showing the effect of dual-task interference at the same stage of display method using SAS® PROC MIXED (adjust=simulate).

Pointing Accuracy

Figure 7.8 shows the variations of the mean pointing distances with respect to the display techniques and trial types. Thereby the distance constitutes the Euclidean distance between the target's center and the resulted intersection point within this target that should reflect the pointing accuracy. As can be seen, the pointing accuracy was not affected by the display conditions, neither under single-task nor dual-task performance, and did not reveal significant differences. The costs of dual-task processing also did not reveal noticeable effects of the display techniques on pointing accuracy. Thus, it is proven that the pointing accuracy is not influenced by the spatial orientation of the target cues.

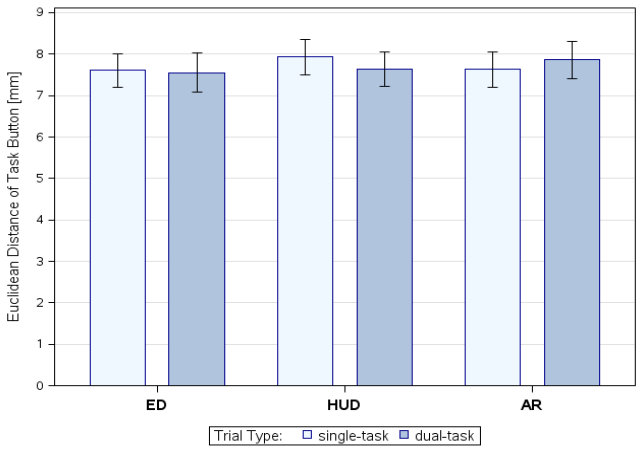


Figure 7.8: Mean Euclidean distance to the center of the target of the visuomotor task per display and trial type (with CI=95%).

7.3.2 Head Rotation Analysis

With analyzing the head rotations related to the rotation axes presented in Figure 7.9 (left), I explored how head movement was controlled by the target presentation schemes during the visuomotor task and to find out which scheme mostly supported head stabilization that could counteract fatigue, affecting the physical effort that can be linked to performance deficits. Therefore the absolute mean head rotations per rotation axis and display condition using 72248 valid tracking observations were evaluated. For calculation the absolute head rotation I took the difference between the 95th and 5th percentile of the rotation values per axis for each subject, task sequence and display condition. Figure 7.9 (right) presents the resulted mean rotations across all participants showing that the amplitude of the head rotation around the yaw axis was the largest using the ED technique. The comparison of the

mean yaw rotations (see Table 7.10) revealed significant effects between all display conditions, whereby the head rotation was minimized most using the HUD scheme. With respect to the roll axis the data also revealed the largest amplitude for the ED scheme and the comparison showed significant effects between the exocentric display (ED) and the egocentric displays (HUD, AR). The rotation around the pitch axis was not significantly affected by the display conditions. Figure 7.10 shows the changes of the head movements for each rotation axis illustrating that the largest rotation amplitudes around the Yaw axis were produced steadily over the session time sequence by the ED condition, while the HUD caused lowest amplitudes constantly over time.

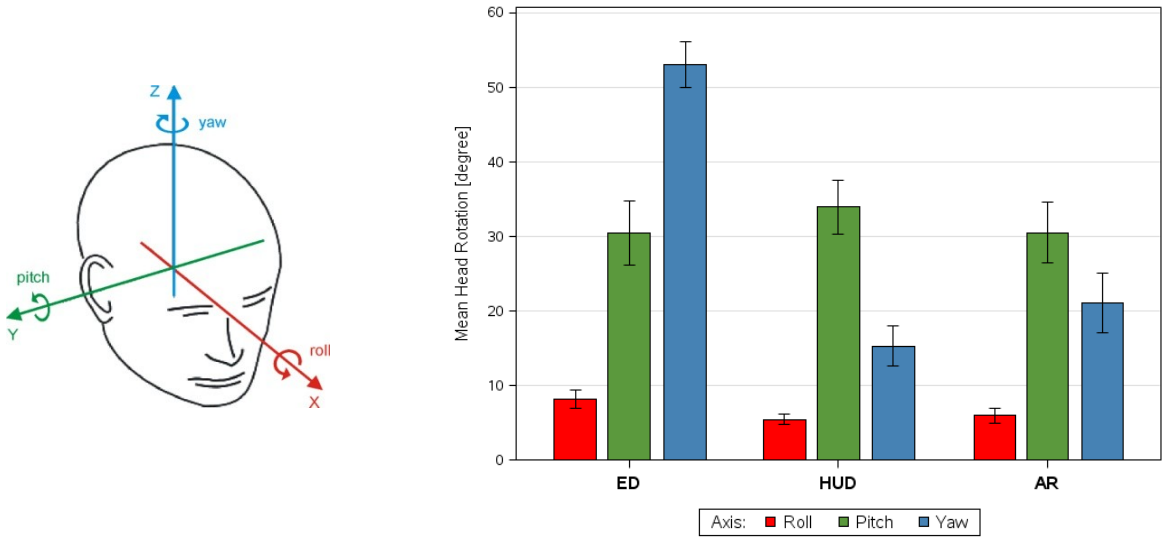


Figure 7.9: Head rotation analysis. *(left)* Head coordinate system with the Z-axis as up-vector, the Y-axis as right vector and the X-axis as forward vector. The head rotations correspond to the Euler angles with yaw as rotation around the vertical axis (Z), pitch as rotation around the side-to-side axis (Y) and roll as rotation around the front-to-back axis (X). *(right)* Absolute mean head rotations around the roll, pitch, yaw angles (with CI=95%) per display method and across all participants.

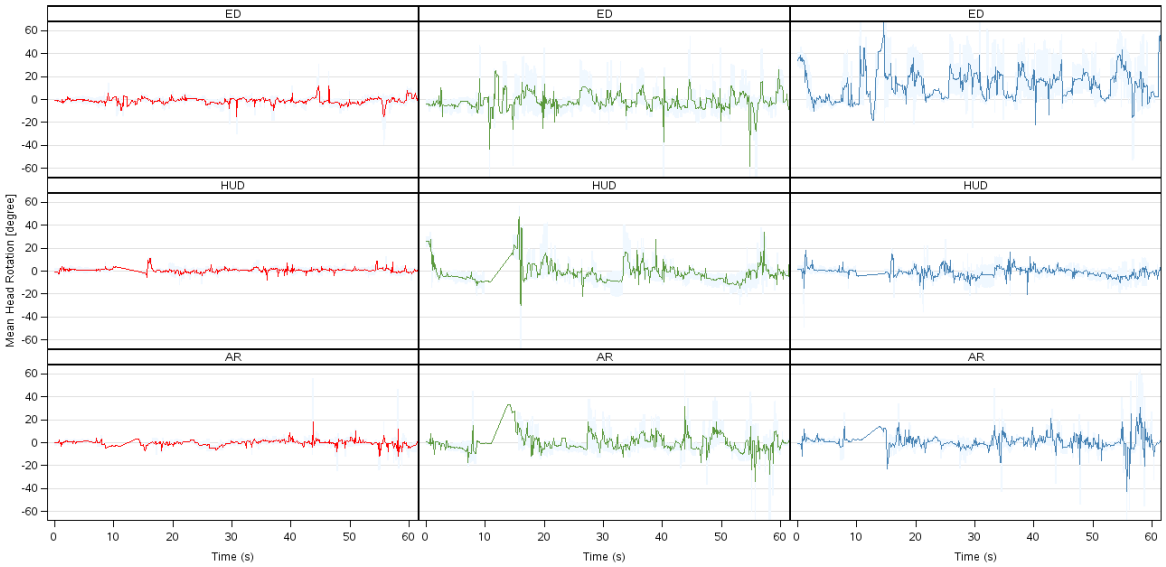


Figure 7.10: Changes in the head orientation by connected mean values with IQR error bars across all participants over the session time: *(left)* for the roll axis, *(middle)* for the pitch axis, and *(right)* for the yaw axis.

Variable	Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Roll (X-axis) [degree]	AR	ED	231	-2.95	0.0035	0.0097	*
	ED	HUD	231	3.72	0.0002	0.0007	*
Yaw (Z-axis) [degree]	AR	ED	231	-13.72	<.0001	<.0001	*
	AR	HUD	231	2.48	0.0140	0.0371	*
	ED	HUD	231	16.19	<.0001	<.0001	*

Table 7.10: Significant differences between the display methods related to comparing the head rotation by the Euler angles (roll, pitch, yaw) using SAS® PROC MIXED (adjust=simulate).

7.3.3 Secondary Task Performance

The assessment of workload by divided attention enables to identify limits on attentional resources during the performance of the primary visuomotor task. Therefore I analyzed the response rate and response time to the auditory stimuli with respect to the secondary reaction-time task over the complete task and separately for the search and the operation process. By means of a timeline analysis over the complete experiment session, I wanted to know which presentation scheme has the greatest potential to support a low workload level for prolonged tasks.

Response Rate and Response Time

The percentage response rates as the ratios between respond to triggered auditory stimuli did not really differ in their means. Thereby the external display ED revealed a mean response rate of 93 % (SD = 9 %), the HUD 94 % (SD = 9%), and the highest response rate, although only slightly, was achieved using the AR condition (M = 95 %, SD = 8 %). The response time is the time for responding to one auditory stimulus and the variations of the means are presented in Figure 7.11, while the results of the comparison are shown in Table 7.11. The evaluation of the overall response times (see Figure 7.11, left) revealed prominent increased values for ED, but only significantly differ from the mean of HUD that was only slightly faster than the AR mean. Just like for the visuomotor performance I considered the search and the operation time separately.

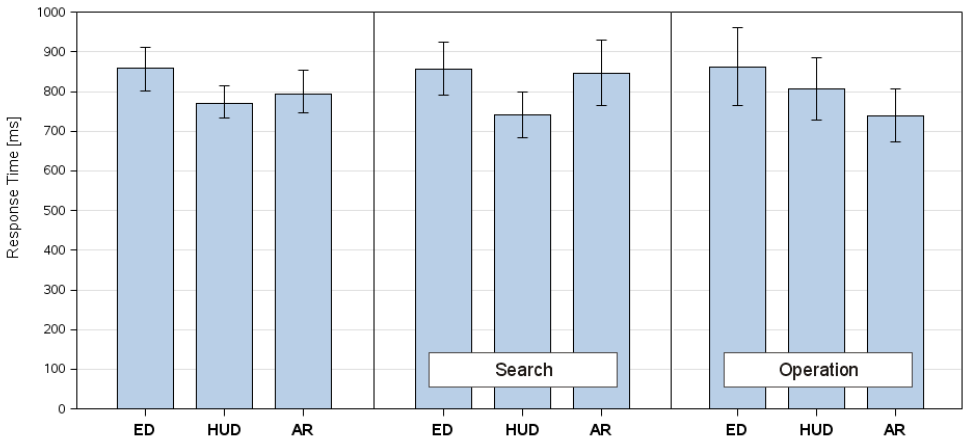


Figure 7.11: Mean response time (with CI=95%) of the reaction-time task per display method over: *(left)* the complete task, *(middle)* the search process, and *(right)* the operation.

For the search process (see Figure 7.11, middle) the data yielded fastest responses with HUD and slowest responses with ED and, contrary to my expectation, also with AR, but only revealed significant differences between the ED and HUD. In contrast, during the operation process (see Figure 7.11, right) the data revealed the shortest responses with AR and the longest with ED, but

multiple comparison did not revealed significant differences between the display conditions. Because the workload demanded by the primary-task is inversely reflected in the secondary-task performance, the outcomes show evidence that the mental workload on divided attention was significantly decreased with under the HUD condition during localization the target area, but showed a tendency for the most decreased workload under the AR condition during the operational process.

Variable	Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Response Time - Overall [ms]	AR	ED	1079	-1.66	0.0974	0.2212	
	AR	HUD	1079	0.61	0.5449	0.8167	
	ED	HUD	1079	2.35	0.0190	0.0491	*
Response Time - Search [ms]	AR	ED	613	-0.22	0.8229	0.9730	
	AR	HUD	613	1.95	0.0512	0.1237	
	ED	HUD	613	2.39	0.0173	0.0447	*
Response Time - Operation [ms]	AR	ED	463	-2.07	0.0392	0.0977	
	AR	HUD	463	-1.17	0.2423	0.4704	
	ED	HUD	463	0.96	0.3358	0.6001	

Table 7.11: Comparison of the response time between the display methods of the secondary task using SAS® PROC MIXED (adjust=simulate).

Resource Trend Analysis

The resource trend analysis complies with a time line analysis that is intended to predict the workload levels and should inform about the tendency of how attentional demand affects the workload during the visuomotor performance of the studied display conditions for prolonged task operations. Therefore I study how the response time to the auditory stimuli was affected over the session time and quantify the association between these two time variables. To conduct the trend analysis, the data needed are prepared in advance, whereby one data set mirrored one auditory signal that was responded containing, amongst other things, the date and start time for a session (participant per display condition), as well as the release and the response time to the auditory signal.

Following steps I performed for the data preparation using SAS® 9.4:

1. Sort by participant, display condition, date, and release time of the auditory stimuli.
2. Get start time of each session grouped by participant and method, and
3. Calculate local time difference by the release time for each associated auditory stimuli entry.
4. Generate uniform time classes and supplement each group (participant, method) with these classes.
5. Perform local regression (SAS® PROC LOESS) by participant and method to predict response time for each time class.
6. Merge individual results of each participant per method.
7. Sort by participant, method and time class.
8. Calculate means of predicted values (SAS® PROC MEANS) across all participant classified by method and time class.

To measure the strength of a linear relationship between the predicted means and the time classes I calculated the Pearson's correlation coefficient (SAS® PROC CORR) and performed linear regression (SAS® PROC REG) to identify a causal effect of time increases upon the response time to the

auditory stimuli. Following Cohen's classification of correlation (1988), the related analysis revealed that the response time and session time were moderately correlated under the HUD condition (r = -.49, p < .0001), strongly correlated under the AR condition (r = -.50, p < .0001) and very strongly under the ED condition (r = -.80, p < .0001). As was indicated by the correlation analysis and is shown by the trend lines in Figure 7.12, all resulted prediction models will decrease over the session time by a negative slope, also referred to as the regression coefficient. Thereby the response time will decrease slowest under the HUD condition, but starts with the lowest intercept. Conversely, the response time will decrease fastest under the ED condition, but also revealed the highest intercept. Thus, the HUD condition will provide slower response times than ED after 284 seconds (HUD: $RT_{[284]}$ = 567.19 ms, ED: $RT_{[284]}$ = 566.70 ms). However, the response time under the AR condition started with a higher intercept, only slightly lower than ED, but was also faster decreased than with the HUD, so that the AR condition would provide faster responses after 355 seconds than HUD (AR: $RT_{[355]}$ = 520.57 ms, HUD: $RT_{[109]}$ = 521.04 ms). Nonetheless, the response time under the ED condition decreased fastest and thus just after 109 seconds it will provide faster responses than AR (ED: $RT_{[109]}$ = 795.95 ms, AR: $RT_{[109]}$ = 796.09 ms). The results of the regression indicated that the session time explaines 24 % of the variance under HUD (F(1,97) = 4.65, p <.0001), 25 % under AR (F(1,97) = 6.24, p <.0001), and even 64 % of the variance under ED (F(1,97) = 3.20, p <.0001).

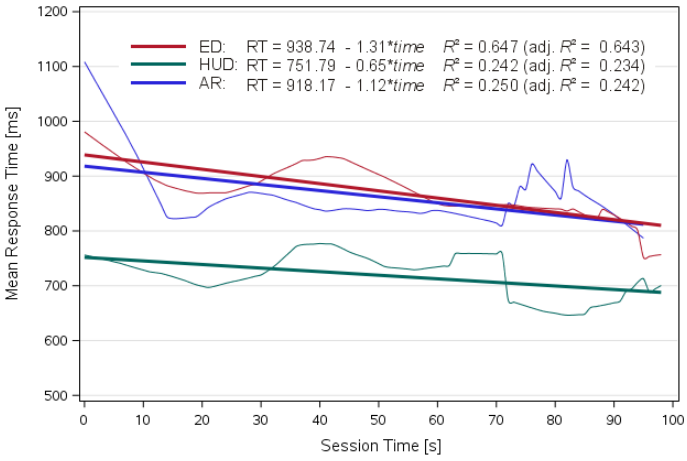


Figure 7.12: Mean response times of the reaction-time task over the session time per display method and the corresponding regression equations across all participants.

7.3.4 Physiological Workload

The physiological workload was assessed by the heart rate variability (HRV) analyzing the changes in the LF/HF ratio that reflects the sympathovagal balance and states the smaller the ratio, the greater the parasympathetic response (relaxed state), and vice versa, the higher the ratio, the greater the sympathetic response (excited state). To limit inter-individual variations, I first normalized the data for each participant using the difference between the parameter value (LF/HF ratio) and the mean of the baseline values that was obtained from the pre and post HRV test performed by each participant. Figure 7.13 shows the distributions of the LF/HF ratio across all participants per display setting and trial type. Under the single-task condition (see Figure 7.13, left) there were no remarkable variations of the mean values for the LF/HF ratio between the display conditions, while under the dual-task condition (see Figure 7.13, right) the LF/HF ratio was most decrease under the HUD condition and therewith caused the most prominent development of the parasympathetic response indicating the most relaxed state. Being aware that the outcomes did not revealed significant differences (see Appendix D), only tendencies, and based on the assumption that the dual-task performance compensate the field condition of normal payload operations, I can assume that the physiological

workload and thus the level of arousal will be decreased most under the HUD condition. But it remains to be clarified whether this arousal level is optimally according to the *Yerkes-Dodson law* (see section 3.3.2), because the findings, prior presented, indicated that an in-view area map claimed least processing resources only during the initial localization of the target area, but also generally provided least head movements around the yaw and roll axes, implying the fewest physical effort spent on receiving the task information.

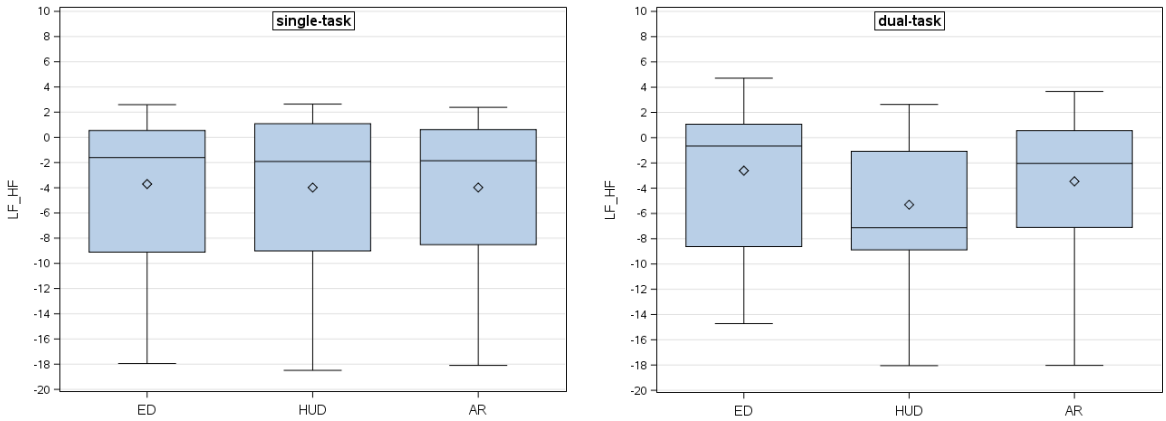


Figure 7.13: Distribution of the HRV parameter LF/HF Ratio categorized by the display methods showing decreased activity of the parasympathetic nervous system for the HUD condition under dual-task condition.

7.3.5 Subjective Workload

For the subjective level of workload I analyzed participants' ratings of the NASA RTLX scale that each participant had to fill in after completing the session covering the performance of twelve task sequences under one display technique. Thereby the rating scores were separately considered for each trial type. Figure 7.14 reflects the mean ratings scores across all participants per display condition and grouped by the trial type. Whether under the single-task condition or under the dual-task condition, the participants rated their experienced workload for nearly all items in favor of the egocentric AR display, while the ED display was predominantly higher rated.

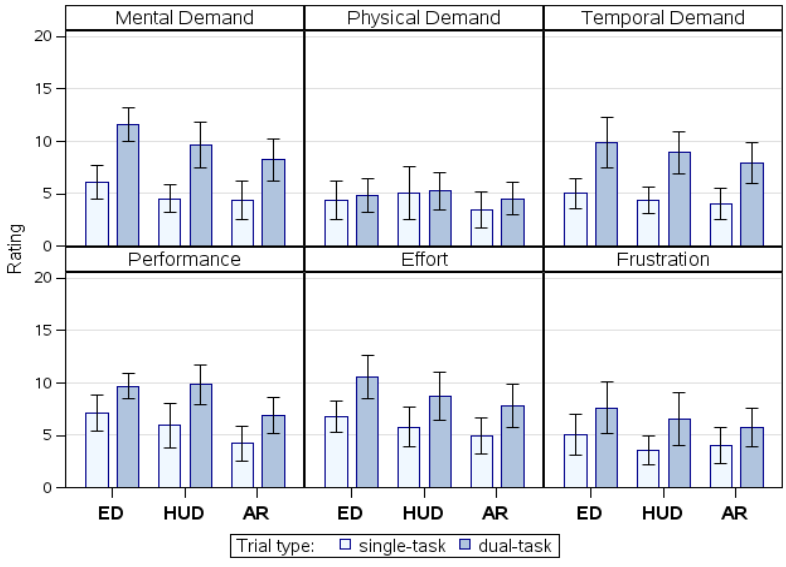


Figure 7.14: Mean ratings (with CI=95%) of the NASA RTLX scale categorized by the display methods and paneled by ratings items.

Analyzing the Wilcoxon scores under the singe-task condition (see Figure 7.15, left) revealed a significant difference, $H(DF=2, N=78) = 14.32$, $p = .0008$, between the rating scores of the display conditions with a mean rank of 138.01 for ED, 115.40 for HUD and 98.28 for AR. Also, the scores under the dual condition (see Figure 7.15, right) yielded a significant difference, $H(DF=2, N=78) = 13.10$, $p = .0014$, with a mean rank of 135.02 for ED, 121.08 for HUD and 96.40 for AR. But the post-hoc analysis per trial type (see Table 7.12) only revealed significant differences between AR and ED under the single- and dual task condition. Overall, the results of analyzing the subjective experienced level of workload suggest that the participants perceived the most decreased workload by the AR condition and significantly differed from the workload perceived by the external display (ED), regardless of the trial type.

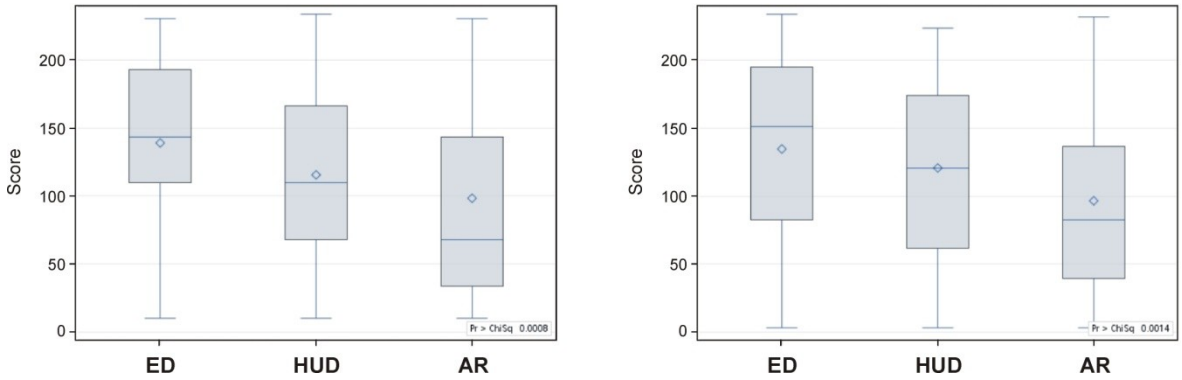


Figure 7.15: Distributions of the Wilcoxon scores across all factors of the NASA RTLX per display method for *(left)* the single-task condition and *(right)* the dual-task condition yielded from SAS® PROC NPAR1WAY.

Trial Type	Display		*z Value*	*Pr > |z|*	*Adj. p-Val*	
single-task	AR	ED	3.80	0.0001	0.0004	*
	AR	HUD	1.69	0.0920	0.2113	
	ED	HUD	-2.21	0.0271	0.0731	
dual-Task	AR	ED	3.55	0.0004	0.0011	*
	AR	HUD	2.23	0.0260	0.0703	
	ED	HUD	-1.33	0.1834	0.3796	

Table 7.12: Comparison of the NASA RTLX rating scores between the display methods per trial type using SAS® PROC GENMOD (dist=multinomial, link=cumlogit, adjust=simulate).

7.3.6 Differentiation by Individual Spatial Navigation Strategy

Finally I present and discuss time effects caused by the studied display conditions with respect to participants' spatial navigation predisposition. According to the spatial orientation of the display setting, I expected that participant using an egocentric navigation strategy would benefit more from the egocentric AR display and vice versa, participants using an exocentric strategy would benefit more from exocentric target cueing. Therefore the margin between the navigation strategies was compared at the same level of display condition related to the task completion time. This analysis was also conducted separately for the initial search and the operation, although the aspect of analyzing the navigation strategies is more connected to localizing the off-screen target area than with in-view target detection during the subsequent operation, because it is less linked to the nature of navigation.

7.3.6.1 Visuomotor Task Performance by Spatial Navigation Strategy

The variations of the mean task completion time per navigation strategy grouped by the display conditions and paneled by the trial type are shown in Figure 7.16. The data revealed that the initial search under ED and AR was affected by participants' navigation strategies (see Table 7.13). It shows that egocentric participants needed significantly longer to localize the target area with the external display (ED) than exocentric one, but this just under the single-task condition. However, under the dual-task condition, the results showed evidence that exocentric participants needed significantly longer with AR. Thus, it can be assumed that the spatial orientation of the ED and AR display seems to be linked to the corresponding individual navigation strategy during the initial search. But it is not clear, which strategy the HUD display supported. As was expected, the subsequent operational process (see Figure 7.16, right) did not show differences between the navigation strategies. Thus, it can be concluded that the primary task performance during the initial localization of off-screen targets was influenced by the participants' navigation strategy, while in-situ target detection was not affected by the participants' navigation strategy.

Variable	Trial	Display		Strategy		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Search Time [s]	single-task	ED	-	ego	exo	154	3.65	0.0004	0.0003	*
	dual-task	AR	-	ego	exo	154	-1.98	0.0496	0.0498	*

Table 7.13: Significant differences between the navigation strategies at the same stage of display method related to the initial search time of the visuomotor task using SAS® PROC MIXED (adjust=simulate).

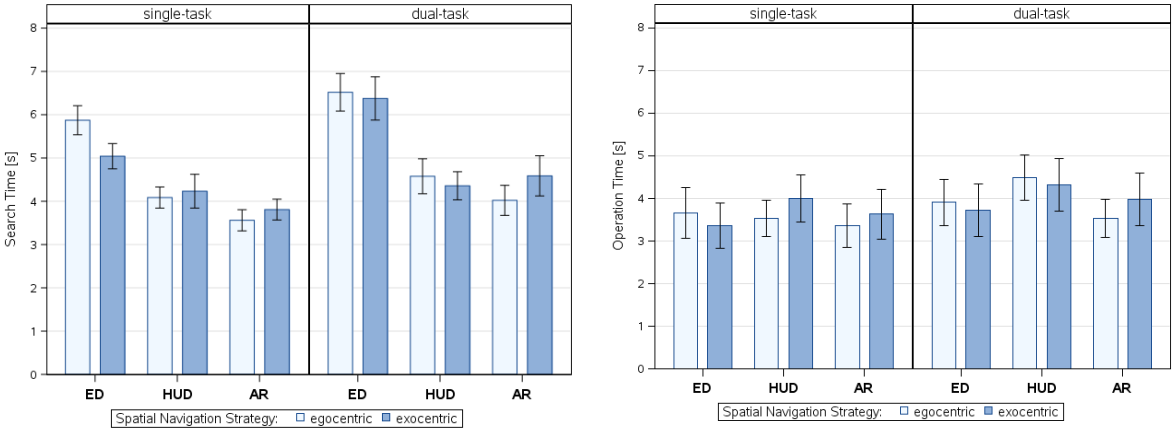


Figure 7.16: Mean completion time (with CI=95%) of the visuomotor task per display method grouped by the navigation strategy and paneled by trial type: *(left)* for the initial search, and *(right)* for the operation.

7.3.6.2 Secondary Task Performance by Spatial Navigation Strategy

Mental demands, undiscovered by the primary task performance, can be detected by concurrent task solving. Thus to verify whether the participants' level of resource allocation complied with their spatial navigation strategy, I analyzed the secondary task performance with respect to the three display conditions. Also for the mental load, I expected a decreased workload under an egocentric orientation of the display for participants using an egocentric strategy and vice versa for exocentric participants. Therefore the response time to the auditory stimuli was evaluated with respect to the process of localizing the task area and operational processes. The related mean value plots are

presented in Figure 7.17, while the results of multiple comparisons of the response times between the navigation strategies at the same stage of display condition are listed in Table 7.14, but showing the significant differences only. The data revealed interesting findings, pointing out that participants using an egocentric strategy, always responded significantly slower than exocentric participants, especially during the initial search, where the differences were significant under all display conditions. More astonishing is that the mean response time during the initial search showed the highest increment for egocentric participants under the AR condition, where they responded 35 % slower than participants using an exocentric strategy. That is surprising because the AR condition should be the most suitable for egocentric participants and not the one demanding most resource allocation. This discovery should continue to be thoroughly investigated in future research. Nevertheless, and in conclusion, these outcomes indicated that such kind of task will induce an increased workload for egocentric participants and such tasks will benefit more from operators using an exocentric strategy.

Variable	Display		Strategy		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Response Time	AR	-	ego	exo	168	4.67	<.0001	<.0001	*
(search) [ms]	ED	-	ego	exo	256	3.82	0.0002	0.0002	*
	HUD	-	ego	exo	186	3.35	0.0010	0.0010	*
Response Time	AR	-	ego	exo	149	4.25	<.0001	<.0001	*
(operation) [ms]	ED	-	ego	exo	144	2.71	0.0075	0.0076	*

Table 7.14: Significant differences between the navigation strategies at the same stage of display method related to the response times of the secondary task using SAS® PROC MIXED (adjust=simulate).

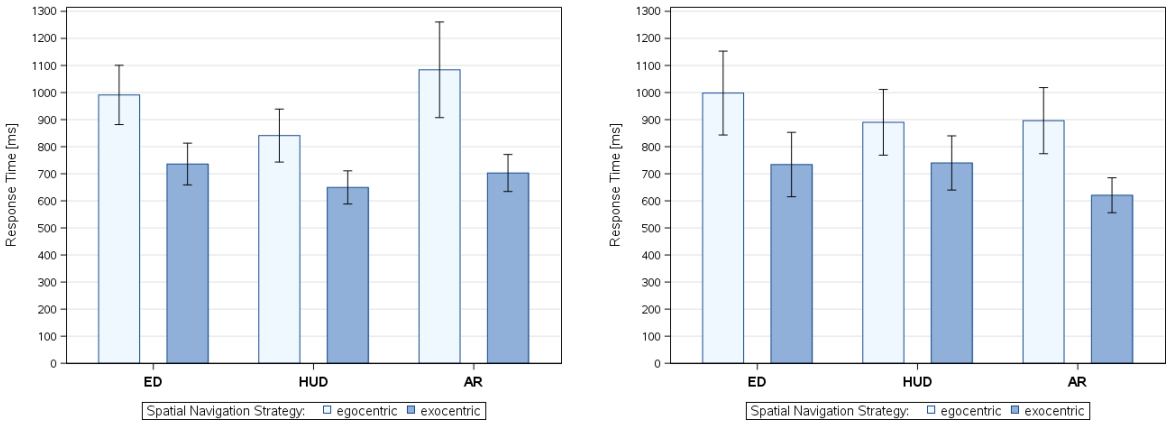


Figure 7.17: Mean response time (with CI=95%) of the secondary task per display method grouped by navigation strategy and paneled by trial type: *(left)* for the initial search, and *(right)* for the operation process.

7.4 Discussions and Revisiting the Research Questions

Because the analysis of the data considered multiple aspects of usability measures including various performance and workload indicators, the discussion is divided into related subsections. At the end of each subsection, its related research questions that were asked in section 7.1, will be answered. To inform the design about display settings for payload operations, the findings will be recapitulated in the following conclusion presented in section 7.5.

7.4.1 Discussion on Visuomotor Task Performance

The effectiveness of the studied display conditions (ED, HUD, AR) was assessed by analyzing the visuomotor performance of goal-directed pointing (see section 7.3.1), which was studied under two levels of task difficulty (single-pointing, multiple directional-pointing) and two types of trial, implying either the condition of single-task performance or performance that was interfered by the request to perform a secondary task, whereby the margin between these trial types will be later discussed in the workload section. The visuomotor task performance was analyzed with respect to time differences as well as differences in accuracy and required head movements. Evidence indicated that compared to the exocentric display (ED), egocentric displays (HUD, AR) significantly improved the visuomotor performance by reduced session and task completion times. This partially supports the findings of Henderson and Feiner (2011a), showing reduced times for task completion with AR compared to their external display, but they could not show significant differences for the overall completion time. In contrast to the spatial orientation of the display, the spatial orientation of target cues within an egocentric display (HUD, AR) did not affect the session and task time that in turn also supports the findings of Henderson and Feiner (2011a), showing no differences between their AR and HUD conditions. Because the execution of a payload operation includes the localization of the target area before performing the task, the visuomotor task was separately analyzed related to the initial search and the task operation. Thereby the exocentric presentation schemes (ED, HUD) showed the complete task information from the beginning, i.e. the task information was already available during the search process, while the egocentric AR display provided an initial navigation to the off-screen target area before getting the task information by 3D-registered overlays. As was expected, the results showed a significant increased search time using the external display (ED) that vastly deteriorated the performance compared to the egocentric displays (AR, HUD), with which the target area was localized 29 % faster on average. This finding reinforces the general characteristic of an exocentric navigation behavior, which implies additional costs of retrieving the task information by initial learning, whereby its level of difficulty depends on the number of targets. In contrast to the spatial orientation of the display, the spatial orientation of the target cues visualized within an egocentric display (HUD, AR) affected the search process differently, depending on the type of trial. Thereby the localization of the off-screen target area under the single-task condition benefited from significant shorter search times using the AR display compared to HUD, as was also shown by Tönnis & Klinker (2006) and Robertson et al. (2008) but against the findings of Khuong et al. (2014) showing that an in-situ 3D visualization located side-by-side can outperform AR overlays. However, the interference from concurrent task performance mitigated this difference and revealed only 4 % faster searching with AR compared to the exocentric HUD visualization. Contrary to slow searching and regardless of the level of task difficulty, the exocentric display (ED) supported fast operation time that suggests that the task information was mainly memorized during the search process and if in-between-checks were required, the user's spatial position and orientation needed to be updated only to the reference frame and not to each object in the environment (exocentric nature). Thus, it is not really surprising that the operation time using the external display (ED) did not significantly differ from the AR display, because the egocentric nature of the AR display requested a spatial update relative to each object if the user's position or orientation has changed, whereby the localization of a target after changing the viewpoint, is affected by the number of objects in the environment (Wang et al., 2006). However, while the exocentric in-view area map (HUD) still supported fast searching, the subsequent operational process was most disadvantaged under this display condition, especially under dual-task conditions, where the HUD visualization caused significantly slower operations for both level of task difficulty. Although the HUD presentation scheme implied an egocentric display, the task

information was presented in an exocentric fashion, resulting in fast localization of the target, but also in slower operation. This may indicate that during the search process an initial memorizing of the task information was not triggered as indicated by the external display, and thus the operation process might has required permanent visual verifications with the in-view area map, which in turn only called for eye movements. Because the fact that the HUD display caused the longest operation times already for the single-pointing task, which only implied one task instruction, it can be assumed that the spatial update followed an egocentric approach as was induced by the egocentric AR display. Regardless of the display condition, the results also indicated that the level of task difficulty is not an indicator for varying time differences during the operation process, in particular under the condition of dual-task processing. In contrast to these time effects, the pointing accuracy was not affected by the spatial orientation of the display and target cues, which is contrary to the findings from Henderson and Feiner (2011b) showing that AR lead to significantly more accurate performance, but was measured during related alignment activities. Last but not least, the performance of the visuomotor task was assessed by analyzing the resulted head movements (see section 7.3.2), which revealed significant differences for the rotation around the vertical axis (yaw) between all display conditions, whereby the head rotation was significantly minimized using the HUD scheme. This is contrary to the generalized conclusion of Henderson and Feiner (2009, 2011a) stating that 3D-registered AR can reduce head and neck movements during instructional tasks. Compared to their work, the research presented here implies a larger working volume. This posits the thesis that the needed body and head movements are addressed to the size of the covered working volume, whereby a larger volume will increase head movements during AR task localizations and thus the physical effort spent on them.

Hence, the research questions related to the visuomotor performance (Q1) can be answered as follows:

- *Do egocentric displays (HUD, AR) speed up the visuomotor performance?*
  Yes. The findings have shown that egocentric displays (HUD, AR) supported faster performance of a task and a session covering several tasks than the external display (ED), and thus extend the findings of the previously conducted proof-of-concept study (see section 6.4), comparing the traditional SSC-based interface with AR support.

- *Do egocentric displays (HUD, AR) enhance the pointing accuracy?*
  No. The results have shown that the pointing accuracy is generally not affected by the spatial orientation of the target cues.

- *Is there a significant improvement in performance induced by egocentric AR target cueing?*
  Partially. It could only be shown that the operational process, after localizing the off-screen target area, was significantly faster performed under dual-task processing with the conformal AR display than with target cueing using an in-view area map provided by the HUD display.

- *Does the egocentric AR presentation generally reduce head movements as concluded by Henderson and Feiner (2009, 2011a)?*
  No. Such guideline is addressed to the covered working volume and thus is depending on its size, whereby a larger volume will increase head movements during task localization, especially with AR. While Henderson and Feiner showed that AR reduced head movements, the presented study has shown that in-view target cueing using the HUD mostly decreases head movements.

- *What kind of spatial reference frame will be triggered by egocentric displays with exocentric target visualization (HUD)?*

  It seems that an in-view area map is featuring the characteristics of an egocentric reference frame, because it did not trigger initial learning of the task information, as was shown by the exocentric ED display and it can be assumed that the in-view area map (HUD) claims egocentric spatial updating, which requires that the user need to update the location information for each object in the environment as the viewpoint changes, which could be explained by slow operation times, even for a low level of task difficulty.

7.4.2 Discussion on Multi-Dimensional Workload

The workload was multi-dimensionally assessed by a dual-task paradigm, physiologically by cardiac responses and subjectively by a rating scale. With the dual-task paradigm I studied the availability of attentional resources yielded by the studied display conditions, whereby the visuomotor performance was interfered by an auditory reaction-time task. The dual-task interference was evaluated by the level of concurrent task integration measured by the time costs of dual-task processing, as well as evaluated by the performance of the secondary reaction time task that should inversely reflect the workload of the primary visuomotor task.

Dual-Task Costs: The results showed that the timing costs for concurrent task integration (see section 7.3.1) revealed significant differences for all display conditions, reaching to around 13 % decreased task completion times. However, considering the session time, covering twelve successive tasks, the dual-task costs only yielded significant effects for the exocentric visualizations (ED, HUD) and not for the egocentric AR visualization, which may be an indicator for improved dual-task integration with AR during longer tasks. Further results varied for the initial search and the operation processes. During the localization of the target area, the dual-task processing significantly increased the visuomotor performance for the external display and the egocentric AR display only, but not for the HUD display. In contrast, during the operation process the dual-task processing caused a significant time increase for the HUD display only, independent of the level of task difficulty. Altogether, the findings have indicated that using the HUD display demanded less attentional resources during the localization of the off-screen target area, while the same display condition was disadvantaged by dual-task integration during the operation process, which has benefited more from the egocentric AR display and even from the external display (ED), whereby the latter one may change with a higher number of targets needed to be learned.

Secondary Task Performance: To evaluate the performance of the secondary reaction-time task I analyzed the response time to the auditory stimuli and the response rate as ratio between missed and triggered signals (see section 7.3.3). While there was no effect of the display settings on the response rate, the response time was affected by the display setting. Thereby the external display caused the slowest responses, but only significantly differs from the exocentric visualization within an egocentric display (HUD) that revealed slightly faster responses than with the egocentric AR display. Contrary to my expectation, the response time under the egocentric AR condition was prominently increased during the search process and was even similar to the mean response time using the external display (ED), indicating that the mental workload during the localization process (visual search) was increased under both conditions, AR and ED. As early hinted, the response time under the HUD condition was faster and significantly different from the external display and thus it has caused the

least workload level on divided attention for the localization process, which was already indicated by the lowest cost of dual-task processing. This suggests that target cueing by an in-view area map (HUD) has allocated less mental processing resources during the initial localization process than off-screen target cueing (ED) and in-situ conformal cueing (AR). In contrast, the response times during the operation process showed fastest responses under AR and the slowest under the external display (ED) condition, but did not yield significant differences. A trend analysis investigated the relationship between the response time and the overall session time with respect to the display conditions. This should provide information about their correlations and was used to predict the development of the workload levels by linear regression. The correlation analysis revealed significant correlations with a negative relationship for all display conditions, indicating as the session time increases, the response time decreases and thus the workload level too. Thereby the two variables moderately correlated for the HUD display, strongly for AR and very strongly for the external display (ED). Roughly the same ratio was shown by the coefficient of determination revealed by linear regression, reflecting how well the data fitted its prediction models. The resulted models have indicated that the workload will be decreased fastest with the external display (ED), but also started with the longest response time. In contrast, using the HUD display provided the lowest level of workload in the beginning, but revealed the slowest reduction over time. However, the predicted values of the egocentric AR display revealed similar characteristics than the ED model. Thereby the session time predicted that response times will decrease slightly faster than under the HUD condition, but showed also that using the AR display causes a higher workload in the beginning. Hence, the analysis of the secondary task performance, including the dual-task costs, revealed different findings, suggesting that an in-view area map provided by the HUD display will demand the least amount of attentional resources during the localization of the off-screen target area, while the in-situ AR display will contribute to a decreased workload during the operation process and for the overall duration. Although the external display significantly showed a poorer primary task performance, it seems that the workload level will not suffer over time in the way as assumed.

Physiological Workload: The physiological workload (see section 7.3.4) was assessed by the LF/HF ratio, a frequency domain parameter of the HRV, indicating the sympathovagal balance. As was already shown by Tümler et al. (2008), Schwerdtfeger et al. (2009) and Grubert et al. (2010), it is hard to get significant differences from data with high individual variations, although normalized by baseline measures. Thus, I can only discuss differences by their mean variations, which have varied according to the trial type. While under the single-task condition the visuomotor performance did not show noticeable variations of the LF/HF ratio between the display conditions, however the performance under the dual-task condition showed the most prominent development of the parasympathetic response for the HUD display, which suggests that target cuing by the in-view area map (HUD) provided the most relaxed state. This could also be related to the fact that the HUD condition caused least head movements. Although higher under dual-task processing as with HUD, the egocentric AR display did not cause variations between the trial types (single-task, dual-task), while the external display showed a slight increase under dual-task processing.

Subjective Workload: Finally I evaluated the workload by subjective experiences of the workload using the NASA RTLX rating scale (see section 7.3.5), whereby the perceived workload was rated as lowest in favor of the egocentric AR display, but yielded significant differences only by comparing with the external display (ED), while not with the HUD display. Showing that the subjective workload most benefitted from AR supports, on the one hand, the findings of the initial

proof-of-concept study (see section 6.4) and, on the other hand, the findings of Robertson et al. (2008) and Medenica et al. (2011).

Although the subjective workload benefited from in-situ target cueing (AR), the indicators based on objectively measures have confirmed this only partially, while even some findings disclosed the contrary. It could be shown that the AR performance over the complete session and during the operational process was not interfered by concurrent task processing and thus needed to allocate less mental resources, whereby the retention of processing resources during the operational part was not exclusively ensured by AR target cueing, because the same was also proven for the external display (ED). However, the evidence that target cueing by an in-view area map (HUD) contributed with consuming significantly least processing resources during the localization process was more clearly confirmed, on the one hand, by non-significant differences between the trial types, exclusively shown for HUD, and on the other hand by fastest response times for the secondary task, although significantly differing only from the external ED. It could be proven in almost the same manner that target cueing using the HUD display disadvantaged the subsequent operational process. Findings related to the physiological assessment could also not confirm the subjective preference in favor of the in-situ AR target cueing, but on the contrary, it could be shown that the HUD supported a more relaxed state under dual-task conditions, which suggests that using the in-view area map (HUD) provided a more optimal level of arousal, supporting an increased performance quality. Consequently and considering only the visual search assisted by the AR display, the related findings support the assumption that AR may cause an increased level of stress or arousal, which in cases of overarousal can induce attentional tunneling (see section 3.3.2) as was already discussed by Grubert et al. (2010).

Hence, the research questions related to the operator workload (Q2) can be answered as follows:

- *How is the spatial orientation of displays affected by dual-task costs?*

  The egocentric AR display and the external display were affected in the same way. Both were interfered by dual-task processing during the initial search, while the subsequent operational process was freed from interferences. The contrary could be shown for the HUD display, not affecting the initial search, while hampering the operation itself. But it can be assumed that longer tasks will benefit most from an egocentric AR display, because the results have shown that a session covering several tasks was not affected by dual-task interferences.

- *Are egocentric displays (HUD, AR) decreasing the workload?*

  In general, yes. Although both egocentric displays made their own contribution, depending on the workload indicator, the results have shown that using an egocentric display can decrease the workload level related to subjective experiences (AR) and the physiological effort spent on task performance (HUD). However, the costs caused by dual-task processing did not differed between the egocentric AR display and the exocentric ED display, while the egocentric HUD display contributed in its own way (see answer of prior question).

- *Is egocentric AR target cueing providing the lowest workload level?*

  Yes and no. While the subjective workload was clearly rated in favor of the egocentric AR display, the resulting workload level assessed by physiological responses was higher than the level resulted from the HUD display, which provided the most relaxed state. Alike was seen by the effort spent on concurrent task performance, whereby the HUD display provided the fastest responses in general and in particular during the initial search, but only differing from the external display (ED). Besides the decreased dual-task costs over the session and during the operation as prior

discussed, the secondary task performance with the AR display showed only a tendency towards faster responses and thus to consume least processing resources during the operational process, but not significantly differing.

- *Does the workload caused by egocentric AR target cueing decrease over time?*
  Yes. A resource trend analysis has shown that the response time of concurrent task performance decreased over the session time, whereby the two time variables strongly correlated. But similar could also be shown for off-screen (ED) and in-view (HUD) target cueing.

Finally it should be noted that this way of assessing the workload level has shown the importance of using multi-dimensional indicators. Usually the assessment of workload is not commonly practiced in AR usability testing (Bai & Blackwell, 2012). In the small number of studies, which have considered the mental or physical effort, the workload level was mainly assessed by only one indicator, which, in most cases, was analyzed by subjective experiences collected by the NASA TLX (Tang et al., 2003; Biocca et al., 2006; Robertson et al., 2008; Schwerdtfeger et al., 2009; De Crescenzio et al., 2011; Mulloni et al., 2011; Medenica et al., 2011), but also, although rarer, physiological assessed by HRV (Tümler et al., 2008; Schwerdtfeger et al.; 2009; Grubert et al.; 2010). Even less studies analyzed the workload by two indicators, such as HRV combined with subjective experiences (Schwerdtfeger et al.; 2009) or with secondary task performance (Grubert et al.; 2010), whereby the dual-task metaphor was only applied during pre-test and post-test phase. Although meaningful, workload subjectively rated by the participants did not consider issues of the task environment and thus can differ from the workload reflecting the effort spent on additional load or physical characteristics. So for example, the prior presented results have shown how strong the workload feedback assessed by the dual-task metaphor can be, because it was most indicative to quantify the amount of processing resources claimed by the primary task performance**.** However, biofeedback, like HRV, enables to identify, for example, physical stress, but its assessment is challenging. Because each method contributes with their own perspective, complementing each other, it is advisable to analyze the workload level by multiple dimensions.

7.4.3 Discussion on Individual Spatial Navigation Strategy

Another aspect of the presented study was addressed by the assumption that task localization can be affected by individual differences in the strategy used for spatial navigation. Therefore each participant performed the Turning Study[76] related to Goeke et al. (2013) that resulted in seven egocentric and six exocentric participants. Related to the outcomes of the visuomotor performance and the workload on divided attention I evaluated the margin between the individual navigation strategies at the same stage of display condition. The findings related to the visuomotor performance provided evidence that the spatial orientation of the exocentric ED display and the egocentric AR display seem to be linked with the corresponding individual navigation strategy. But it was not clear which navigation strategy the HUD display has supported and thus I can assume that the target presentation provided by the in-view area map (HUD) will be equally suitable for both groups. In contrast to the primary task performance, the analogous analysis, with respect to the performance of the secondary task, revealed different and interesting findings. It showed striking high response times yielded from egocentric participants and, even more astonishing, this observation was independent from the used display condition. This suggests that such task operation will induce an increased workload level on divided attention for egocentric participants and that such tasks will more benefit

[76] Turning Study. URL: http://www.navigationexperiments.com/TurningStudy.html, last visit: 28.09.2016

from operators using an exocentric strategy. This and the assumption that the exocentric ED display is linked with the exocentric navigation strategy, could also explain why the workload level under the ED condition did not suffer overall in the way assumed. Such affected workload induced by egocentric participants, seems very important to me as selection criterion to find the most suitable operator for such kind of tasks and where the amount of attentional resources, providing a decreased level of mental workload, will be a crucial key factor. Consequently, these findings suggest that astronauts' individual predisposition for visual spatial processing should be determined and studied related to the display setting for payload guidance, and more generally in the context of Augmented Reality. Thus, future research on AR interaction needs to consider this issue, especially where spatial navigation will be involved. The neglect of this spatial predisposition may be the reason why some studies investigating presentation schemes for task localization (see section 4.1.2), revealed conflicting findings. So, for example, Tönnis et al. (2005) have shown that an exocentric bird's eye view outperformed an egocentric 3D arrow visualization to direct a car driver's attention. After optimizing the 3D arrow, the follow-up study revealed opposite results (Tönnis & Klinker, 2006). The same can be applied to the contrary performance results of the works conducted by Schwerdtfeger et al. (2009) and Grubert et al. (2010). While Grubert et al. reasoned, amongst other, a significant better AR performance by using an attentional funnel navigation, whereby the same navigation metaphor applied by Schwerdtfeger et al. caused a higher error rate and supported only a slightly faster picking process. Such attentional funnel has exclusively egocentric-aligned features, and thus it is possible that the number of participants, using an egocentric strategy by nature, was increased during the study from Grubert et al. (2010). The study results presented by Biocca et al. (2006) and Schwerdtfeger & Klinker (2008) can be similarly discussed, because they also compared attentional funnel navigation with alternative presentation schemes for task localization, both revealed conflicted findings. Schwerdtfeger and Klinker (2006) have already recognized that there is a need "[...] for a suitable classification for users concerning their amenability to spatial concepts, as needed in AR/VR applications and games" (p. 1006). In general all performed studies investigating ego- and exocentric localization methods (such as presented in section 4.1.2) need further verification by considering the individual navigation strategy.

The research questions related to participants' spatial navigation strategies (Q3) can be answered as follows:

- *Does the spatial orientation of the displays correspond with the individual navigation strategy?*
  Yes. The results related to the primary visuomotor performance revealed such correlation for the egocentric AR display and the external display (ED) for the initial search, but there was no visible assignment for the in-view area map provided by the HUD display.
- *Does the performance with the egocentric AR display benefit from egocentric users?*
  Yes. The initial visual search with the egocentric AR display was faster performed by participants using an egocentric strategy, but just under the dual-task condition.
- *Does the performance with the exocentric ED display benefit from exocentric users?*
  Yes. The visual search with the external display (ED) could benefit more from participants using an exocentric strategy, but this only under the single-task condition, although the performance by exocentric participants generally revealed a better quality related to the amount of processing resources than participants using an egocentric strategy.

- *Does the individual predisposition for visual spatial processing need to be considered in display settings for payload guidance?*

  It cannot be answered with a simple yes or no. According to the findings, showing that participants using an egocentric strategy will consume more processing resources in general, it would consequently be the best that such operations are exclusively performed by operators using an exocentric strategy. Alternatively, it is suggested that at least the target cues should be presented by an in-view area map, as was provided by the HUD, because the localization of the off-screen target area using this display was not affected by the individual navigation strategy.

7.5 Conclusion

Firstly, it should be emphasized that the assessment of task performance partially revealed different results under dual-task processing than under single-task processing. While the investigation of specific tasks constitutes real-world scenarios and related experiments are mostly performed in isolated laboratories, it seems important to me that such confounding factor should be compensated by dual-task performance. Taking this into account, the following conclusion is based on the findings revealed from the dual-task performance and aims to inform about the spatial orientation of the display and visualization settings for space payload operations.

The findings related to the visuomotor performance are suggesting that for payload operations the traditional external display needs to be replaced by an egocentric display, like a HMD. When using such a display, the spatial orientation of the presented instruction visualization depends on whether the localization of the off-screen target area should be supported, or the operation task itself. During the initial search process the instruction visualization is not dependent on the spatial orientation, although an exocentric area map within an egocentric display (HUD) significantly minimizes head rotations, which will be important when the working area, given by a task, implies greater dimensions. By contrast, once the target area was localized, the subsequent operational process should be supported by a 3D-registered superimposed AR visualization. Similar strategy is suggested by the findings related to the mental workload level, but providing clearer guidelines. They support the findings of the primary visuomotor task performance, showing that an egocentric AR display is most suitable to integrate concurrent task performance during the operational process. But the workload measure using the dual-task metaphor has also proven that the initial localization of the target area will significantly benefit from in-view target cueing provided by the HUD display. Hence, to inform the design about display settings for payload operations, I would recommend a hybrid approach, using a static in-view area map (HUD) to localize off-screen target area and conformal in-situ guidance provided by AR for the subsequent operational process. Although the egocentric AR display was subjectively perceived as this display providing the least mental workload, in cases where applying AR exclusively, especially issues on navigation support, it should be thoroughly examined whether perceptual narrowing by the tunnel effect has been avoided. This study has also shown that such operations can be affected by individual differences in spatial navigation behavior. To ensure the best quality of performance, it is recommended to use only operators having an exocentric navigation strategy.

7.6 Summary

In this chapter I investigated the effects of the spatial orientation of the display and the target visualization for payload guidance on performance and workload. For comparison I set up three different presentations schemes, each with its own characteristics in the spatial reference frame for the display and the visualization. This resulted in an exocentric off-screen display, commonly used for payload guidance, and two egocentric HMD displays (AR, HUD), which differed in their spatial orientation of the target visualization, whereby it was not clear which kind of spatial updating will be triggered by an in-view area map provided by the HUD display. In a within-subject design thirteen participants conducted a series of visuomotor tasks, varying in their level of task difficulty. This series of tasks needed to be conducted twice, whereby one trial represented the usual single-task condition and the other one the same performance but interfered by a simple reaction-time task. The usability of the display settings was assessed in a multi-dimensional way using primary task measures in conjunction with various workload indicators, such as: assessing the amount attentional resources, analyzed by secondary task measures, and the costs of dual-task processing, subjective experiences obtained from analyzing the rating scores of the NASA RTLX and physiological responses assessed by the heart rate variability using the LF/HF ratio. Besides the session and task completion, to analyze the primary and secondary task performance, an additional distinction was made between the initial search of the off-screen target area and the subsequent operation. Furthermore, it was also examined whether individual differences in spatial navigation behavior are affecting such visuomotor tasks. Therefore it was previously determined which spatial reference system can be assigned to each participant. Thus, the study revealed a corresponding number of results, whereby the most important findings will be summarized here. With respect to the main purpose of this study, it was proven that an egocentric display provided by an HMD outperformed the exocentric display usually used for payload operations. Although subjectively rated as display providing the least workload, the analysis of related objective data revealed different results regarding the egocentric AR display. On the one hand, it caused a more aroused physiological state than the HUD display under the dual-task condition, where additional processing resources were needed for concurrent task performance. On the other hand, handling the effort spent on dual-task performance using the AR display showed that the resulted workload level on divided attention was affected in the same way as the external display, which consumed more processing resources than target cueing by the HUD display, especially during the initial search. The same was indicated by the costs of dual-task processing, assessed by the margin between the trail types. Thereby the results showed that in-view target cueing by the HUD display was the one display that was not interfered by concurrent task solving during the initial search, but was also the one that was interfered during the subsequent operational process. Consequently, to support payload operations it is suggested that the display setting should be following a hybrid approach, using target cueing by an HUD display to localized the off-screen area, while the subsequent operation should be supported by in-situ conformal cueing using an AR display. Apart from these display related issues, the findings also revealed that the individual predisposition of humans for spatial navigation seems to correspond with the spatial orientation of the exocentric external display and the AR display, This should be considered in AR research where spatial navigation is involved. Related to this, an interesting finding pointed out that tasks like studied here will induce an increased workload level for humans using an egocentric navigation strategy and will benefit from operators using an exocentric strategy.

The presented study was conducted under normogravity conditions, but it must be taken into account that the environmental condition for the application implicates changes in gravity that in principle

impair human performance (see section 2.4). Previous studies under long-term and short-term microgravity showed that the absence of gravity leads to human spatial disorientation, which can be additionally influenced by the intrinsic and extrinsic spatial reference frames (Gurfinkel et al., 1993; Glasauer & Mittelstädt, 1997, 1998; Harm et al., 1998; Lipshits et al. 2005). To study the effect of short-term microgravity on the spatial orientation of target cueing, a related proposal for parabolic flight experimentation was submitted to ESA, which was accepted in August 2015. The first associated experiment was already approved and ready to fly in the end of 2016. The fact that target cueing by an in-view area map provided by the HUD display decreases the workload and minimizes head rotations around the vertical axis, could be the key factors under altered gravity conditions.

Chapter 8

Placement of AR Input Devices in Altered Gravity: Effects of Short-Term Hyper- and Microgravity on Goal-Directed Pointing *(3DPick Study)*

During payload operations, astronauts interact with the input devices provided by the SSC laptop. This is needed to be adapted when AR is intended to be applied to such operations. Thereby one fundamental question is asked regarding their spatial placement, which is closely linked with the underlying hand-eye coordination. While in the previous chapter the focus was on the spatial orientation of the task cues, this chapter deals with the spatial placement of interactive AR interfaces used for control and symbolic input tasks. Such interfaces can be placed in reference to the world, the body or the head. Because the payload tasks are conducted aboard the ISS, the visuomotor coordination is affected by the absence of gravity and, thus, it is not clear from which placement reference the astronaut is benefitting more. Related microgravity research does not deliver an adequate directive. As a consequence, it is required to investigate the trade-off between the three placement references in consideration of altered gravity conditions. Taking this issue into account, this chapter reports on experimentation conducted under short-term altered gravity conditions, as provided by parabolic flight. Parts of this chapter have been published previously in Markov-Vetter et al. (2012, 2013b, 2013c, 2013e).

8.1 Research Objective

To navigate through a procedure or to enter a command-based instruction, the astronauts are used to operate the touchpad and keyboard of the SSC laptop. It is clearly evident that this no longer suits the framework of AR. Works on instructional AR support, neither beyond (see section 4.1.1) nor ahead the spaceflight domain (see section 4.2), have been considered to compensate the loss of input modalities, or used an additionally physical device (Henderson & Feiner, 2009), which is generally interfering with an AR display, like a HMD. As already discussed in section 5.2.3, before designing alternative AR interfaces, their spatial placement is more fundamental. Assuming that the elements of such AR input devices are selected by direct pointing, the placement is substantial for the associated hand-eye coordination. With such placement it can be determined whether the pointing targets are coded inside or outside of the human body frame of reference, are providing haptic cues or whether the interface is handled in a unimanual or asymmetrically bimanual way. As shown in Figure in 8.1, these characteristics can be assigned to a world, body or head reference, each implying its own specification. Thereby the body reference is restricted to hand-held interfaces, holding the interface in the non-dominant hand, while the dominant hand is used for pointing. Corresponding to the environmental condition of payload operations (see section 5.1.1, *Environmental Analysis*), the absence of gravity affects astronauts' visuomotor coordination and therefore also their hand-eye coordination, which in turn can be influenced by the spatial placement reference. As a result of the

discussion in section 5.2.3, it can be stated that although a world-referenced interface is promoted by unimanual handling and tactile sensation, it is disadvantaged by a outside-coded location that requires a spatial coordination transformation, which in turn can be affected by the absence of gravity. However, additional head movements, needed to orient towards such interfaces, can in turn contribute to more precise pointing. In contrast to the outside-coded world reference, both the body and the head reference enable to point towards inside-coded interfaces, not requiring a transformation between the instrinsic and extrinsinc coordinate systems. Although a body-referenced interface is favored by such inside coding, its handling will be a demanding one, because weightlessness affects limb proprioception, needed, for example, for goal synchronization of bimanual skills. On the other hand, its effect size under microgravity is reduced by the fact that the pointing movements are conducted in the body-related frame of reference. Just like for the world reference, the hand-held reference provides tactile sensation by haptic cues, which supports the proprioceptive restoration of possible movement disorder. This is not the case when pointing towards head-referenced interfaces, although their underlying hand-eye coordination decisively benefits from the inside coded location and the unimanual handling mode. Until now, there was no related microgravity research, studying the trade-off between these reference conditions, less than ever in the scope of AR, nor provided findings to give adequate indications which placement reference can significantly contribute to most reliable and fastest hand-eye coordination. Hence, the research reported in this chapter investigates the effect of altered gravity, especially of microgravity, on the spatial placement of AR interfaces, used for control and symbolic tasks. As also discussed in section 5.2.3, it can be assumed that bimanual handled interfaces contribute most to a reliable and fast hand-eye coordination under altered gravity conditions, which is representing the hypothesis of this research.

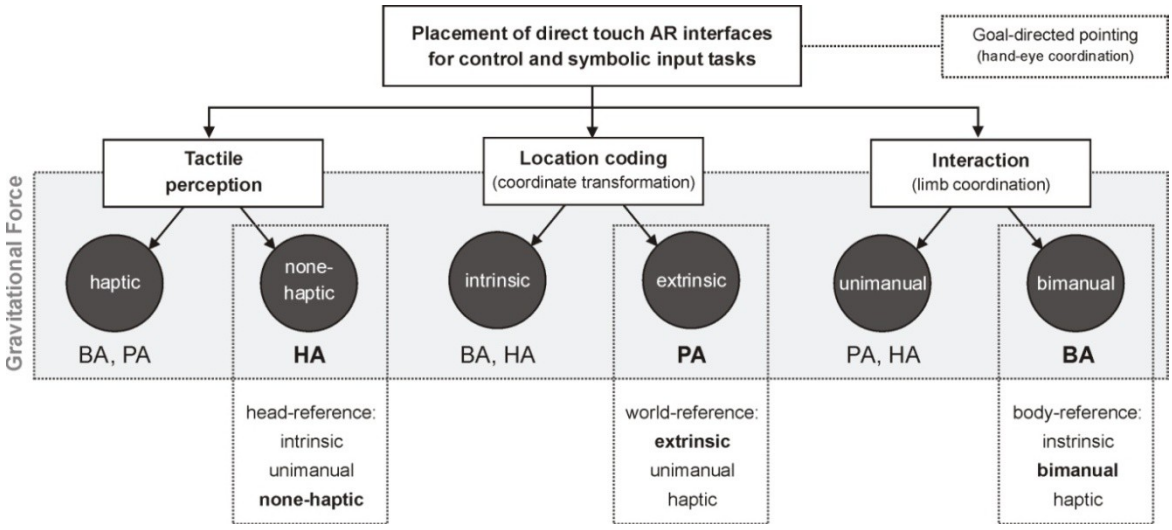


Figure 8.1: Placement condition studied under altered gravity resulted from sensorimotor indicators involved in hand-eye coordination to handle touch interfaces in AR for control and symbolic input tasks.

While the research presented in the last chapters was closely linked to payload operation, this research is detached from such a task requirement. Its findings are utilizable for applications intended to close the gap of missing input devices in an AR surrounding in regard to the broader issue on optimal placement, supposing direct touch selection by goal-directed pointing movements. The most important aspect of this research is the consideration of environmental changes induced by modified conditions of gravity. Thus, this research aimed to inform the design of interactive AR interfaces in their spatial placement, not only in normogravity, but also in altered gravity and particularly in

microgravity, which will be important for applications, supporting astronauts' work, for example, aboard the ISS. By comparing the visuomotor performance of hand-eye coordination resulted from world-, body- or head-referenced interfaces, the goals of this research were:

- to verify the feasibility of operating AR interfaces under altered gravity conditions,
- to investigate the effect of altered gravity on visuomotor performance and its demanded workload,
- to find out the most efficient placement of AR interfaces for input tasks.

For me this raised the following research questions, but bearing in mind that this research is considering altered gravity conditions:

Q-1: How the interface placement condition is affecting the underlying visuomotor coordination?

- **Q-1.1:** Does inside coded interfaces improve the visuomotor performance and the workload?
- **Q-1.2:** Does bimanual handling improve the visuomotor performance and the workload?
- **Q-1.3:** Does the absence of haptic cues deteriorate visuomotor coordination?

Q-2: How changes of gravity are affecting the visuomotor coordination?

- **Q-2.1:** Do changes of gravity cause timing effects?
- **Q-2.2:** Do changes of gravity affect the pointing frequency?
- **Q-2.3:** Do changes of gravity influence the pointing accuracy?

Although not fully comparable, considering research conducted under normogravity conditions, the most similar research was conducted by Harrison et al. (2011) and Lindeman et al. (1999). While Harrison et al. (2011) compared the pointing accuracy caused by hand-referenced interfaces bimanually handled and world-referenced interfaces unimanually handled, Lindeman et al. (1999) investigated the effect of haptic feedback on sliding and pointing tasks performed with a world- and a body-referenced interface, as well as they additionally compared the reference types. But both works have not considered aimed pointing towards targets referenced to the head. This was studied by Bowman and Wingrave (2001) comparing a head-referenced floating menu and a hand-referenced panel, with their *TULIP* system, whereby menu items are virtually attached to the fingers or palm and operated by pinches of the related hand's thumb.

8.2 Methodology

To study the trade-off between world-, body- and head-referenced AR interfaces, I introduced three placement conditions that are presented in Table 8.1. Thereby the world-referenced interface was stationary placed on a physical surface in front of the user (**PA** = physical aligned). Consequently, it provided haptic cues and required unimanual pointing towards outside-coded targets. For the body-referenced interface (**BA** = body aligned) I used a handheld interface that needed to be operated bimanually und provided haptic cues too. However, the head-referenced interface was aligned to the HMD's field of view (**HA** = head aligned) and required to point with only one hand, as was also requested by the PA interface, but without the presence of haptic cues. Both the BA and HA condition implied pointing towards inside-coded targets.

Interface	Body frame	Referenced to	Aligned to	Handling	Haptic Cues
PA	outside	world	physical surface	unimanual	yes
BA	inside	body	handheld	bimanual	yes
HA	inside	head	HMD	unimanual	no

Table 8.1: Identified placement conditions used for interactive AR interfaces and their characteristics.

Furthermore, to modify the condition of the gravity load, the study was conducted under parabolic flight conditions, providing short-term hypergravity and microgravity. Thereby the parabolic flight maneuvers provided a period of up to 22 seconds reduced gravity (microgravity, $\approx 0g$) per parabola, surrounded by increased gravity periods (hypergravity, $\approx 1.5g$ - $1.8g$). In general, during one parabolic flight campaign (PFC) researchers are able to perform their experiments within three days with one parabolic flight per day. Thereby one flight usually provides 31 parabolas in series that are interrupted by a two minute stable 1*g* period. More detailed information about parabolic flights is presented in section 8.2.1. To answer the research questions, I performed two related experiments during two PFCs, considering multiple levels of gravity load (1*g*-IN, 1.8*g* 0*g*), investigating the effect of altered gravity on interface placement. As shown in Figure 8.2 To reveal possible effects of flight conditions, the experiments were additionally conducted under normogravity conditions on the ground. Thereby I distinguished between performances directly before (1*g*-PRE) and after experimentation (1*g*-POST) that were compared to the normogravity performance in flight (1*g*-IN). Investigating Human Factors under parabolic flight conditions is challenged by potential occurrences of motion sickness. At this point, the experimenter needs to decide between the acceptance of invalid trials caused by motion sickness (high losses) and the treatment of motion sickness by Scopolamine (commonly used) with the risk of adverse effects. Because the approved number of participants for one flight is very limited and participants who suffer from motion sickness cannot be exchanged during the flight, it was decided that all participants received Scopolamine. Therefore the ground-based conditions (1*g*-PRE, 1*g*-POST) additionally enable to investigate the effect of Scopolamine on task performance and workload by comparing it with the normogravity level in flight (1*g*-IN). While under the 1*g*-PRE condition the participants did not receive the medication, the 1*g*-POST level took place directly after the flight in condition with medicine residue. For tests under hypergravity only the upward phases of the parabolic flight sequences were used, because these are usually more stable than the downward phases.

The trade-off study was framed by two independent variables (interface placement, gravity load) and complied with a within-subject design. During the first experiment three levels were assigned to the interface placement (HA, PA, BA) and five levels for the gravity load (1*g*-PRE, 1.8*g*, 0*g*, 1*g*, 1*g*-POST). Based on the results of the first experiment, the second one was only framed by two levels for the interface placement (PA, BA) and just four levels for the gravity load (1*g*-PRE, 1.8*g*, 0*g*, 1*g*). To detect possible effects of altered gravity on interface placement, I investigated the underlying hand-eye coordination resulting from the placement conditions. Therefore the participants were requested to perform a visuomotor task (see section 8.2.2), whereby the effectiveness and efficiency of the placement conditions were quantified by performance and workload measures. Because the experiment schedule and setups differed for both experiments, I would like to consider separately them for further description at this point. The description of the first experiment and its outcomes is presented in section 8.3, while the second experiment is presented in section 8.4.

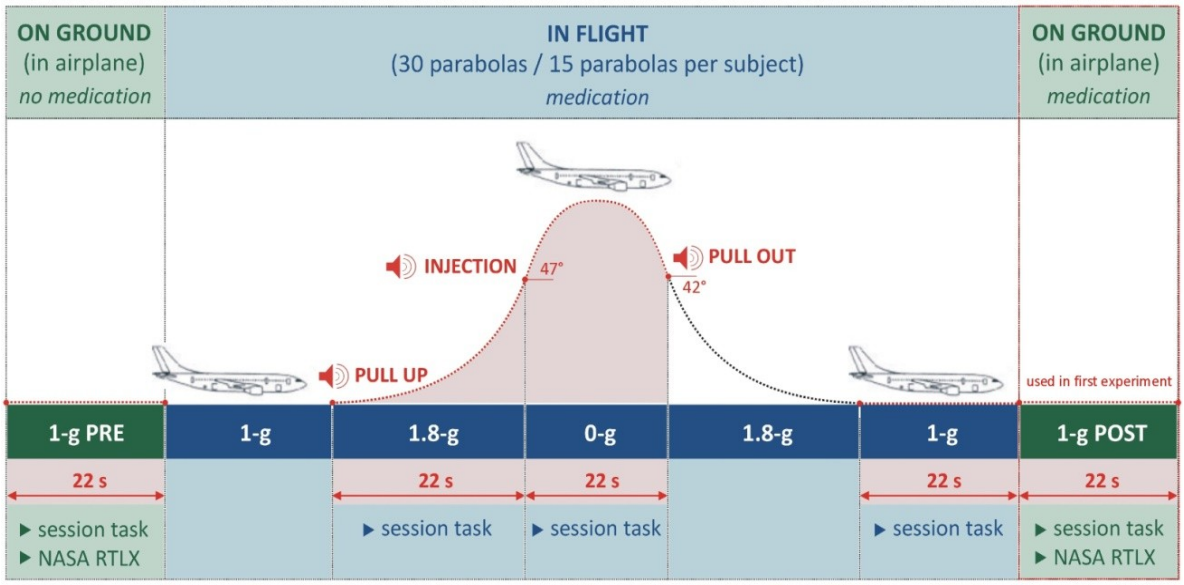


Figure 8.2: The *g*-levels and their characteristics used for the experiments.

8.2.1 Parabolic Flight

In Europe parabolic flights are managed and organised by the French company Novespace[77] based in Bordeaux–Mérignac since 1996 and initially used to train astronauts aboard the Zero-G Airbus A300 (see Figure 8.3, left) that was used until 2014 and was replaced by an Airbus A310 in 2015. In cooperation with the French Space Agency (CNES), the European Space Agency (ESA) and the German agency DLR, Novespace organises multiple campaigns per year, where scientists are selected to perform experiments and technological tests of space hardware under weightlessness conditions. The aircraft provides 230 m^3 of experimentation area (see Figure 8.3, right) and can accommodate 40 passengers. Per parabolic flight campaign (PFC) around ten experimentation teams are taking part.

Figure 8.3: The Novespace Zero-G Airbus A300: *(left)* the airbus during the upward hypergravity period [photo from *"NOVESPACE A300 Zero-G Rules and Guidelines"*, 2012]; and *(right)* front view of the experiment test area inside the airbus.

[77] NoveSpace: microravity, airbone missions. URL: http://www.novespace.fr/, last visit: 28.09.2016

Each campaign has its own flight profile (see Figure 8.4), which depends on the participated experiments, and usually consists of 93 parabolas with 31 per flight day. During a period of two hours a parabolic flight is a series of aircraft maneuver that provides approximately 22 seconds of weightlessness that is surrounded by an upward and downward hypergravity phase (1.5*g* - 1.8*g*). Thereby the first parabola is always used to get familiar with the environment and to test the experiment setup. After the initial test parabola, parabolas are usually executed in sets of five, followed by a short break of a five minute period, while after the 16th parabola the break takes eight minutes. The period between the start of each parabola is three minutes and consists of approximately one minute parabolic phase, followed by a two minute period of a steady horizontal flight (1*g*). With a microgravity time of 22 seconds this results in about twelve minutes weightlessness per flight and in about 34 minutes per campaign.

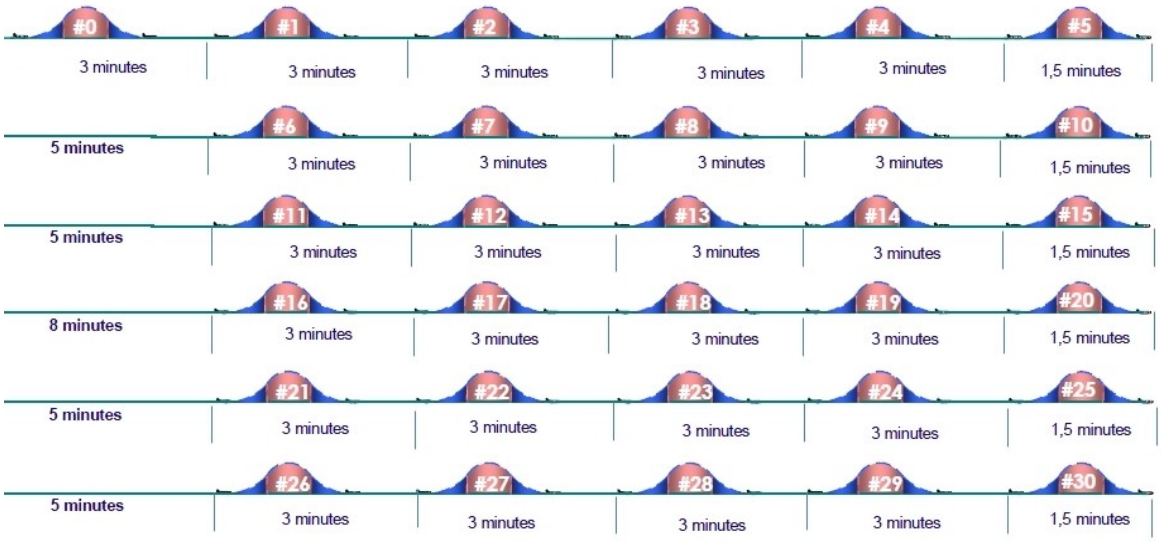


Figure 8.4: The parabola sequences per flight day of the 56th ESA PFC. [provided by Novespace, 2012]

As shown in Figure 8.5 and detailed described by ESA (2015), each aircraft maneuver begins with a steady horizontal period under 1*g* with a flight altitude of 6000 meters and a speed of 810 km/h. Then the pilot gradually pulls up the nose of the aircraft and it starts climbing at an angle of about 40 degrees that causes that the aircraft achieved nearly the double of the gravity level (1.5*g* - 1.8*g*). Then at an altitude of about 7500 meters, with an angle of around 47 degrees to the horizontal and a speed of 650 km/h, the aircraft engine thrust is reduced to the minimum required to compensate for the air-drag. At this point the aircraft follows a free-fall ballistic trajectory, i.e. a parabola, lasting approximately 20 seconds, during which weightlessness is achieved. The peak of the parabola is achieved at around 8500 meters, at which point the speed has dropped to about 390 km/h. After the microgravity period, the aircraft must pull out of the parabolic arc, which causes the second hypergravity period, und returns into steady horizontal flight (ESA, 2015a). During the flight, the pilot gives audio announcements through the cabin speakers and inform the experimenters about times and angles, as well as makes commands for the beginning of the hypergravity periods (*"pull up"*, *"pull out"*) and the microgravity period (*"injection"*). The audio commands were used for timing the experiments in flight.

The design of an experiment to fly aboard the Zero-G Airbus is subject to strict guidelines specified by the company Novespace. Designing an experiment usually includes the construction and setup of a

flight rack. Related to the Airbus interface, the provided guidelines detailed describe the approved rack limitations for size, weight, maximum load, power supply, materials, etc. For example, until 2014 it was only allowed to use anodized aluminum profiles manufactured by Item (Solingen, Germany). From 2015 it has only been allowed to use profiles built by Bosch (Stuttgart, Germany). Additionally the experimenters must provide a complete hazard analysis where all risks are itemized that may arise from the experiment setup.

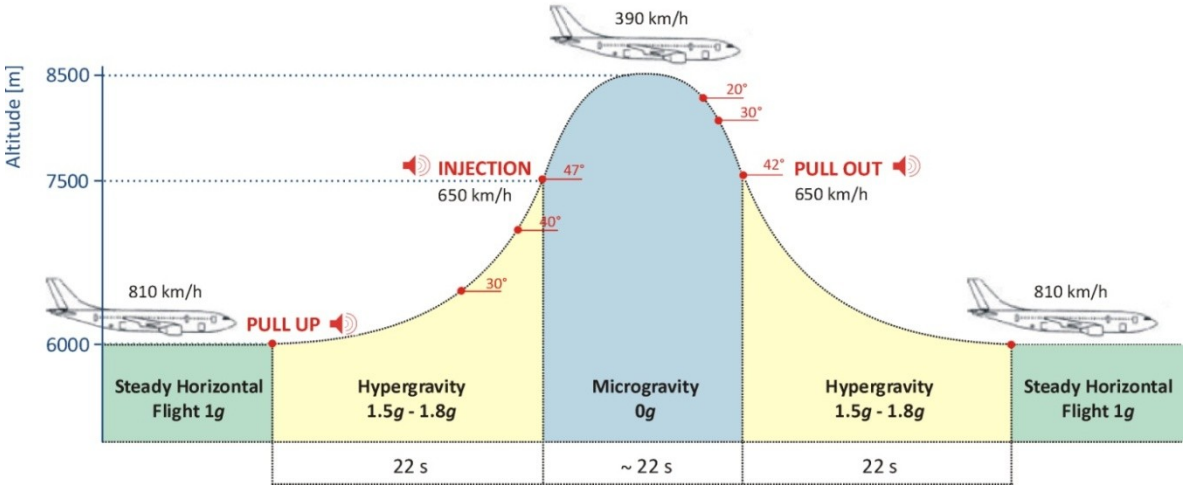


Figure 8.5: Aircraft maneuver of a parabola sequence showing the flight attitudes, the flight angles, the speed and pilot's audio announcements (red marked).

8.2.2 Experiment Task

The design of the task to be performed during parabolic aircraft maneuvers requires a short, repeatable, but still realistic task. For investigating the performance and workload of controlled AR pointing I used a discrete aimed pointing task and implemented a virtual keyboard located in 3D space (see Figure 8.6) with respect to the QWERTZ layout, which is provided by a usual keyboard. Wearing an optical see-through HMD the participant was requested to confirm letters by goal-directed pointing movements in response to visual stimuli. The pointing movements were performed by the index finger of participant's dominant hand. The stimulated pointing targets were computer randomized and signaled in green. By confirming a correct target, it was signaled in red.

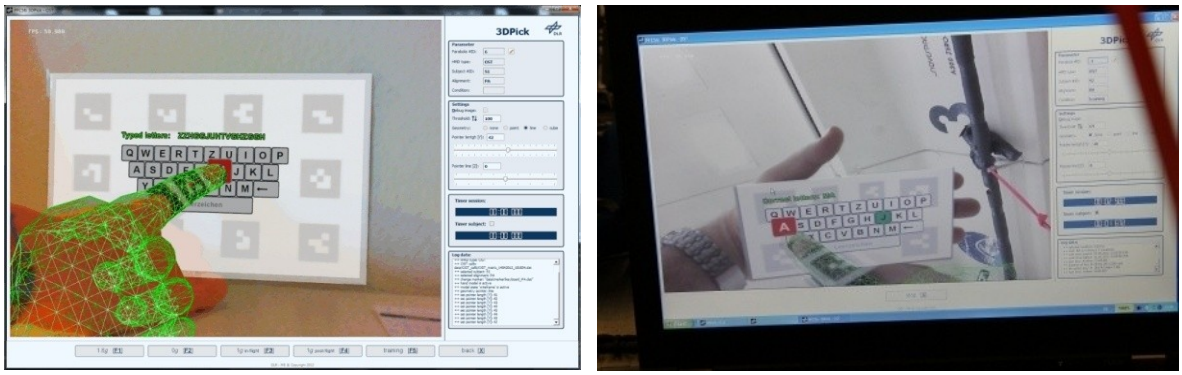


Figure 8.6: The experiment task of goal-directed pointing towards a virtual AR keyboard used for parabolic flight studies: *(left)* with the world-referenced PA interface, and *(right)* with the bimanually BA interface.

Pointing towards a false target was also indicated in red, while the target key was still signaled. Visual targets were successively triggered for 22 seconds as this was the duration of a single phase of hyper- or microgravity. For the bimanual handling mode (BA) the participant's non-dominant hand should held a surface for displaying the virtual keyboard (see Figure 8.6, right). In case of the PA condition (see Figure 8.6, left) a physical surface was installed in front of the participant allowing him to reach the pointing targets in a comfortable way. The head-referenced keyboard was virtually aligned to the HMD at a proper distance for interaction space and was individually adjusted for each participant. To get familiar with the experiment task and in handling the AR interfaces each participant completed an intensive training session using the final flight rack.

8.2.3 Apparatus

The experiments were conducted aboard the Novespace Zero-G Airbus A300 (see Figure 8.3, left). To interface AR, the monocular optical see-through dataGlass2/a from Shimadzu was used (see Appendix B and Figure 8.7, left), which was equipped with an optical sensor for optical inside-out marker tracking. The HMD was mounted on a bicycle helmet, that allowed a quick change of the HMD setup and was connected to the data processing unit (Lenovo Thinkpad T420s, 2.8 GHz CPU, NVIDIA Quadro NVS 4200M). To compute the position of participant's eye relative to the optical sensor, each participant performed a self-calibration related to Kato and Billinghurst (1999) in the aircraft and before starting the first test series. To realize pointing towards the interfaces different physical marker configurations for optical tracking were applied: (1) a single marker for pointing that is attached to the fingertip of the index finger of participant's dominate hand, (2) a multi-marker configuration for the hand-held panel (BA methods) for participant's non-dominate hand, and (3) a multi-marker configuration for physical external panel (PA methods).

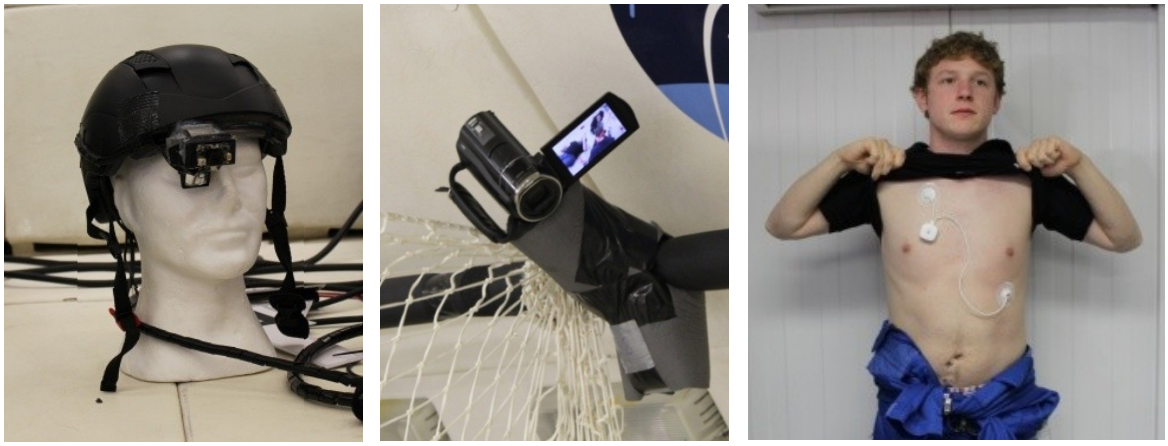


Figure 8.7: Parts of the apparatus used for parabolic flight studies: *(left)* the optical see-through HMD, *(middle)* the video camera installed at the handrail, and *(right)* the wireless HRV sensor attached to participant's thorax.

To ensure correct fitting of the pointing marker, a virtual hand model was visualized and the subject was requested to align his fingertip to the virtual fingertip, which was extended with a virtual line and a small sphere at the end, whereby the sphere was responsible for intersecting with the virtual interface targets. During experimentation, the operator could individually readjust the line together with the sphere in real-time, to ensure that the fingertip intersects with the foremost point. I implemented the experiment software in C++ using Qt-4.8.5 for the 2D operator interface, osg-3.0.1

for rendering the 3D content and osgARTProfessional-1.1.3 for marker tracking. For cases of unexplainable effects and to reconstruct the in-flight performance later, the complete period of the experiment was recorded using a video camera (Sony HD Handycam®, see Figure 8.7, middle) that was mounted at an aircraft's handrail above (during the first study) and directly opposite the site (during the second study) of the experiment setup. During the second experiment the physiological workload was additionally assessed and measured by the heart rate variability (HRV). As it has already been used in the study presented in Chapter 7, the participant was equipped with a wireless eMotion HRV sensor from Mega Electronics. The HRV electrodes were placed under the flight suit to participant's thorax (see Figure 8.7, right).

To fly the experiments aboard the Zero-G Airbus A300 I designed and built up flight racks with respect to the strict guidelines (see section 8.2.1). The schematic and setup of the flight racks were different between both experiments. Therefore the experiment setups are explained in the chapters presenting the experiment. Appendix E additionally provides the schematics of installation the experiment in the aircraft.

8.2.4 Measurements and Data Analysis

To evaluate the efficiency of hand-coordination yielded from the studied placement conditions (HA, PA, BA) and the effect of altered gravity (1*g*-PRE, 1.8*g*, 0*g*, 1*g*, 1*g*-POST) on the visuomotor performance and workload, different types of measures were used as dependent variables that are listed and specified in Table 8.2. Because the time to complete a test series was dictated by the time period of the hyper- and microgravity phase, the visuomotor performance was assessed by the frequency of correctly responded targets. Furthermore I present the percentage error rate, the pointing accuracy and assessed the performance by the response time that the participant needed to respond to a visual stimulus. For a uniform data collection I used a timer for the operator and one for the participant. For scheduling the experiment in flight the operator timed the procedure strictly by pilot's audio announcements of trajectory: *"Pull Up"* (increased Gz-load up to 1.8*g*) and "*Injection"* (rapid fall of Gz-load to ≈0 *g*). By inception of the hypergravity and the microgravity level, the operator manually started his timer and the first target was triggered. The timer for the participant was not started before the participant responded to the first stimulus. After 22 seconds or by operator's manual command, the session stopped. The time elapsed between the start times was considered as initial reaction time that the participant needed to orient towards the targets.

For workload assessment (see section 3.4.3) the subjective experience by the NASA RTLX was measured and analyzed. Thereby the subjects rated their workload for all three placement conditions (HA, PA, BA) after performing the 1*g*-PRE, the in-flight test and 1*g*-POST tests. Because of time limitations, additional differentiation between the gravitation levels in flight was not possible. Thus the rating for the in-flight tests includes 1.8*g*, 0*g* and 1*g* at once. During the second study the physiological workload was additionally assessed by cardiac response analyzing the heart rate variability (see section 3.4.3, *Physiological Workload*) with respect to the Stress Index (SI), which indicates the activity of the sympathetic regulation. I additionally performed a motion analysis related to the second experiment. Thereby the underlying goal-directed pointing movements were evaluated by the mislocalization of the pointing targets and by the path trajectories of the pointing movements, considering the two-component model introduced by Woodworth (1899).

Dependent Variable	Description
Visuomotor Performance	
Session time of operator [s]	Fix specified by 22 seconds. Timer was manually started by the operator. After 22 seconds the timer automatically stopped or was manual stopped by the operator.
Session time of participant [s]	Time elapsed between first target hit and the operator timer was terminated (< 22 seconds).
Initial reaction time [ms]	Time elapsed between the start of operator timer and the first target hit. Time that was initially needed to orient towards the interface.
Pointing frequency	Number of targets that was correctly responded within 22 seconds.
Error rate [%]	Percentage of number of targets that were falsely responded to the number of triggered targets within 22 seconds.
Pointing response time [ms]	Time elapsed between the visual stimulus of a target onset and motoric response onset.
Pointing accuracy [mm]	Euclidean distance between the center of the pointing target and the resulted pointing intersection within this target.
Subjective Workload	
Loading Factors	NASA RTLX rating scale (see Appendix A).
Physiological Workload by HRV	
Stress Index (SI)	Stress or strain index: $$SI = \frac{nD}{2D * (maxRR - minRR)}$$ with D as most frequently occurring value in the cardio interval; nD as frequency of D and $(maxRR - minRR)$ as variation range

Table 8.2: Performance and workload measures as dependent variables of the PF studies.

Data are presented by mean value plots that are marked with a confidence interval (CI) of 95%, as well as by distribution plots. Appendix F additionally provides the descriptive statistic for the outcomes showing measures of the central tendency and variability related to the first experiment, while Appendix G shows the same for the second parabolic flight study. In cases of skewed distributions the appendices also show the related box plots. To determine whether the visuomotor performance changed as a function of the independent variables (placement conditions, gravity load), multiple pairwise comparisons were conducted using SAS® PROC MIXED for repeatedly measured data and SAS® PROC GLIMMIX for objective count data (see section 3.4.2, *Data Analyzing),* whereby the significance level ($\alpha = 5\,\%$) was guaranteed by the Simulate method. To identify differences between the placement conditions, the performance indicators were compared at the same stage of *g*-level and across all *g*-levels, while they were compared between the in-flight *g*-levels (1.8*g*, 0*g*, 1*g*) at the same stage of placement condition to identify effects of altered gravity. To identify effects of Scopolamine and of the general flight condition, the performance was compared between the normogravity levels at the same stage of placement condition that aimed to verify that these conditions have not influenced ($p > .05$) the visuomotor performance. This was done between (1*g*-PRE, 1*g*, 1*g*-POST) for the first study and between (1*g*-PRE, 1*g*) for the second one. For analyzing main effects on the NASA RTLX rating scores between the placement conditions, a non-parametric test using SAS® PROC NPAR1WAY with the WILCOXON option was conducted.

As already described in section 3.4.2 (*Data Analyzing*), this induces a Wilcoxon rank sumtest in cases of two sample populations only (second study) and a Kruskal-Wallis rank test in cases of more than two random samples (first study), whereby the latter one only reports a main effect and required subsequent post-hoc testing that was done by multiple pairwise comparisons using the SAS® PROC GENMOD with Simulate adjustment. For analyzing the physiological workload, the Stress Index (SI) per interface placement condition across all *g*-levels is presented and discussed.

8.2.5 Confounding Factors

I identified eight confounding factors (see Table 8.3) that were not studied, but that could still have affected the experiment, and provided corresponding counteractions. For reconstruction subjects' performance and to identify further confounding factors, I utilized video sequences that were recorded during all in-flight sessions.

No.	Confounding factor	Counteraction
1	HMD calibration error.	Recalibration (in flight: only possible in breaks).
2	Unstable lighting conditions for optical tracking.	Adjustment of the lighting threshold value in operation time by the operator.
3	Incorrect fitting of the physical markers.	Observation by the operator and if necessary correction in operation time by the subject or operator.
4	Transition from the upwards 1.8*g* phase to the 0*g* phase.	The operator stopped manually the 1.8*g* phase with pilot's audio announcement *"40"* before the 0g phase was injected.
5	Unstable posture of the subject.	Assistant slightly fixed the subject.
6	Motion sickness.	Medication (Scopolamin).
7	Unstable flight conditions, like weather or pilot's performance.	Post-check of resulted parabola flight profiles provided by Novespace (for sample parabolas see Appendix E).
8	Fatigue effects:	-

Table 8.3: Confounding factors of the PF studies and the corresponding counteractions.

8.3 First Experiment (56th ESA Parabolic Flight Campaign)

The first experiment took place in May 2012 during the 56th ESA PFC (see Figure 8.8) in Bordeaux-Merignac, which was organized and funded by the European Space Agency. The experiment was approved with three flight days, for two participants and one operator per flight.

Figure 8.8: The participated teams of the 56th Parabolic Flight Campaign. [Photo © ESA, A. Le Floc'h, 2012]

8.3.1 Experiment Design and Procedure

In a within-subject design six participants performed the visuomotor task (see section 8.2.2) under three levels of interface placement (HA, PA, BA) and five levels of gravity load (1*g*-PRE, 1.8*g*, 0*g*, 1*g*, 1*g*-POST). The study was conducted during three flight days with 31 parabolas per flight. Thus ninety test series per *g*-level were expected (450 in total), whereby a test series constituted the successive performance of the visuomotor task for 22 seconds. During the in-flight procedure, the operator was responsible for the experiment coordination and the data collection unit, while the active participant performed the experiment task for 15 parabolas. In a counterbalanced measured design (see Figure 8.9) the active participant conducted the experiment task under one placement condition during five parabolas. With respect to the hazard report requirements the safety of the active participant had to be guaranteed by the inactive participant who served as assistant.

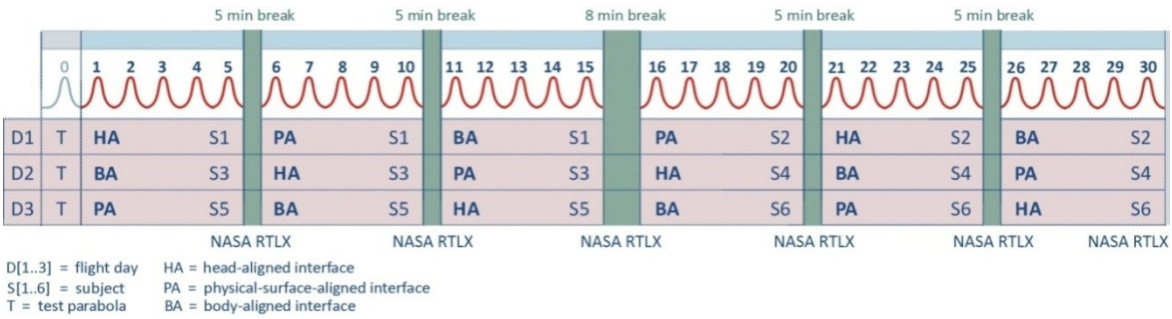


Figure 8.9: Study in-flight schedule with distributions of participants and conditions for the 56th PFC.

After the 16th parabola the participants changed their role. Because the experiment was performed for the first time, it was decided that the participant performs the task in a lying posture (see Figure 8.11) that was applied to all gravity loads. In microgravity the subject came up in a horizontal floating posture around 5 cm above the floor that was controlled by the assistant (see Figure 8.11, right).

8.3.2 Participants

The experiment was conducted by six subjects (1 female and 5 male) with an average age of 39.6 years (see Table 8.4). Four subjects had experienced parabolic flights previously and two subjects were novices. Four subjects wore glasses and all subjects were right-handed. Four subjects had a dominant right eye, while the remaining two subjects had a dominant left eye. All subjects received Scopolamine by injection as anti-motion-sickness medication. The subjects were precisely informed about the experiment protocol and signed an informed consent form.

ID	Gender	Age	Glasses	Dominant eye	Scopolamine (in mg)	PF experience (# parabolas)
S1	male	34	no	right	0.5	yes (550)
S2	female	45	no	right	0.5	yes (155)
S3	male	50	yes	left	0.7	yes (605)
S4	male	35	yes	left	1.0	no (0)
S5	male	44	no	right	1.0	no (0)
S6	male	30	no	right	1.0	yes (62)

Table 8.4: The inter-individual characteristics of the participants of the 56th PFC.

8.3.3 Flight Rack Schematic and Aircraft Setup

To fly the experiment aboard the Zero-G Airbus I built up a flight rack related to Novespace's guidelines. Inside the aircraft the rack was installed on the aircraft rails near the wall and longitudinal (X axis) located in the middle of the experimental area with a right transverse position (Y axis). For providing haptic cues for the PA interface, a swinging arm equipped with a board that could be vertically moved and rotated was designed (see Figure 8.10, left). Therewith the board could be individually positioned in front of the participant's head with sufficient space to perform pointing movements.

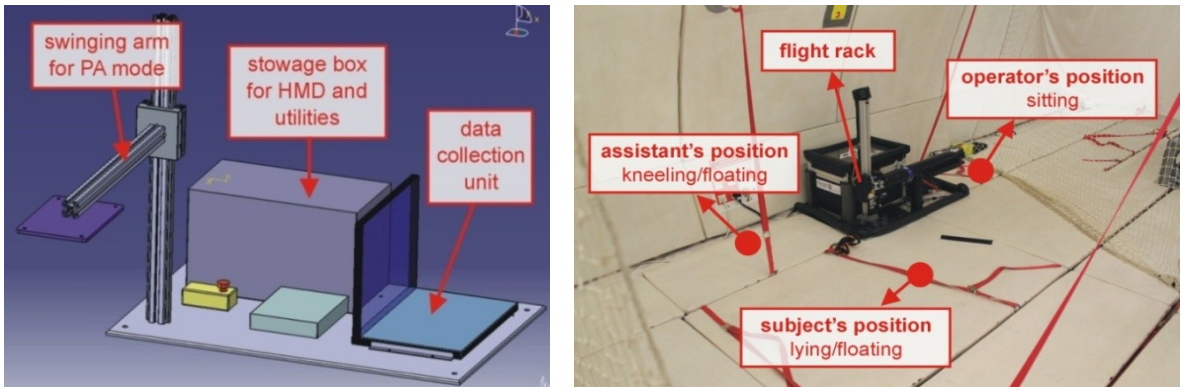


Figure 8.10: The experimental setup during the 56th PFC: *(left)* schematic of flight rack, and *(right)* final setup of the experimental area with labeled positions for the operator, the assistant and the active participant.

8.3.4 Results

The expected number of data sets was 450 data sets (270 in flight, 180 on ground in aircraft). Because one subject did not perform the 1*g*-PRE test, 435 (270 flight, 165 ground in aircraft) test series were finally evaluated. The visuomotor task could be successfully fulfilled under all placement conditions and *g*-levels. The operation software worked with a mean rate of 41 frame per second (SD = 11 fps). Because the experimental area inside of the Zero-G aircraft does not support daylight and provide stable lighting conditions, the optical marker-based tracking system worked very stable and did not require additional adjustment of the video lighting threshold value during operation. No participant suffered from motion sickness. For analyzing the data I compared the performance between the placement conditions across all *g*-levels and on same stage of *g*-level, as well as between the *g*-levels on same stage of interface placement separated for the in-flight levels (1.8*g*, 0*g*, 1*g*) and for the normogravity levels (1*g*-PRE, 1*g*, 1*g*-POST). Thereby timing effects, effects of pointing frequency and faulty pointing, as well as effects related to the pointing accuracy were considered.

Figure 8.11: Participants operating the interfaces in flight during the 56th PFC: *(left)* the bimanual handling of the BA interface, *(middle)* the world-referenced PA interface, and *(right)* task performance with PA under 0*g*.

8.3.4.1 Pointing Performance

Session Time of Operator and Participant

While the timer of a session was started manually by the operator and was running for 22 s or manually stopped, participant's timer was started after responding to the first triggered target. This resulted in different session times. Across all placement conditions operating the system resulted under 1*g* conditions (1*g*-PRE, 1*g*, 1*g*-POST) in a mean session time of 22.01 s (SD = 0.01), under 1.8*g* in a mean of 21.21 (SD = 1.44) and under 0*g* in a mean of 21.66 (SD = 1.55). In contrast, participants' performance across all placement conditions resulted in session times that differed between all *g*-levels (**1g-PRE:** M = 20.06 s, SD = 0.84; **1.8g:** M = 19.09 s, SD = 1.93; **0g:** M = 19.36 s, SD = 2.47; **1g:** M = 19.90 s, SD = 1.55; **1g-POST:** M = 20.03, SD = 1.01).

Initial Reaction Time

The difference of the session times reflected the reaction time that the participant needed to initially orient towards the interface. The distribution of the variances exhibited skewness, which was adjusted by natural logarithmic transformation. Figure 8.12 shows the mean variations of the log transformed reaction time for the placement conditions across all *g*-level and grouped by *g*-level. Visual inspection indicated that orientation towards the interface yielded longest reaction times under HA, while using the BA interface yielded shortest times for orientation under all *g*-levels, particularly

under 0*g*. Comparing the log transformed reaction time across all *g*-levels yielded significant differences between all placement conditions (see Table 8.5) that suggest that the reaction time needed to orient towards the interface significantly benefits from the BA condition. Comparing the placement conditions on same stage at *g*-level revealed significant faster orientation using the BA interface, compared to HA and PA (see Table 8.5). But, the comparison between the in-flight *g*-levels (1.8*g*, 0*g*, 1*g*) at the same stage of placement condition did not yield significant differences, suggesting that altered gravity did not affect visuomotor performance during the initial orientation phase. The same applies to comparing the 1*g* levels (1*g*-PRE, 1*g*, 1*g*-POST), which enabled to identify possible effects caused by Scopolamine and flight conditions. It also did not show significant differences, which suggests that these conditions did not affect the visuomotor performance during the initial orientation phase.

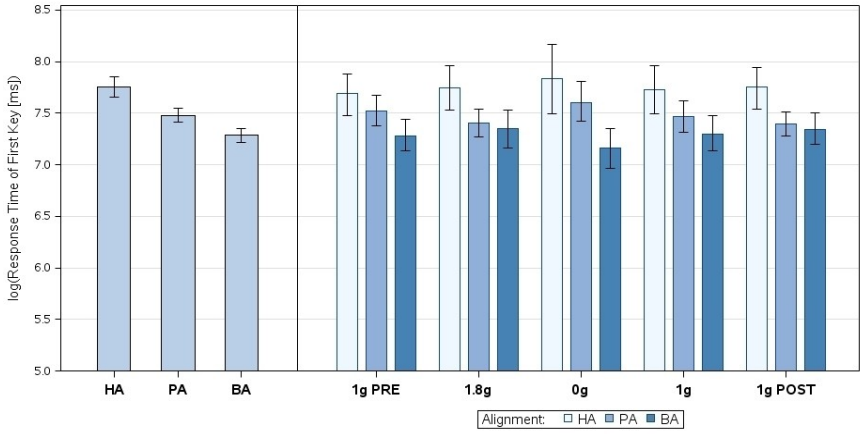


Figure 8.12: PFC56 - Mean log(reaction time first key) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*
Effect of placement across all g-levels							
HA	PA	-	-	434	4.75	<.0001	<.0001 *
BA	HA	-	-	434	-8.07	<.0001	<.0001 *
BA	PA	-	-	434	-3.33	0.0009	0.0027 *
Effect of placement at the same stage of g-levels							
BA	HA	0*g*	0*g*	422	-5.25	<.0001	<.0001 *
BA	PA	0*g*	0*g*	422	-3.48	0.0006	0.0405 *

Table 8.5: PFC56 - Significant differences of the log(reaction time first key) between the placements using SAS® PROC MIXED (adjust=simulate).

Frequency of Correct Target Hits

The frequency of correct targets hits is the number of targets that the participant had responded to in a correct way within 22 seconds. As clearly shown in Figure 8.13, participants pointed with the lowest frequency under the HA condition under all *g*-levels, while the means of PA and BA did not vary prominently. Most prominent variations can be seen under 0*g* with a mean frequency of 6.80 (SD = 3.60) for HA, compared to PA (M = 11.47, SD = 3.14) and BA (M = 12.80, SD = 3.65). The comparison of the placement conditions across all *g*-levels and per *g*-level (see Table 8.7) verified this observation by significant differences between HA and PA, as well as between HA and BA, except under 1*g* in flight.

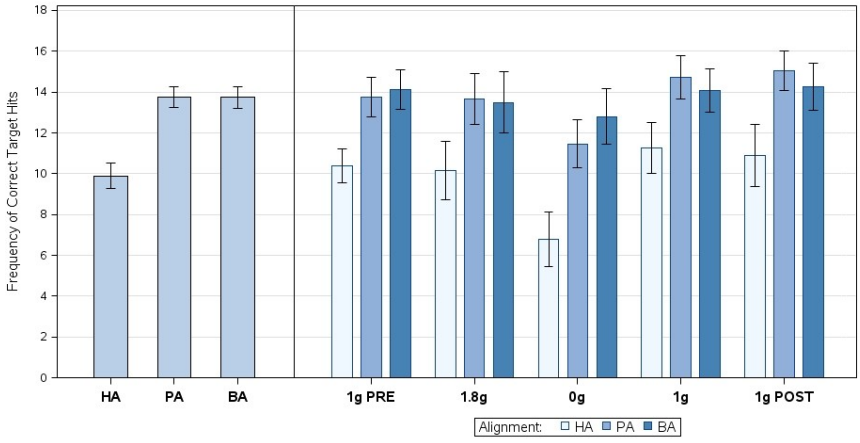


Figure 8.13: PFC56 - Mean frequency of correct target hits (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of altered gravity (1.8g, 0g, 1g) at the same stage of placement								
HA	HA	0*g*	1.8*g*	422	-4.45	<.0001	<.0001	*
HA	HA	0*g*	1*g*	422	-5.70	<.0001	<.0001	*
PA	PA	0*g*	1*g*	422	-3.49	0.0005	0.0053	*

Table 8.6: PFC56 - Significant differences of correct target hits between the *g*-levels on same stage of placement using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate).

Comparing the frequency of correct target hits between the in-flight *g*-levels revealed significant differences for HA between 0*g* and 1.8*g*, as well as between 0*g* and 1*g* (see Table 8.6). This suggests that altered gravity affected aimed pointing towards the HA interface. It also revealed that pointing towards the PA interface was affected under 0*g* conditions that significantly differed from 1*g* in-flight. Comparing the normogravity levels did not revealed any significant difference, i.e. Scopolamine and flight conditions did not influence the visuomotor performance with respect to the pointing frequency.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of placement across all g-levels								
HA	PA	-	-	422	-9.63	<.0001	<.0001	*
BA	HA	-	-	422	9.79	<.0001	<.0001	*
Effect of placement at the same stage of g-levels								
HA	PA	1g PRE	1g PRE	422	-3.41	0.0007	0.0023	*
BA	HA	1g PRE	1g PRE	422	3.74	0.0002	0.0008	*
HA	PA	1.8g	1.8g	422	-3.91	0.0001	0.0002	*
BA	HA	1.8g	1.8g	422	3.74	0.0002	0.0003	*
HA	PA	0g	0g	422	-5.91	<.0001	<.0001	*
BA	HA	0g	0g	422	7.30	<.0001	<.0001	*
HA	PA	1g	1g	422	-3.71	0.0002	0.0006	*
BA	HA	1g	1g	422	3.04	0.0025	0.0060	*
HA	PA	1g POST	1g POST	422	-4.46	<.0001	<.0001	*
BA	HA	1g POST	1g POST	422	3.68	0.0003	0.0008	*

Table 8.7: PFC56 - Significant differences of correct target hits between the placements using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate).

Percentage Error Rate

The percentage error rate is the ratio of the number of false responded targets to the number of triggered targets. The means of the log transformed error rate (see Figure 8.14) show prominent variations for HA by an increased error rate under all g-levels, but only showed significant differences across all *g*-levels compared to PA and BA (see Table 8.8). Comparing the error rate between the in-flight levels did not show effects of altered gravity. The same applies to comparing the normogravity levels, indicating that Scopolamine and flight conditions did not affect the pointing error rate.

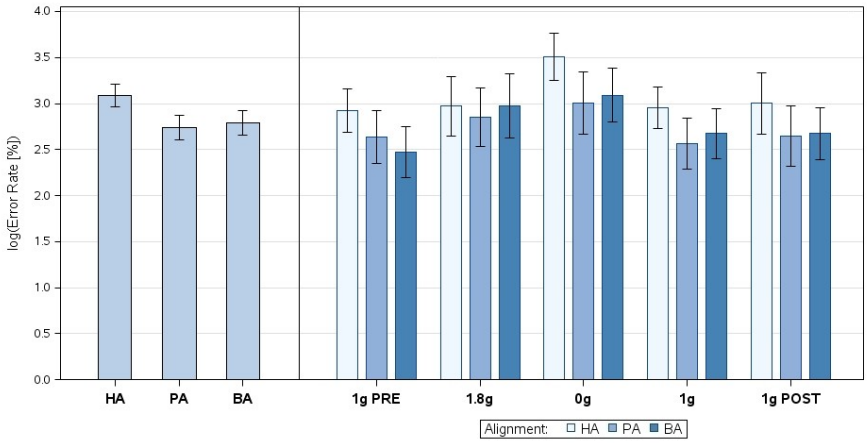


Figure 8.14: PFC56 - Mean log(error rate) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
HA	PA	-	-	354	3.77	0.0002	0.0006	*
BA	HA	-	-	354	-3.24	0.0013	0.0037	*

Table 8.8: PFC56 - Significant differences of the log(error rate) between the placements across all *g*-levels using SAS® PROC MIXED showing (adjust=simulate).

Response Time

For evaluation the response time as time that was needed for responding to a visual stimulus, I only considered data sets where the frequency of false targets hits was zero, and needed to adjust the data by natural logarithmic transformation because of skewed distributions. The variations of the mean response time (see Figure 8.15) show longest response times for the HA condition under all gravity loads, while the mean response times yielded from PA and BA did not specifically vary. Similar to the pointing frequency, the comparison of the placements across all *g*-levels and on same stage of *g*-level revealed significant differences (see Table 8.9) between HA and PA, as well as between HA and BA (except 1*g* in-flight). But comparing the response time between the in-flight and normogravity levels at the same stage of interface placement did not reveal significant differences indicating that there was no effect of altered gravity, Scopolamine and flight conditions on the time needed to respond.

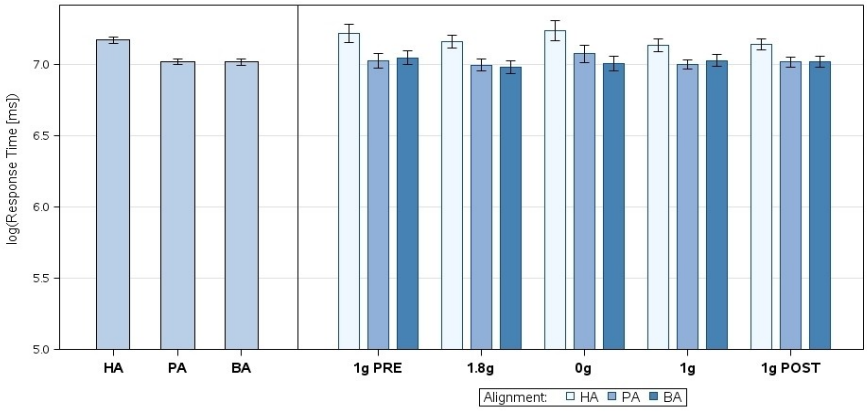


Figure 8.15: PFC56 - Mean log(response time) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of placement across all g-levels								
HA	PA	-	-	4913	9.57	<.0001	<.0001	*
BA	HA	-	-	4913	-9.86	<.0001	<.0001	*
Effect of placement at the same stage of g-levels								
HA	PA	1g PRE	1g PRE	4901	5.15	<.0001	<.0001	*
BA	HA	1g PRE	1g PRE	4901	-4.65	<.0001	0.0003	*
HA	PA	1.8g	1.8g	4901	4.79	<.0001	0.0002	*
BA	HA	1.8g	1.8g	4901	-5.27	<.0001	<.0001	*
HA	PA	0g	0g	4901	3.79	0.0002	0.0129	*
BA	HA	0g	0g	4901	-5.49	<.0001	<.0001	*
HA	PA	1g	1g	4901	4.26	<.0001	0.0021	*
HA	PA	1g POST	1g POST	4901	3.81	0.0001	0.0126	*
BA	HA	1g POST	1g POST	4901	-3.78	0.0002	0.0130	*

Table 8.9: PFC56 - Significant differences of the log(response time) between the placements using SAS® PROC MIXED (adjust=simulate).

Accuracy by Euclidean Distance

The accuracy is reflected by the Euclidean distance between target's centre and the final intersection within this target. For evaluation of the accuracy I only considered data sets where the frequency of false targets hits was zero. Visual inspection of the variations of the mean distances (see Figure 8.16) revealed increased mean distance under the PA condition and decreased mean distance under the HA condition across all *g*-levels that also showed significant differences (see Table 8.10, top).

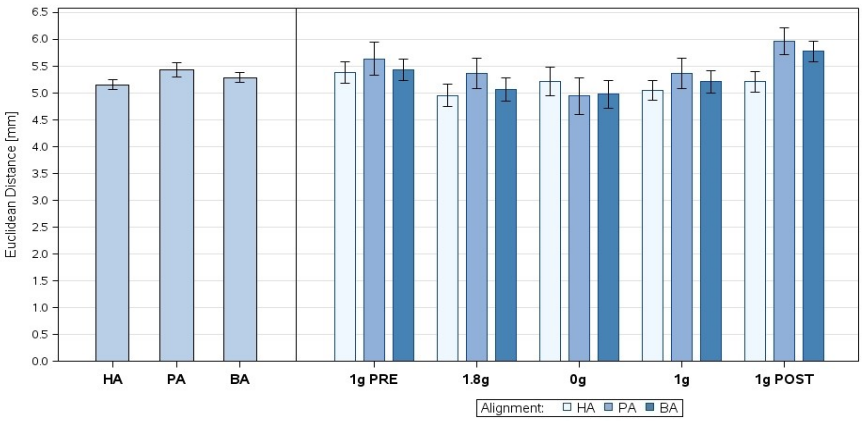


Figure 8.16: PFC56 - Mean Euclidean distance (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

The comparison of the placements on same stage of gravity load only revealed differences on the 1*g*-POST level between HA and PA, as well as HA and BA with decreased distances under the HA condition. Comparing the distances for the in-flight levels (1.8*g*, 0*g*, 1*g*) at the same stage of placement did not show effects of altered gravity on the accuracy. Comparing the distances for the normogravity levels (1*g*-PRE, 1*g*, 1*g*-POST) also did not yield effects of Scopolamine (1*g*-PRE, 1*g*), but showed an effect of the flight condition (1*g*, 1*g*-POST) for BA with a less accurarcy in 1*g*-POST.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*
Effect of placement across all g-levels							
HA	PA	-	-	2746	-3.45	0.0006	0.0015 *
Effect of placement at the same stage of g-levels							
HA	PA	1g POST	1g POST	2734	-4.01	<.0001	0.0050 *
BA	HA	1g POST	1g POST	2734	3.68	0.0002	0.0189 *
Effect of Scopolamine or flight condition (1g-PRE, 1g, 1g-POST) at the same stage of placement							
BA	BA	1g	1g POST	1692	-3.87	0.0001	0.0035 *

Table 8.10: PFC56 - Significant differences of the Euclidean distance between the placements for same *g*-level and between the *g*-levels on same stage of placement using SAS® PROC MIXED (adjust=simulate).

8.3.4.2 Subjective Workload

The participants rated their experienced workload by the NASA RTLX after changing the placement conditions. Because the interface placement changed after five parabolas during the in-flight test, the participants rated their workload afterwards across all related *g*-levels (1.8*g*, 0*g*, 1*g*). Therefore I compared the rating scores for the placement conditions (HA, PA, BA) related to the pre-test (1*g*-PRE), the post-test (1*g*-POST) and summarized for all in-flight tests. Figure 8.17 shows the mean ratings scores across all participants per placement condition and *g*-level, as well as paneled by the rating factors. The loading factors were in general rated lower for the PA interface under all *g*-levels indicating decreased workload, while the factors were highest rated for the HA interface, except for the *"Physical Demand"* that was similar to the ratings for the BA condition.

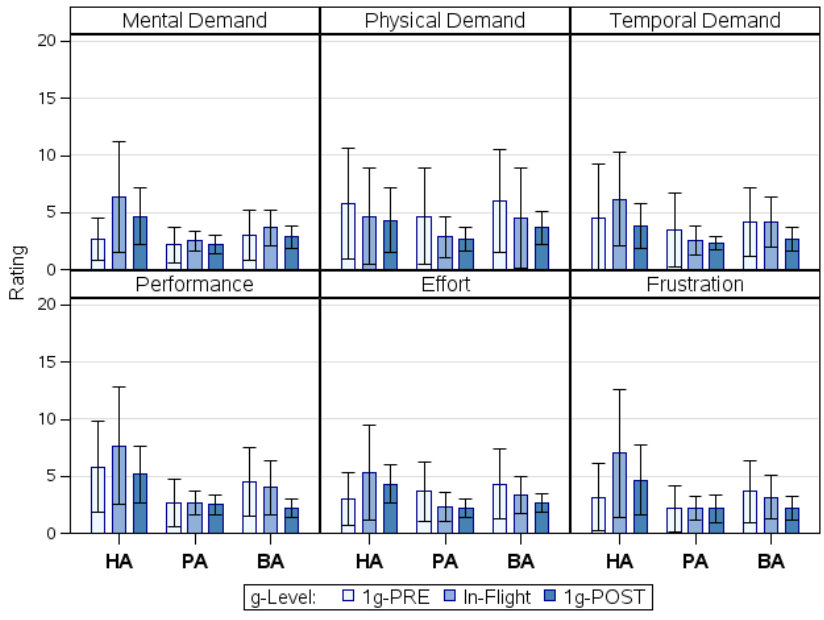


Figure 8.17: PFC56 - Mean ratings (with CI=95%) of the NASA RTLX scale grouped by the placements and *g*-levels, and paneled by ratings items.

Analyzing the Wilcoxon scores by the Kruskal-Wallis test for the 1*g*-PRE level (see Figure 8.18, left) did not reveal a significant difference, $H(DF=2, N=36) = 3.68$, $p = .1592$, between the rating scores of the placement conditions with a mean rank of 56.19 for HA, 46.79 for PA and 60.51 for BA. But the scores for the in-flight levels (1.8*g*, 0*g*, 1*g*) (see Figure 8.18, middle) yielded a significant difference, $H(DF=2, N=36) = 22.46$, $p < .0001$, with a mean rank of 71.76 for HA, 37.38 for PA and 54.36 for BA. Also the scores for 1*g*-POST level (see Figure 8.18, right) yielded a significant difference, $H(DF=2, N=36) = 21.98$, $p < .0001$, with a mean rank of 73.19 for HA, 40.90 for PA and 49.40 for BA. Post-hoc analysis (see Table 8.11) revealed significant differences under the in-flight levels (1.8*g*, 0*g*, 1*g*) between all placement conditions, and under the 1*g*-POST level between BA and HA, as well as HA and PA. The discrepancy between 1*g*-PRE and the in-flight levels indicates that something has affected the self-reported workload, but it is not clear whether it was caused by altered gravity, Scopolamine, or flight conditions.

g-level	Display		*z Value*	*Pr > |z|*	*Adj. p-Val*	
1*g*-PRE	BA	HA	-0.61	0.5407	0.8147	
	BA	PA	-1.87	0.0616	0.1438	
	HA	PA	-1.26	0.2084	0.4133	
In-flight (1.8*g*, 0*g*, 1*g*)	BA	HA	2.78	0.0055	0.0159	*
	BA	PA	-2.41	0.0160	0.0386	*
	HA	PA	-4.82	<.0001	<.0001	*
1*g*-POST	BA	HA	3.75	0.0002	0.0003	*
	BA	PA	-1.25	0.2128	0.4294	
	HA	PA	-4.72	<.0001	<.0001	*

Table 8.11: PFC56 - Comparison of the NASA RTLX rating scores between the placements per *g*-level using SAS® PROC GENMOD (dist=multinomial, link=cumlogit, adjust=simulate).

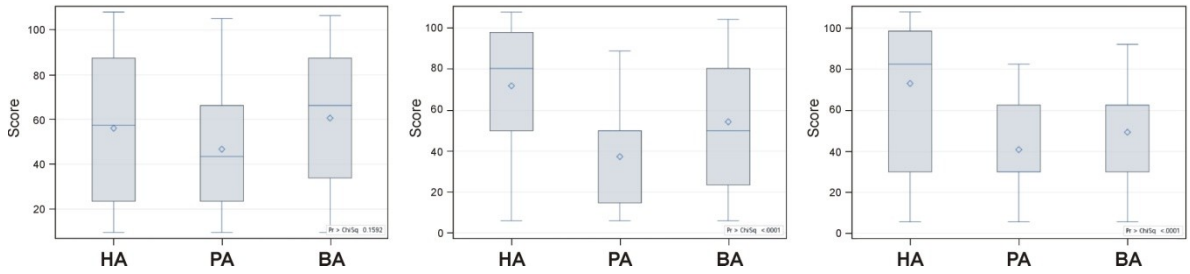


Figure 8.18: PFC56 - Distributions of the Wilcoxon scores across all ratings items of the NASA RTLX by the placements for the 1*g*-PRE level *(left)*, for the in-flight levels (1.8*g*, 0*g*, 1*g*) *(middle)* and for the 1*g*-POST level *(right)* yielded from SAS® PROC NPAR1WAY.

8.3.5 Discussion

The results, presented above, indicate that the visuomotor performance and self-reported workload during direct AR selection are affected by the interface placement, considering interfaces that are referenced to the world, hand or the head, each varying in its specification of location coding, handling mode and the presence of haptic cues, supporting tactile sensation. The most important finding revealed that the underlying visuomotor coordination benefited from haptic cues, regardless of the gravity load. It was proven that the performance was significantly deteriorated by using the head-referenced interface, not providing haptic cues, especially under microgravity conditions. These findings were related to all time-based metrics, the error rate and the number of correct target hits,

while the same interface condition surprisingly supported the most accurate pointing, but differed significantly only from the world-referenced interface across all gravitational levels. Considering Woodworth's two-component model (1899), this finding suggests that the correction phase of the pointing movement was prolonged, which consequently shortens the preceding ballistic phase. Because this phase is responsible to bring the finger fast in the vicinity of the target, its shortening may explain why the head-referenced interface deteriorated as a function of time. Nonetheless, due to the increased amount of impairments, the head reference is considered as not fully operational. However, comparing the pointing performance between the world-referenced and handheld interface, both providing haptic cues, only revealed time differences under microgravity related to the initial orientation phase in favor of the hand reference, although it implied bimanual handling, but not requiring transformation between intrinsic and extrinsic coordinates. Besides differences between the placement references, comparing the performance between the in-flight levels (1.8*g*, 0*g*, 1*g*) at the same level of placement made it possible to investigate possible effects of altered gravity. Such modified conditions of gravity only affected the pointing frequency revealed from the world and head references, while the performance by the hand reference was never affected by gravity changes. To ensure that neither the intake of Scopolamine nor the flight condition affected the visuomotor coordination, the performance at the same level of placement condition was additionally compared between the normogravity conditions (1*g*-PRE, 1*g*, 1*g*-POST). Such phenomena could really not be observed. It was only shown that pointing towards the hand-referenced interface caused significantly less precise pointing under the post-test on ground compared to the 1*g* condition in flight. But it is much more likely that this impairment resulted from the exhaustion and fatigue caused by the strenuous performance in flight and the bimanual handling, requiring more physical effort. In contrast to indicators of performance bias, the mental workload was assessed in a subjective way, analyzing the rating scores of the NASA RTLX. Related data could be assessed separately under 1*g*-PRE and 1*g*-POST. This was not possible under the flight condition because of physical restrictions. Thus the workload under the in-flight levels (1.8*g*, 0*g*, 1*g*) was rated in a summarized way. While the rating scores for the 1*g*-PRE test did not revealed differences between the placement conditions, it did for the in-flight test. Thereby the world-referenced interface led to the significantly lowest self-reported workload, while the head reference caused the significantly highest workload. Similar was shown for the post-test, but not revealing differences between the world and hand references. The resulting discrepancy between 1*g*-PRE and the in-flight levels suggests that either altered gravity, Scopolamine, or the flight condition affected the self-reported workload.

Overall, with this experiment I provided evidence that haptic cues are strongly suggested for the visuomotor coordination requested by AR interfaces that are operated by direct touch selection. Although a head-referenced interface provides inside-coded pointing targets and is needed to be handled by only one hand, the lack of tactile sensation seems to clearly preponderate, especially in microgravity, but also under normogravity conditions, which confirms the findings by Lindeman et al. (1999). Thus, the research question Q-1.3 (see section 8.1) can be answered with "Yes", confirming that the absence of haptic cues deteriorates visuomotor coordination, not only in performance, but also caused the highest self-reported workload, which in turn confirms the normogravity findings revealed from a comfort questionnaire by Bowman and Wingrave (2001), showing, amongst others, that pointing towards head-referenced floating targets causes a large amount of arm strain. Although the performance using the body-referenced interfaces constitutes an increase in physical effort by the bimanual handling mode, the feature to perform with haptic cues seems to have a greater emphasis. It could also be shown that the same reference type provided the fastest initial orientation towards the interface and was not affected by altering the gravity.

Nonetheless, the findings are not sufficient to prove distinct differences between the interfaces, providing tactile sensation and thus it is not clear whether the location coding or the handling mode contributed to more efficient visuomotor performance, although the self-reported workload benefited most from unimanual pointing towards the outside coded interfaces, provided by a world reference. Hence, the other research questions, asked in section 8.1 under Q-1, cannot be answered yet. But upon visual inspection of the performance data, it seems that the inside-coded handheld interface has the potential to outperform the outside-coded world-referenced one, as was hypothesized. To verify this assumption, a follow-up experiment was conducted that is presented in the following section. Besides answering the remaining questions of Q-1, the second experiment has also been intended to complete answering the research questions related to Q-2. While the first experiment already provided evidence that the visuomotor coordination revealed from the world and the head reference is affected by microgravity, as well as also under hypergravity from the head reference, the second experiment was also intended to collect motion data, enabling the analysis of the underlying goal-directed pointing movements by, for example, Woodworth's two-component model.

8.4 Second Experiment (58th ESA Parabolic Flight Campaign)

The second experiment took place in May/June 2013 during the 58th ESA PFC (see Figure 8.19) in Bordeaux-Merignac, which was organized and funded by the European Space Agency. This experiment was also approved with three flight days and six participants. In the first experiment I could show that haptic cues are required for direct AR selection under hyper- and microgravity conditions. Therefore the placement conditions were reduced by eliminating the head-referenced interfaces, which did not support tactile sensation. Additionally I decided to exclude the post-test (1*g*-POST), because participants were extremely tired and exhausted from the strenuous performance of the flight tests. Thus, the second experiment was intended to evaluate differences in performance and workload between the visuomotor coordination revealed from goal-directed pointing towards the world- and the hand-referenced interface.

Figure 8.19: The participated teams of the 58th Parabolic Flight Campaign. [Photo © ESA, A. Le Floc'h, 2013]

8.4.1 Experiment Design and Procedure

Corresponding to the elimination of the HA reference and the 1g-POST level, the resulted within-subject design only consisted of two levels of interface placement (PA, BA) and four levels of gravity load (1*g*-PRE, 1.8*g*, 0*g*, and 1g). Because of the reduced level for the placement condition, a procedure design with a repetition rate of seven was applied, which means that seven parabolas were used for one subject under one placement condition (see Figure 8.20). Therewith the study was conducted during three flight days with 31 parabolas per flight, but only using 28 parabolas for data collection. Thus, 84 test series per *g*-level (336 in total) were expected, whereby one test series constituted the successive performance of the visuomotor task (see section 8.2.2) for 22 seconds, as was already done during first experimentation. This time the experiment was approved to fly with four attendees per flight, including two participants, one operator and one assistant, who should ensure participant's safety in flight. The same assistant flew on all three days and was very experienced under parabolic flight conditions (> 900 parabolas), enabling a uniform assistance. Being more experienced under those conditions, this time the participant performed the experiment task in an upright posture. This implied that in microgravity the participant came up in an upright floating posture, but was slightly fixed by the assistant. Because three participants (see section 8.4.2) were experienced in parabolic flights, they performed the test series in flight in the first part. Thus the unexperienced participants could spend this inactive time to become familiar with the gravitational condition before conducting the experiment, while the active participant conducted the experiment task under one placement condition during seven parabolas in a counterbalanced measured design.

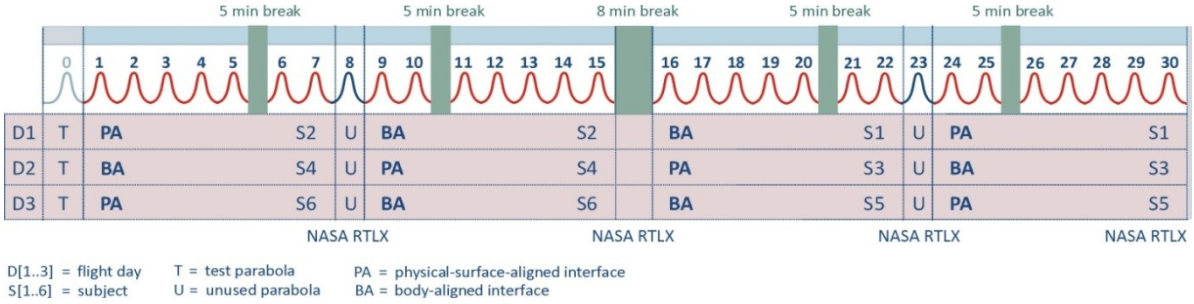


Figure 8.20: Study in-flight schedule with distributions of participants and conditions for the 58th PFC.

8.4.2 Participants

The experiment was performed by six participants (1 female and 5 male) from 26 to 51 years (average 32.2). Table 8.12 presents participants' inter-individual characteristics. Three participants had experi-

ID	Gender	Age	Height (in cm)	Dominant eye	Scopolamine (in mg)	PF experience (# parabolas)
S1	female	26	171	left	0.8	no (0)
S2	male	51	185	left	0.5	yes (900)
S3	male	29	171	right	0.7	no (0)
S4	male	24	172	right	0.4	yes (93)
S5	male	32	180	left	0.9	no (0)
S6	male	31	188	right	0.7	yes (124)

Table 8.12: The inter-individual characteristics of the participants of the 58th PFC.

ences in parabolic flights previously and already took part in the first campaign. Three subjects were novice in parabolic flight. All subjects were right-handed. Three subjects had a dominant right eye, while the other three subjects had a dominant left eye. Three participants had experience working with AR and three subjects were novices in this field. They came from the fields aerospace, biology and media. All subjects received Scopolamine by injection as anti-motion-sickness medication. The subjects were precisely informed about the experiment protocol and signed an informed consent form.

8.4.3 Flight Rack Schematic and Aircraft Setup

To fly the experiment aboard the Zero-G Airbus I used the flight rack from first experiment during the 56th PFC and built up a second rack (see Figure 8.21, left) to provide haptic cues for the PA interface, so that the participant could operate in an upright posture. The PA panel was installed at the handrail of the aircraft in front of the main flight rack. Inside the aircraft the rack was installed on the aircraft rails near the wall and longitudinal (X axis) located in rearmost part of the experimental area with a right transverse position (Y axis). While the active participant always performed in front of the PA panel, the inactive subject lay next to the operator.

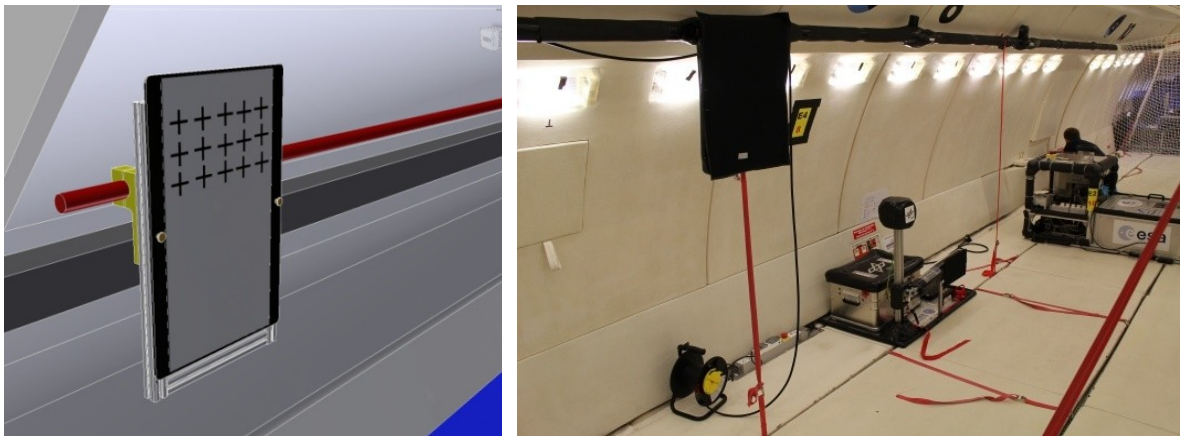


Figure 8.21: The experimental setup during the 58th PFC: *(left)* schematic of the PA panel installed at the aircraft's handrail, and *(right)* final setup of the experimental area.

8.4.4 Results

Similar to the first experiment, the visuomotor task could be successfully fulfilled under all placement conditions and *g*-levels (see Figure 8.22) and no participant suffered from motion sickness. The operation software worked with a mean rate of 41 frame per second (SD = 12 fps). Therewith the expected number of data sets of 336 (252 in flight, 84 on ground in aircraft) could be evaluated. The data were inspected and analyzed in same way as for the first experiment by comparing the performance between the placement conditions across all *g*-levels and on same stage of *g*-level, as well as between the *g*-levels on same stage of placement separated for the in-flight levels (1.8*g*, 0*g*, 1*g*) and for the normogravity levels (1*g*-PRE, 1*g*). Besides evaluating the workload by self-reported experiences, I additionally evaluated the stress index assessed by the heart rate variability (HRV).

Figure 8.22: Participants operating the interfaces under 0*g* during the 58th PFC: *(left)* task performance with the world-referenced PA interface, and *(right)* task performance with the body-referenced bimanual BA interface.

8.4.4.1 Pointing Performances

Initial Reaction Time

The variances of the reaction time as time initially needed for orienting towards the interface exhibited skewed distribution which was adjusted by natural logarithmic transformation. Figure 8.23 shows the mean variations of the log transformed reaction time per placement condition and *g*-levels. Except less than 1*g*-PRE, orientation towards the interface yielded the longest reaction times under the PA condition, particularly under 0*g*. Comparing the log transformed reaction time across all *g*-levels yielded significant differences (see Table 8.13) between BA and PA with faster reaction time for BA. Also, comparing the placement conditions at same stage of *g*-level revealed significant faster orientation using the BA interface under 0*g* and 1*g*. The comparison of the log transformed reaction time between the in-flight *g*-levels (1.8*g*, 0*g*, 1*g*) at the same stage of placement condition yielded significant differences for PA indicating that initial orienting is timely affected by microgravity (0*g*, 1*g*), but not by hypergravity (1*g*, 1.8*g*), and yielded differences for BA that were also affected by microgravity (0*g*, 1*g*). Comparing both normogravity levels (1*g*-PRE, 1*g*) showed significant differences for PA that suggests that Scopolamine or flight conditions affected the visuomotor performance during the initial orientation phase under this placement condition.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of placement across all g-levels								
BA	PA	-	-	328	-7.13	<.0001	<.0001	*
Effect of placement at the same stage of g-levels								
BA	PA	0g	0g-	328	-6.95	<.0001	<.0001	*
BA	PA	1g	1g	328	-4.85	<.0001	<.0001	*
Effect of altered gravity (1.8g, 0g, 1g) at the same stage of placement								
PA	PA	0g	1.8g	328	7.08	<.0001	<.0001	*
PA	PA	0g	1g	328	5.19	<.0001	<.0001	*
BA	BA	0g	1g	328	3.09	0.0022	0.0444	*
Effect of Scopolamine or flight condition (1g-PRE, 1g) at the same stage of placement								
PA	PA	1g	1g PRE	328	3.76	0.0002	0.0050	*

Table 8.13: PFC58 - Significant differences of the log(response time to first key) between the placements and *g*-levels using SAS® PROC MIXED (adjust=simulate).

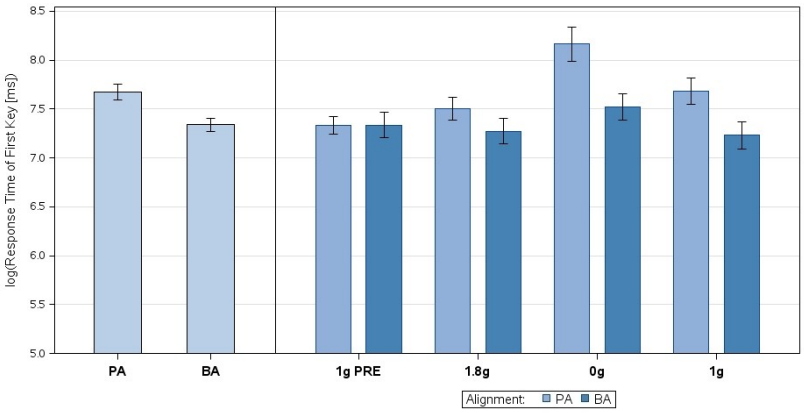


Figure 8.23: PFC58 - Mean log(response time first key) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Frequency of Correct Target Hits

As can be seen in Figure 8.24 and Table 8.14, the variations of the mean pointing frequency showed a significantly decrease for the PA interface across all *g*-levels. While during the 1g-PRE test the pointing frequency did not differ significantly between the conditions, under the flight conditions participants pointed with the lowest frequency using the PA condition, with most prominent variation under microgravity. Comparing the pointing frequency between the in-flight levels at the same stage of placement yielded that PA and BA were affected by hyper- and microgravity. Effects of Scopolamine and flight conditions on pointing frequency were not observed by comparing the pointing frequency under normogravity conditions (1*g*-PRE, 1*g*).

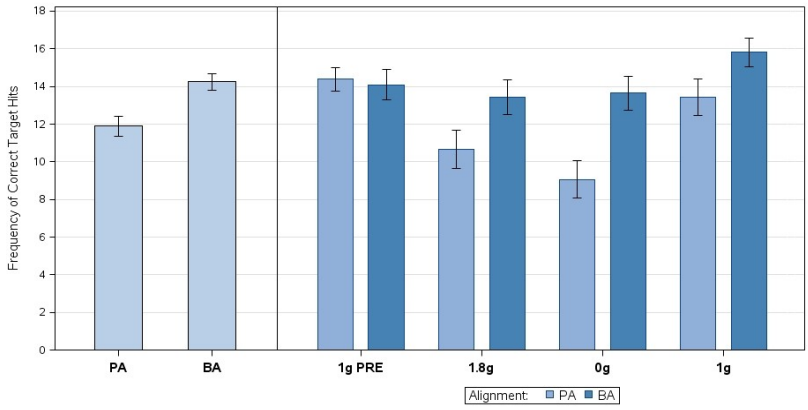


Figure 8.24: PFC58 - Mean frequency of correct target hits (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of placement across all g-levels								
BA	PA	-	-	328	6.38	<.0001	<.0001	*
Effect of placement at the same stage of g-levels								
BA	PA	1.8g	1.8g	328	3.64	0.0003	0.0002	*
BA	PA	0g	0g	328	6.17	<.0001	<.0001	*
BA	PA	1g	1g	328	2.85	0.0046	0.0047	*
Effect of altered gravity (1.8g, 0g, 1g) at the same stage of placement								
BA	BA	1.8g	1g	328	-2.85	0.0046	0.0267	*
BA	BA	0g	1g	328	-2.59	0.0102	0.0520	*
PA	PA	0g	1g	328	-5.91	<.0001	<.0001	*
PA	PA	1.8g	1g	328	-3.64	0.0003	0.0022	*

Table 8.14: PFC58 - Significant differences of correct target hits between the placements and *g*-levels using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate).

Percentage Error Rate

The variations of the mean log transformed error rate are presented in Figure 8.25, while Table 8.15 shows the related significant differences. Compared to BA it shows a significantly increased error rate for PA across all g-levels, as well as under hyper- and microgravity. Comparing the in-flight levels (1.8*g*, 0*g*, 1*g*) for the same placement showed an increase in error rate for PA affected by hyper- and microgravity, as well as for BA that was affected by microgravity. Comparing the normogravity levels (1*g*-PRE, 1*g*) did not revealed effects of Scopolamine and flight conditions.

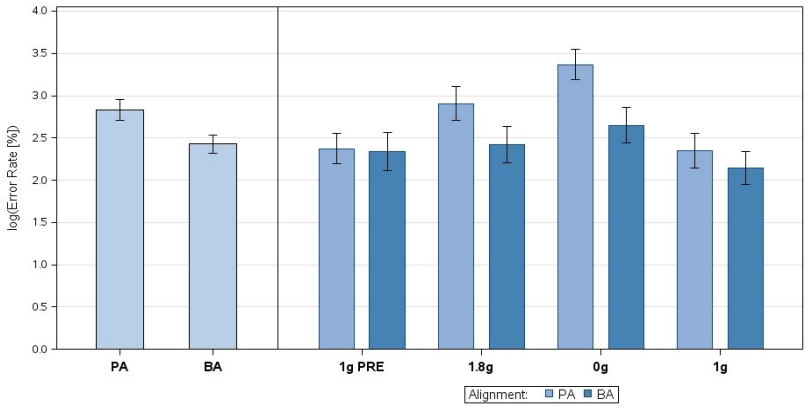


Figure 8.25: PFC58 - Mean log(error rate) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Effect of placement across all g-levels								
BA	PA	-	-	233	-4.85	<.0001	<.0001	*
Effect of placement on same stage of g-levels								
BA	PA	1.8g	1.8g	233	-3.48	0.0006	0.0129	*
BA	PA	0g	0g	233	-5.68	<.0001	<.0001	*
Effect of altered gravity (1.8g, 0g, 1g) on same stage of placement								
PA	PA	1.8g	1g	233	4.17	<.0001	0.0011	*
PA	PA	0g	1.8g	233	3.74	0.0002	0.0054	*
PA	PA	0g	1g	233	7.58	<.0001	<.0001	*
BA	BA	0g	1g	233	3.40	0.0008	0.0172	*

Table 8.15: PFC58 - Significant differences of log(error rate) between the placements and *g*-levels using SAS® PROC MIXED (adjust=simulate).

Response Time

For evaluation of the response time I only considered data sets where the number of false targets hits was zero. The mean variations are presented in Figure 8.26 and the related significant differences in Table 8.16. Across all *g*-levels the participant pointed significantly faster with BA that also provided significantly faster pointing under altered gravity conditions (1.8*g*, 0*g*). The response time yielded from both conditions was not affect by hypergravity (1.8*g*, 1*g*) and microgravity (0*g*, 1*g*) altered gravity conditions. But the performance under the BA condition was affected by Scopolamine or flight conditions (1*g*-PRE, 1*g*).

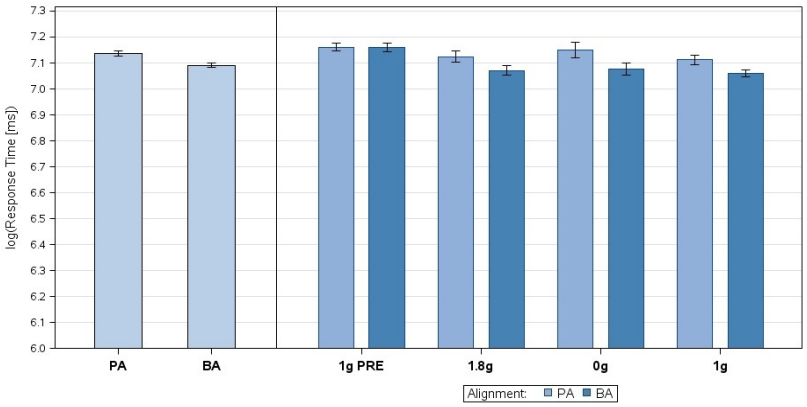


Figure 8.26: PFC58 - Mean log(response time) (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*
Effect of placement across all g-levels							
BA	PA	-	-	3870	-5.21	<.0001	<.0001 *
Effect of placement at the same stage of g-levels							
BA	PA	1.8g	1.8g	3870	-3.14	0.0017	0.0359 *
BA	PA	0g	0g	3870	-3.52	0.0004	0.0104 *
BA	PA	1g	1g	3870	-3.63	0.0003	0.0073 *
Effect of Scopolamine or flight condition (1g-PRE, 1g) at the same stage of placement							
BA	BA	1g	1g PRE	3870	-6.72	<.0001	<.0001 *

Table 8.16: PFC58 - Significant differences of the log(response time) between the placements and *g*-levels using SAS® PROC MIXED (adjust=simulate).

Accuracy by Euclidean Distance

For pointing accuracy I assessed the Euclidean distance between target's center and the final intersection within this target, and for evaluation I only considered data sets where the frequency of false targets hits was zero. The variations of the mean distances (see Figure 8.27) showed an increase in distance under *g*-levels. Comparing the distance (see Table 8.17) between the placement conditions revealed significant differences under all *g*-levels, except the under 1*g*-PRE, while the comparison between the in-flight gravity loads at the same stage of placement conditions only showed an effect of microgravity for BA. But the pointing accuracy resulted from both interfaces was affected by Scopolamine or flight conditions, as revealed from comparing the normogravity levels (1*g*-PRE, 1*g*).

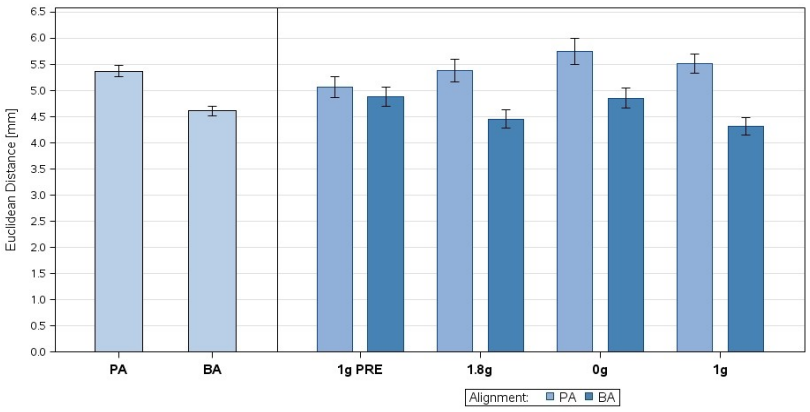


Figure 8.27: PFC58 - Mean Euclidean distance (with CI=95%) of the placements: *(left)* across all *g*-levels, and *(right)* grouped by *g*-level.

Placement		g-level		DF	t Value	Pr > |t|	Adj. p-Val	
Effect of placement across all g-levels								
BA	PA	-	-	3870	-11.03	<.0001	<.0001	*
Effect of placement at the same stage of g-levels								
BA	PA	1.8g	1.8g	3870	-6.15	<.0001	<.0001	*
BA	PA	0g	0g	3870	-5.38	<.0001	<.0001	*
BA	PA	1g	1g	3870	-9.24	<.0001	<.0001	*
Effect of altered gravity (1.8g, 0g, 1g) at the same stage of placement								
BA	BA	0g	1g	3870	4.09	<.0001	0.0014	*
Effect of Scopolamine or flight condition (1g-PRE, 1g) at the same stage of placement								
PA	PA	1g	1g PRE	3870	3.39	0.0007	0.0165	*
BA	BA	1g	1g PRE	3870	-4.46	<.0001	<.0001	*

Table 8.17: PFC58 - Significant differences of the Euclidean distance between the placements and *g*-levels using SAS® PROC MIXED (adjust=simulate).

8.4.4.2 Subjective Workload

Just as in the first experiment, the participants rated their experienced workload by the NASA RTLX after changing the placement conditions. Because the interface placement changed after seven parabolas during the in-flight test, the participants rated their workload afterwards across all related *g*-levels (1.8*g*, 0*g*, 1*g*). Therefore I compared the ratings between the placement conditions (PA, BA) for the pre-test (1*g*-PRE) and summarized for the in-flight tests. Figure 8.28 (left) shows the mean rating scores paneled by the rating factors across all participants per placement and *g*-level. The loading factors were generally rated lower for the pre-test, suggesting that the flight conditions caused an increased workload. Visual inspection related to the studied interfaces indicated that all loading factors were rated highest for PA during the in-flight tests, while the ratings during the pre-test did not clearly vary between PA and BA. For analyzing the rating scores of the NASA RTLX, a nonparametric Wilcoxon rank sum test was performed. Figure 8.28 (right) shows the distributions of the Wilcoxon score values across all loading factors for the 1*g*-PRE and the in-flight condition. The test performed for the 1*g*-PRE condition showed a higher mean score for PA ($M = 37.13$) than for BA

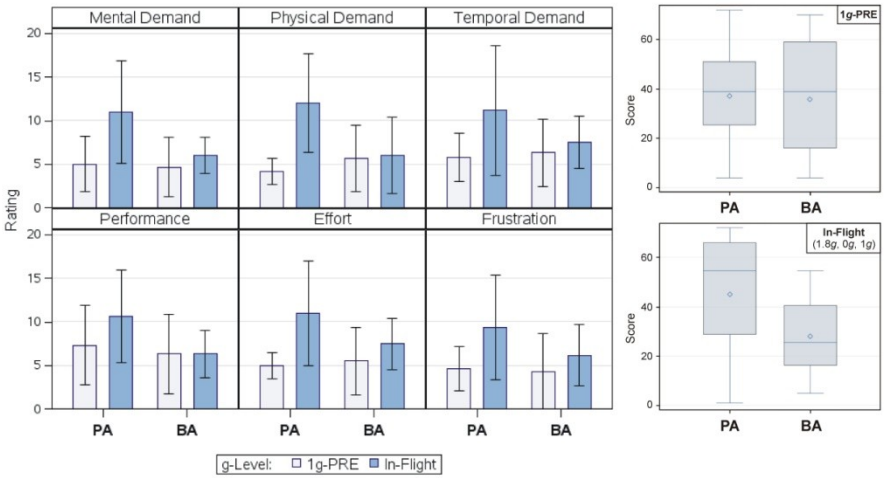


Figure 8.28: PFC58 - Subjective Workload: *(left)* Mean ratings (with CI=95%) of the NASA RTLX grouped by the placements and *g*-levels, and paneled by the loading factors. *(right)* Distribution of the Wilcoxon scores of the placements across all loading factors for 1*g*-PRE and the in-flight levels.

($M = 35.87$), but did not indicate significant differences, $Z = 0.2503$, $p = 0.4012$, $r = 0.0295$. The same test performed for the in-flight condition indicated that the workload was significantly more increased with PA ($M = 44.99$) than with BA ($M = 28.01$), $Z = 3.4458$, $p = 0.0003$, $r = 0.4061$.

8.4.4.3 Physiological Workload

For the physiological workload I evaluated the heart rate variability (HRV) that was assessed in the upright posture during the experiment and will be reported by the Stress Index (SI) revealed from the in-flight performance. The SI describes the equilibrium within the autonomic nervous system (ANS) which consists of the sympathetic (exciting) and the parasympathetic (calming) part. Therefore, it can be applied to measure the physiological and cognitive workload during a defined period of time. As presented in Table 8.18, the data showed that the physiological and cognitive workload was at the lowest point for the participants during the inactive phase of the experiment, except one subject (S5), who was unexperienced in parabolic flight and suffered a panic attack initially, but calmed down before starting the experiment after the 15th parabola. Apart from that, the results suggest that the workload was higher during the application of the BA mode for five subjects, while one subject was more strained using the PA method that could have been caused by the inter-individual characteristics.

ID	Gender	Height (cm)	Weight (kg)	BMI (kg/m^2)	PF	**Stress Index** (SI)		
						inactive	**PA**	**BA**
S1	female	171	90	30.78	no	81.55	117.47	104.34
S2	male	185	82	23.96	yes	35.28	40.19	54.75
S3	male	171	78	26.68	no	30.29	35.33	36.30
S4	male	172	69	23.32	yes	78.22	155.13	174.76
S5	male	180	107	33.03	no	332.23	170.56	196.82
S6	male	188	80	22.64	yes	65.93	95.96	110.10

Table 8.18: PFC58 - Participants' inter-individual characteristics and Stress Indexes for the inactive, the PA and the BA phase during the in-flight condition. (PF = experienced in parabolic flight)

8.4.5 Discussion

This follow-up study was intended to supplement the findings of the first experiment (see section 8.3), investigating the effect of altered gravity on visuomotor coordination revealed from world-, hand- and head-referenced interfaces. The initial experiment has already provided evidence that tactile sensation is not only essential in normogravity, but also, and of special importance, in short-term hyper- and microgravity. Thus, the follow-up study did no longer consider head-referenced interfaces, reducing the levels of placement condition, which in turn raised the sample size for the remaining world and hand reference. Besides this level reduction, the follow-up study also implied changes in participants' posture, which was modified from a lying (horizontal floating under 0*g*) to a an upright one (vertical floating in 0*g*) that is more realistic for the intended application. It can be assumed that such changes in participants' posture would not cause different results for the head-referenced interface within the meaning that it not could compensate the lack of tactile sensation.

The findings of the follow-up experiment showed that during the pre-tests on ground and without Scopolamine (1*g*-PRE) the visuomotor performance did not differ between the world-referenced interface and the hand-referenced bimanual interface, replicating the findings of the first study. However, under normogravity in flight (1*g*) the performance differed between the interfaces for the pointing frequency, the response time and pointing accuracy, and showed a significantly improved performance in favor of the handheld interface. This supports the findings from Lindeman et al.

(1999) showing that faster pointing was achieved by using a bimanually handled handheld interface compared to an unimanually handled world-referenced interface, but is not confirming the results of Harrison et al. (2011) showing that pointing towards world-referenced targets provided more precise pointing than pointing towards hand-projected targets. This could be explained by the way of presenting the targets. While Harrison et al. projected the targets directly on the hand surface, during the parabolic flight study a surface was held at the palm of the hand, which provides a more stable base. Considering the changed level of gravity (1.8*g*, 0*g*), the results showed that the visuomotor coordination is affected in the same way than under 1*g* in flight, revealing that the hand-referenced bimanual interface significantly outperformed the world-referenced interface under hypergravity (1.8*g*) and microgravity (0*g*). The findings also revealed that initial orienting towards the interface timely benefited from the bimanual interface under 1*g* in-flight condition and, more important, under microgravity that I have already observed during the first experiment. Consequently, the research question Q-1.2 can be answered with "Yes" related to its performance issue, confirming that bimanual handling improve the performance of visuomotor coordination. Because the handheld interface also provides inside-coded targets, the research question Q-1.1 can be partially confirmed, although the head reference implies the same target location coding, but its lack of haptic cues weakens its advantage of inside-coding.

As was done for the first experiment, the data were also analyzed with respect to possible effects of altered gravity. Therefore the in-flight levels (1.8*g*, 1*g*, 0*g*) were compared on the same stage of placement condition. In cases where I observed such effects, the visuomotor performance was always deteriorated under hyper- and microgravity conditions (Q-2). The hand-referenced bimanual interface was affected by microgravity on the reaction time during initial orientation, the error rate and pointing accuracy, as well as by hyper- and microgravity on the pointing frequency. However, the world-referenced interface was affected by microgravity in the same way as the bimanual interface, except the pointing accuracy, which was not affected. Like the bimanual interface, hypergravity was affecting the pointing frequency, but the error rate and initial reaction time in addition. The world reference caused also differences between hyper- and microgravity for the initial reaction time and error rate, showing that microgravity worsen the performance even more. Regardless from the placement reference, the response time was never affected by changes of the gravity load. Thus, research questions related to gravity changes can be answered as follows: Changes in gravity reveal timing effects, but only during the initial orientation, while the pointing response time otherwise remains unaffected (Q-2.1). However, the question asking for effects on the pointing frequency can be clearly answered with "Yes" (Q-2.2). While these two answers are valid for both placement conditions, the question related to the pointing accuracy can only be confirmed for the handheld interface, while accuracy revealed from the world-referenced interface was not affected by changes of gravity (Q-2.3).

As was mentioned in the beginning of the discussion, the visuomotor coordination during the pre-test (1*g*-PRE) was not affected by the placement conditions, but benefitted from the hand reference during the normogravity test in flight (1*g*). Thus, it can be assumed that something caused a performance bias. To verify this, the normogravity conditions (1*g*-PRE, 1*g*) were compared on the same level of placement condition, which enabled me to investigate whether Scopolamine or flight condition have influenced the performance. While the pointing performance during the first experiment was not affected by such phenomena, this time the visuomotor performance showed such effects for both reference types. While the performance using the world-referenced interface was timely affected during the initial orientation phase, the body-referenced bimanual interface was affected on the

pointing response time. But both placement conditions were affected on the pointing accuracy. Thereby the flight condition caused a decline for the time effect for both references, while a decline for the accuracy was only revealed by the world reference. The handheld interface, however, yielded more precise pointing under the flight condition. It is not clear for me which condition has caused these effects, Scopolamine or flight, but I can assume that the modified condition of participants' posture was responsible in addition.

Besides performance measures, visuomotor coordination was also evaluated by workload indicators. This was subjectively assessed in the same way as during the first experimentation. Because the physical characteristic of parabolic flights do not allow to decide between the in-flight levels, this time the workload was additionally assessed by cardiac responses revealed by the heart rate variability (HRV). Thus, during the follow-up study I distinguished between self-reported and physiological workload. Replicating the findings of the first experiment, the rating of the self-reported workload did not differ under the 1*g*-PRE condition on ground, which indicates that the condition either of altered gravity, Scopolamine or flight was affecting the subjective workload. However, the summarized workload in flight (1.8*g*, 0*g*, 1*g*) revealed a rating in favor of the hand-referenced interface, which is not replicating the finding of the first experiment, but could also be explained by changing participants' posture from a lying to an upright one. In contrast, although the analysis of performance indicators and the subjective workload showed that visuomotor coordination benefitted more from the hand-referenced bimanual interface, the assessment of the physiological workload revealed an opposite effect. By analyzing the strain index assessed by the HRV, the data showed that most participants were more strained using the bimanual interface than the world-referenced interface. This can be caused by the interlimb coordination during the bimanual handling mode, which required more physical effort resulting in a higher heart rate. That leads to a lower heart rate variability implying higher strain. Thus, the second part of the research question Q-1.3 cannot be answered unambiguous. On the one hand, it can be confirmed that bimanual handling improves the self-reported workload, while, on the other hand, a positive answer needs to be rejected by physiological observations. Comparing related findings revealed from normogravity research cannot be considered, because the review of such works (see section 4.1.3) revealed that any workload assessment has been neglected, although the presented experiments have shown its supplementary value.

In conclusion, the results based on the in-flight performance of the follow-up study indicate that visuomotor coordination significantly benefits from an interface specification, which combines inside-coded target locations and tactile sensation by haptic cues. Providing these characteristics implies a body-referenced interface, like the hand, which requires asymmetric bimanual handling. Although such handling mode in turn is apparently causing an increase of physical stress, hand-referenced interfaces also contribute to least subjective workload. Bearing in mind that only the in-flight conditions are discussed here, it was proven that these findings are not only valid in normogravity, but also, and most important in the context of this research, in hyper- and microgravity conditions. Otherwise, the results revealed from the performance on earth's surface suggest in turn that the visuomotor coordination is not depending on the target location coding and handling mode. It only must be ensured that tactile sensation is provided, as already deduced from the initial experiment. The performance bias between ground- and flight-based performances cannot be answered unambiguously, because it is not clear whether the intake of Scopolamine, the flight condition or even their interaction was responsible. The most relevant issue in microgravity research deals with the proof and confirmation of altered gravity effects. In this context, comparing the

performance revealed from the gravity loads in flight, such effects could be proven and was not dependent on the interface specification, but varying in its indications. Therewith, the results showed clear evidence that the visuomotor coordination resulted from an world- and hand-referenced interface is deteriorated in hyper- and microgravity, whereby the absence of gravity partially caused more decreases. Although the related research questions (Q-2.1- Q-2.3) are answered so, it is not clearly answered to the overall question (Q-2), asking how changes in gravity affects visuomotor coordination. Hence, to answer this question, a motion analysis of the underlying goal-directed pointing movements was conducted, which is presented in the subsequent section.

8.5 Motion Analysis: Changes in Goal Directed Pointing Movements

In the prior presented parabolic flight study (see section 8.4) I showed that short-term hypergravity and microgravity deteriorated the visuomotor coordination resulting from an outside-coded world interface (PA) and an inside-coded bimanual interface (BA). Besides the fact that the visuomotor coordination generally benefits more from the BA interface than from PA, both reference types were affected by hyper- and microgravity in nearly same way. To understand these visuomotoric disturbances and differences between the world- and hand-references interfaces, it is useful to study the underlying goal-directed pointing movement. Hence, I performed motion analysis of the pointing finger that investigates the motor control during aimed pointing for both interface types under altered gravity using the in-flight levels (1*g*, 0*g*, 1.8*g*) and raised the question: How altered gravity influenced the goal-directed pointing movement?, which corresponds to the research question Q-2, asked in section 8.1. To answer these questions the motion analysis was twofold and investigated on the one hand the mislocalization of the pointing targets and on the other hand the path trajectories of the pointing movement.

8.5.1 Objectives

The mislocalization of a target indicates how a target was reached with respect to the body's longitudinal z-axis. Several studies under altered gravity conditions observed the phenomenon of the *"elevator illusion"* that results from spatial disorientation and generates perceptual mislocalization of objects as a function of the active gravitational force level (Cohen, 1973; Lackner & DiZio, 1996). This means that the position of stable objects will be visual perceived below its real position under microgravity and above its position under hypergravity. Thereby aimed pointing movements towards such objects can lead to either overshooting or undershooting of the target, but findings are often inconsistent about the level of the gravitational force. For example, Fisk and colleagues (1993) observed a higher number of overshooting under microgravity for rapid arm movements, while Carriot and colleagues (2004) observed undershooting under microgravity. Also, Bringoux and colleagues (2012) found that subjects undershot the target orientations in microgravity, while they overshot in hypergravity.

With respect to targets' mislocalization and the studied placement conditions I aimed to answer the following questions:

1. What type of target's mislocalization (overshoot, undershoot) was most prominent?
2. Has altered gravity affected the mislocalization of the targets?

For studying the characteristics of the underlying movement trajectories, the two-component model for goal-directed pointing movements introduced by Woodworth (1899) was applied. The two-component model states that an aimed movement consists of two different phases. Thereby the initial portion, often termed as *ballistic phase*, is relatively rapid and brings the limb, like the pointing finger, into target's proximity. The latter portion of the movement is referred as *correction phase* and used to adjust the movement in order to come closer to the target and finally to hit the target. Thereby various forms of discrete feedback, like visual feedback, are used to reduce errors in aiming that are caused by the initial ballistic phase. Beside of numerous works in movement and motor behavior science (for survey: Elliott et al. 2001; Elliott & Kahn, 2010), there are only few researches on aimed movements for direct object selection in virtual and augmented environments (Bowman et al., 1999; Grossmann & Balakrishnan, 2004; Teather & Stuerzlinger, 2011; Looser, 2007), and especially with focus on the two-component model (Liu et al., 2009; Liu & Liere, 2009). Because the gravitational force is always integrated in the organization of aimed pointing movements under earth like conditions (Lackner, 1993), it becomes interesting how these two phases are looked alike under altered gravity, in particular under the absence of gravity.

With respect to the two-component model and the studied placement conditions I aimed to answer also the following questions:

3. Did pointing movements revealed typical trajectory pattern under altered gravity?
4. Did altered gravity cause timely changes in the path lengths of the ballistic and correction phase?

8.5.2 Target Mislocalization

To investigate the phenomenon of the *"elevator illusion"* the position of the target's centre $P_{centre}(x, y, z)$ with the final position of intersection $P_{hit}(x, y, z)$ within this target was compared. Thereby I only considered the value of the upright Z-axis and determined a target as overshot by $P_{hit}(z) > P_{centre}(z)$ and as undershot by $P_{hit}(z) < P_{centre}(z)$, as well as calculated the difference d_z by $P_{centre}(z) - P_{hit}(z)$. Targets' mislocalizations are presented by the frequencies of over- and undershooting and the mean difference d_z per interface placement and gravity load across all participants. Analyzing the impact of altered gravity on the ratio of the frequencies of under- and over shoot targets was done by a test of independence (SAS® PROC FREQ with option *"chisq"*) at the same stage of placement method. Also the difference d_z with respect to the gravity vector was evaluated by analyzing the variance (SAS® PROC GLM) between the in-flight *g*-levels (1*g*, 0*g*, 1.8*g*) on same stage of placement and 1*g* as control level. The significance level was set at 5 %. Appendix H additionally provides the descriptive statistic for the outcomes showing measures of the central tendency and variability related to target's mislocalization.

The results showed that the targets were overshot in general with both placement conditions under all in-flight *g*-levels. Using the PA method (see Figure 8.29, left) revealed most undershoots under 1.8*g* (20.31 %), while the frequency of undershot targets under 0*g* (16.80 %) and 1*g* (16.49 %) were approximately the same. But analyzing the ratio of under- and overshoot targets between the *g*-levels showed that altered gravity did not affect the ratio of frequencies, $\chi^2(2, N = 1393) = 2.86$, $p = .2454$. Considering the difference d_z for PA (see Figure 8.29, right) showed that participants pointed with the smallest discrepancy in the upright Z-axis under 1.8*g* ($M = 3.99$ mm), while under 0*g* ($M = 4.88$ mm) and 1*g* ($M = 4.68$ mm) they pointed with about the same discrepancy. Statistical analyzing revealed a significant main effect of altered gravity on the difference d_z, $F(2, 1392) = 4.62$, $p = .01$.

Post-hoc analysis using the control level 1*g* with Bonferroni corrections (α/2 = 0.025) showed that the effect of microgravity was not significant, $F(1, 943) = 0.43$, $p = .5111$, while the difference d_z was significantly affected by 1.8*g*, $F(1, 1010) = 6.06$, $p = .0140$.

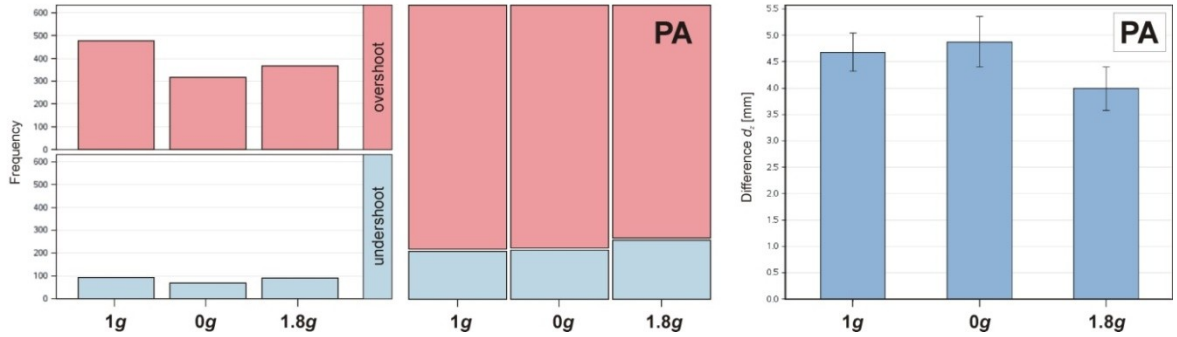


Figure 8.29: Target mislocalization per in-flight *g*-level using the PA method:*(left)* frequency of over- and undershot targets and percentage distribution of over- and undershot targets, and *(right)* difference d_z.

Using the BA method (see Figure 8.30, left) revealed most undershoots under 0*g* (28.62 %) and fewest under 1.8*g* (18.62 %). Analyzing the ratio of under- and overshoot targets between the *g*-levels under the BA condition revealed a significant difference, $\chi^2(2, N = 1801) = 19.59$, $p < .0001$. For post-hoc analysis, two separate tests with Bonferroni correction were required resulting in a significance level of α/2 = 0.025 for each test. Comparing the ratio between 1*g* and 0*g* revealed a significant effect of microgravity, $\chi^2(1, N = 1237) = 12.44$, $p = .0005$, while the comparison between 1*g* and 1.8*g* showed that hypergravity did not affect the ratio of frequencies, $\chi^2(1, N = 1228) = 0.39$, $p = .5625$. Considering the difference d_z for BA (see Figure 8.30, right) showed that participants pointed with the smallest discrepancy in the upright Z-axis under 0*g* ($M = 2.72$ mm), while under 1.8*g* ($M = 3.00$ mm) and 1*g* ($M = 3.01$ mm) they pointed with about the same discrepancy.

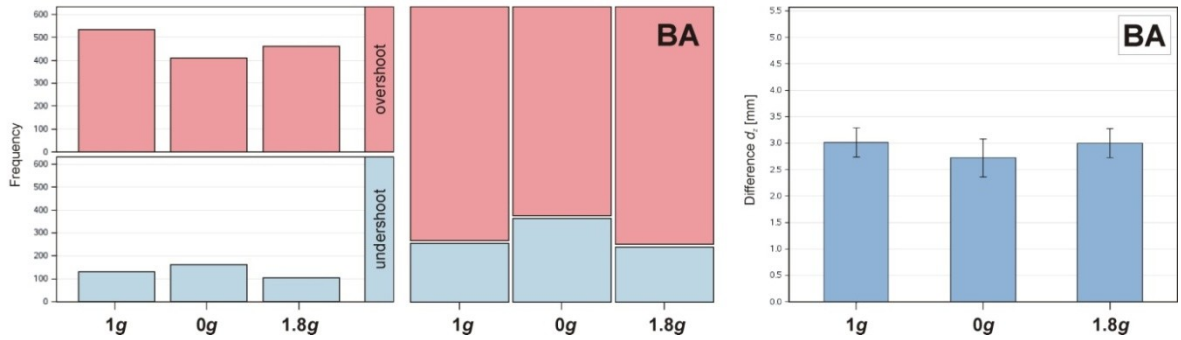


Figure 8.30: Target mislocalization per in-flight *g*-level using the BA method:*(left)* frequency of over- and undershot targets, *(middle)* percentage distribution of over- and undershot targets, and *(right)* the difference d_z.

But statistical analyzing revealed that altered gravity did not affect the difference d_z, $F(2, 1800) = 1.09$, $p = .3370$. Although the PA interface caused greater differences along the gravity vector, visual inspection of the left-hand bar charts of Figure 8.29 and Figure 8.30 revealed that the number of over- and undershoots was higher with the BA than with PA interface, especially under microgravity. This may be an indicator why the BA interface was affected in the pointing accuracy under microgravity, while the PA interface is not (see section 8.4.4.1).

8.5.3 Two-Component Model

To investigate the effect of altered gravity (0*g*, 1.8*g*) on the trajectory of the pointing movement for a discrete response I analyzed the displacement in movement, the pointing velocity and acceleration over time and for each placement condition (PA, BA) separately. By comparing their mean values against the 1*g* level, tests should reveal whether the condition of altered gravity has affected the displacement, velocity and acceleration. Furthermore, Woodworth's two component model (1899) was applied to determine the ballistic and corrective portion of the movement by means of the velocity profile. With respect to the two phases, I analyzed the impact of altered on the duration of each phase. All statistical tests were performed by a linear mixed model using PROC MIXED (SAS®) with simulated adjustment ("*adjust=simulate"*), 1*g* as control level (*"pdiff=control('1g')"*) and a significance level of 5 %.

8.5.3.1 Data Preparation

Motion data of the pointing finger of each participant were recorded, which had to be preprocessed. The location of the finger was determined in 3D Cartesian coordinates (x,y,z) to produce the movement displacement, velocity and acceleration over time. One target hit yielded an average of 62 tracking frame across the in-flight *g*-levels for PA, and an average of 59 tracking frames for BA.

The following steps were performed for the data preparation of the tracking paths using SAS®:

1. Generate a unique identifier for each target hit and assign each tracking frame to its identifier.
2. Perform the following steps for each tracking path:
    2.1 Calculate the time difference and the differences of displacement in x, y, z direction from the previous entry.
    2.2 Calculate the 3D Euclidean distance by the differences of displacement.
    2.3 Calculate the velocity from the Euclidean distance and the time difference.
    2.4 Calculate the velocity difference from the previous entry.
    2.5 Calculate the acceleration by velocity and time difference.
    2.6 Calculate the relative duration between the first and current entry related to the target hit.
3. Delete the first entry of each target hit.
4. Sum the time differences, their durations and count the number of tracking paths for each target hit.
5. For each data set generate a time class by relative duration ("*0.1 + int(10*duration)/10* ").
6. Select target hits that fulfill the following criteria:
    - 3D Euclidean distance > 0
    - number of false target hit = 0
    - number of frames > 20
    - 0.8 s < absolute duration of target hit < 5 s
7. Determine the assignment of these time classes and select a maximal time class t_{max} across all participants, presuming that the eligible time classes are fairly frequented.
8. Calculate the median displacement per axis for each participant and across all participants.
9. Calculate the median velocity for each participant and across all participants.
10. Calculate the median acceleration for each participant and across all participants.

Because the mean response time of the PA condition across all in-flight *g*-levels yielded 1283.28 ms, the related maximal time class t_{max} was set at 1.3s. The mean response time of the BA condition yielded 1207.56 ms, thus t_{max} for BA was set at 1.2s. The allocations of the number of tracking paths to the time classes are presented in Appendix H, which also provides time series plots showing the distribution of the pointing velocity per interface condition.

8.5.3.2 Movement Displacement, Velocity and Acceleration

Figure 8.31 shows the displacement in movement over time across all participants per gravity load (1*g*, 0*g*, 1.8*g*) and for both interface placement conditions (PA, BA). It reflects the changes of the position of the pointing finger during the movement of a target hit. The displacement profiles show the smallest position changes in the beginning of the movement followed by a fast increase for both placement conditions and under all *g*-levels. All trajectories reached their global maximum at 0.5 s. Visual inspection of the maximal position peaks under 1*g* and 1.8*g* revealed similar displacement under the PA condition, while it was the same under the BA condition. Related to the total movement I can assume that hypergravity did not cause notable changes in position for both placement conditions. But it is clear to see that microgravity (0*g*) caused smaller position changes in the area around the peak of the trajectories for both placement conditions, and always smaller changes in the total movement trajectory for the BA condition. Comparing the placement conditions showed that the pointing finger with PA was moved with bigger changes in general than with BA. Related to the global peaks in position I noted that BA caused a decrease of about 20 % in position changes compared to PA.

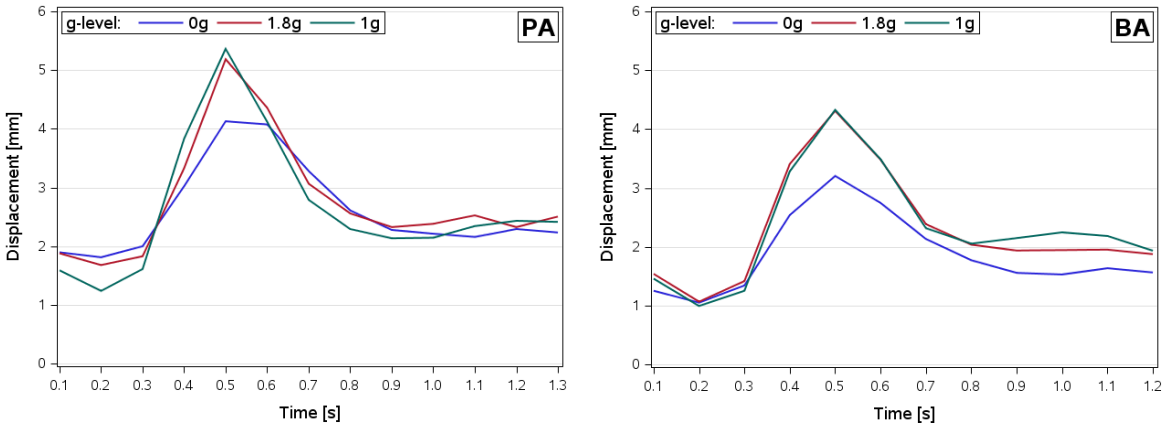


Figure 8.31: Displacement in pointing movement for the in-flight *g*-level across all participants: *(left)* under the PA condition and *(right)* under the BA condition.

The velocity profiles (see Figure 8.32) revealed similar characteristics of the resulted trajectories as for the profiles of the displacement. After the pointing movement started with the lowest velocity the movement speeded up fast and reached its upper reversal point at the end of the first third of the movement. The following movement showed a fast decrease in pointing velocity, while the remaining movement runs out with minimal changes in velocity. The peak in velocity was similar affected by hypergravity and did not show notable differences between 1*g* and 1.8*g*, in particular for BA. In contrast, microgravity caused a decrease of about 20 % for PA and about 30 % slower movement under the BA condition. Just as the pointing displacement, microgravity continuously caused the slowest movement in the total trajectory with BA. Comparing the placement conditions showed that the pointing finger was faster moved with PA than with BA under all *g*-levels, but

differed most under microgravity, where BA caused a decrease of about 15 % in pointing velocity compared to PA.

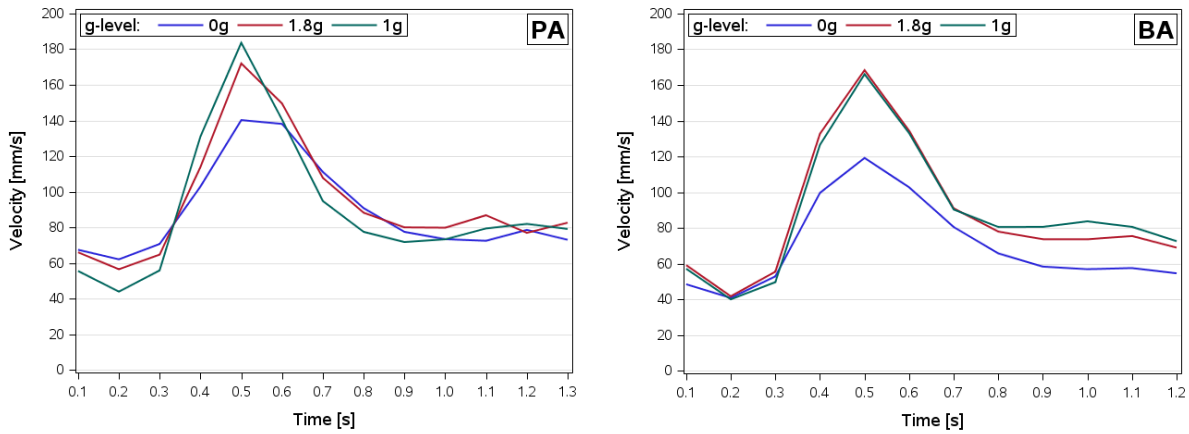


Figure 8.32: Pointing velocity for the in-flight *g*-level across all participants.

With respect to the pointing acceleration (see Figure 8.33), the movement accelerated fast, after an initial slow down at the beginning, and reached its global maximum, as expected, just before the peak in displacement and velocity at 0.4 s. The movement rapidly slowed down and revealed the biggest decrease in acceleration directly after the global peak. Thereafter the acceleration rose constantly, but remained below half of the global maximum. The condition of microgravity caused a decrease of acceleration in the area around the peak of the trajectories, but was not stalling that much during the rapid fall phase than under 1g and 1.8*g*. The remained movement was not notable affected by altered gravity conditions. Comparing the interface placement conditions showed that using BA led to higher accelerated motions than with PA under all g-levels, in particular under altered gravity conditions. Thus, the slower pointing performance revealed from the world-referenced PA interface (see section 8.4.4.1) can be explained by a less accelerated pointing movement, especially in the beginning.

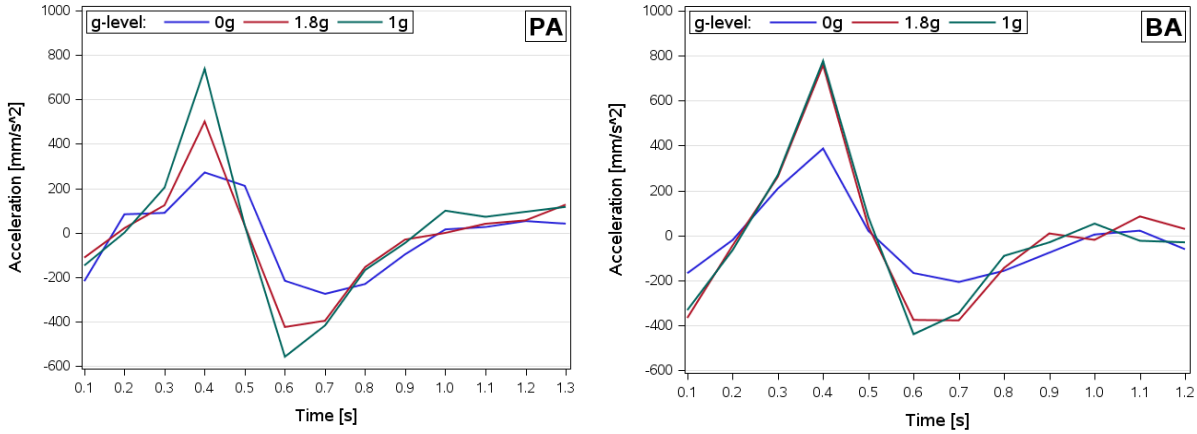


Figure 8.33: Pointing acceleration per in-flight *g*-level across all participants.

Analyzing the variation of the mean displacement, velocity and acceleration revealed different findings. While pointing towards the PA interface was not affected by altered gravity, the pointing movements under the BA condition showed effects for the displacement and the velocity (see Figure

8.34). Thereby changes in the position of the pointing finger and the pointing velocity were significantly decreased by microgravity (see Table 8.19), while hypergravity yielded no effects.

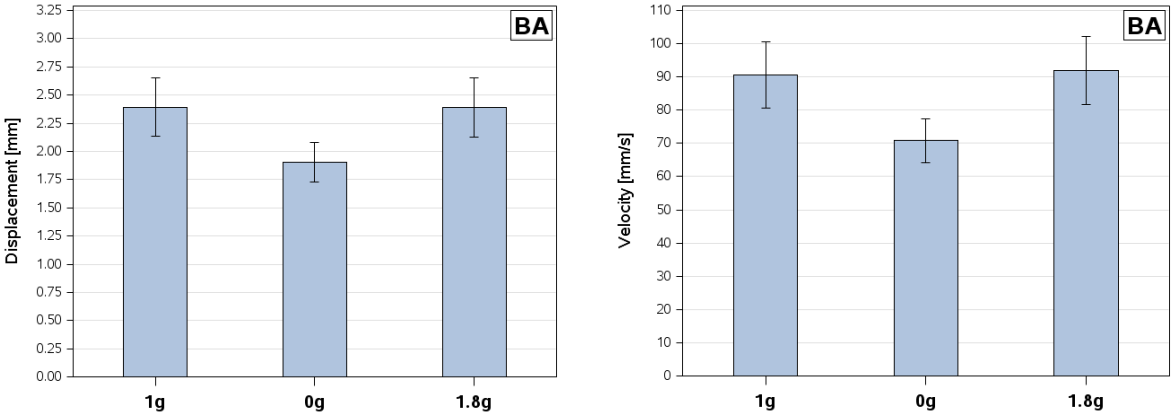


Figure 8.34: Mean movement displacement and mean pointing velocity of BA per in-flight *g*-level (CI=95%).

Measure	*g*-level		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Displacement (BA)	0g	1*g*	249	-2.93	0.0037	0.0071	*
	1.8g	1*g*	249	-0.01	0.9885	0.9998	
Velocity (BA)	0g	1*g*	249	-3.08	0.0023	0.0044	*
	1.8g	1*g*	249	0.23	0.8213	0.9633	

Table 8.19: Comparison of displacement and velocity between 1*g* and the altered *g*-levels (0*g*, 1.8*g*) under BA using SAS® PROC MIXED (pdiff=control('1*g*'), adjust=simulate).

8.5.3.3 Two Phases Movement Times

To apply Woodworth's two-component model, I subdivided the total movement of goal-directed pointing into the ballistic and correction phase. Therefore two criteria for the subdivision as a function of the resulted pointing velocity were defined (Meyer et al., 1988). With the first criterion I assigned the global maximum of the velocity as part of the primary sub movement that constitutes the ballistic portion of the movement. Directly after the peak in velocity, I started from that point to identify the end of the ballistic phase using the second criterion indicating that the following correction phase is induced by the first local minimum in velocity. For exposing effects of altered gravity, the resulted path lengths by the two-phase movement times were compared that reflect the duration of the ballistic and correction phase. The criterion to divide the total movement path has been applied on the prepared data with respect to the velocity, separately, for both placement conditions. Figure 8.35 shows the velocity profile per gravity load (1*g*, 0*g*, 1.8*g*) for the PA interface (top) and the profiles for the BA interface (bottom). All plots are marked with a vertical reference line splitting the movement trajectories into the ballistic and correction phases. As can be seen, aimed pointing movement revealed same length of the ballistic phase under normogravity (1*g*) and hypergravity (1.8*g*) under both placement conditions, while microgravity (0*g*) caused longer ballistic phases. As I expected, the longest ballistic phase was yielded from PA under 0*g* in general, because this interface type was most slowed down by initial orienting and in its responses (see section 8.4.4.1).

Figure 8.36 shows the mean total movement time subdivided by the two phases for both placement conditions per *g*-level. Also here, it can be seen that for both placement conditions, microgravity caused the longest durations of the ballistic phase, consequently shortening the correction phases,

which are generally used for fine-tuning the goal-directed pointing movement. Because the outside-coded PA interface led to bigger changes in displacement regardless of the gravity load (see section 8.5.3.2), it is possible that this interfered with fine-tuning at the end of the movement. Consequently, this could be the reason why the PA interface led to significantly less precise pointing than the BA interface under all in-flight levels (see section 8.4.4.1). But comparing the duration of the phases between the in-flight gravity loads at the same stage of placement did not reveal significant effects for both interface placements.

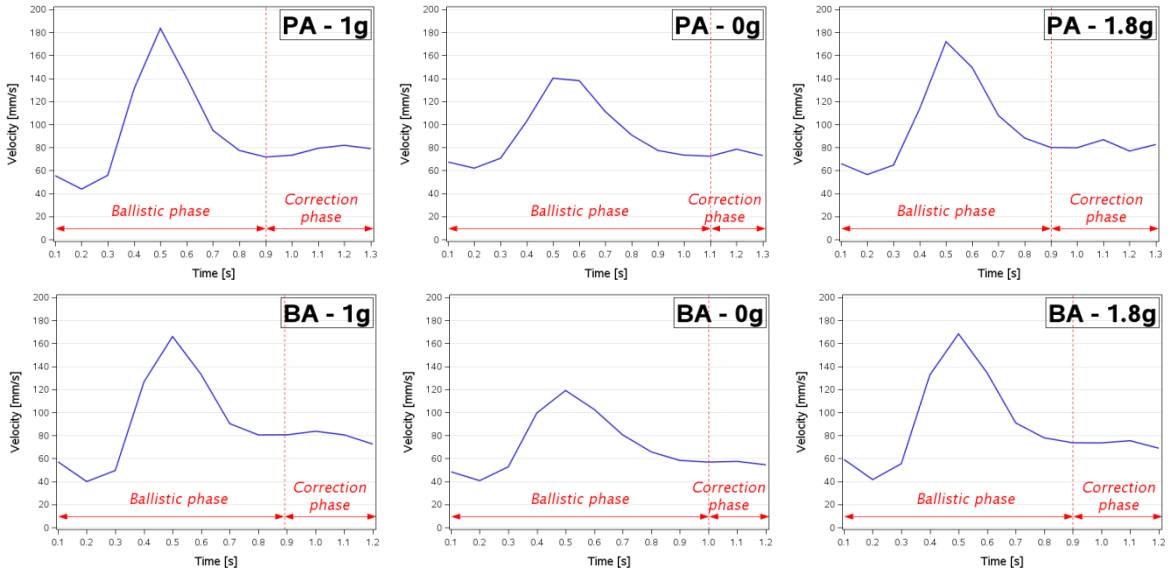


Figure 8.35: The ballistic and correction phases: *(top)* of the PA condition, and *(bottom)* of the BA condition.

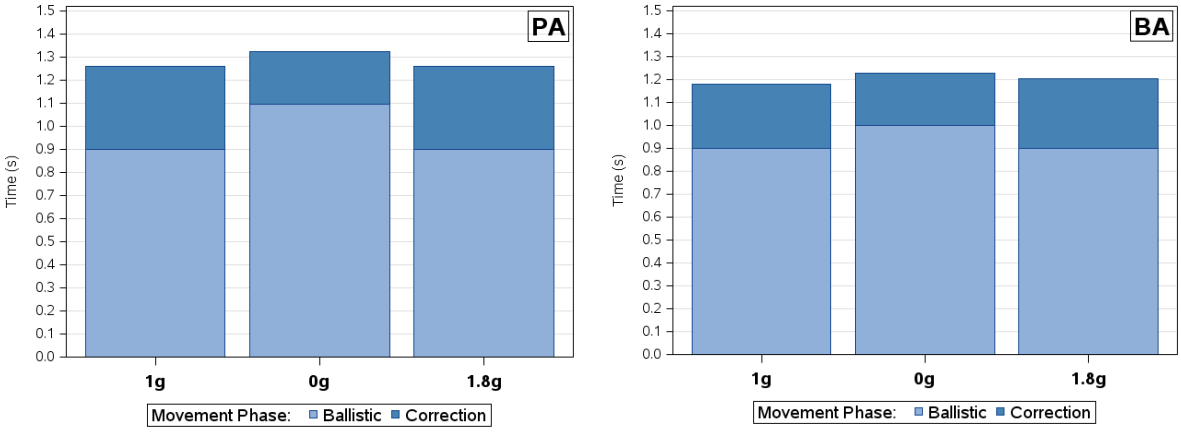


Figure 8.36: Mean movement time, showing the portion of the ballistic and the correction phase: *(left)* under the PA condition, and *(right)* under the BA condition.

8.5.4 Discussion

In section 8.4 it was shown that hyper- and microgravity affect pointing towards an outside-coded interface (PA) and an inside-coded interface (BA), significantly increasing the initial reaction time and error rate, as well as decreasing the frequency of correct target hits. Thereby the BA interface was additionally affected in accuracy by microgravity, while the hypergravity shortcomings revealed from the PA interface were extended by an increased error rate. Nonetheless, pointing towards the

inside-coded targets generally contributed to a better performance than the outside coded interface. To understand these findings I performed motion analysis of the underlying pointing movement of hitting a target. By evaluating the mislocalization of the targets and applying Woodworth's two-component model (1899) I investigated how altered gravity has affected the position of the pointing finger, the pointing velocity and acceleration over time and divided the movement in its ballistic and corrective part by means of the velocity profile. The motion analysis was performed for aimed movements yielded from the outside-coded interface and the inside-coded bimanual interface under hyper- and microgravity, as well as under normogravity condition in flight as a control level.

With respect to the target's mislocalization, I can conclude that in general the targets were mostly overshot, regardless of the interface placement und gravity load. This may be explained by a too high positioning of the intersection sphere that was attached to the virtual hand model (see section 8.2.3). Nonetheless, pointing towards the PA interface led to more undershoots and significantly smaller discrepancy in the upright Z-axis under hypergravity, while microgravity did not cause changes. In contrast, pointing towards the bimanual BA interface, led to significant more undershoots under microgravity, while the discrepancy in the upright Z-axis was not affected by modified conditions of gravity. It could also be shown that the BA interface caused higher numbers of under- and overshoots than the PA interface, in particular in the absence of gravity, which may explain why BA caused significantly most precise pointing under microgravity. Exploring the moving trajectories over time related to the displacement, the velocity and the acceleration revealed the same characteristics of the profiles for both placement conditions under all in-flight *g*-levels (1*g*, 0*g*, 1.8*g*). Thereby the profiles showed an initial rapid increase, which led to the global maximum, followed by a rapid decrease and running out until the end of the movement. Compared to the pointing motion under normogravity, the condition of hypergravity did not reveal notable differences in the profiles, although the PA interface showed slight decreases. In contrast, all profiles showed that microgravity caused a drastically decrease for both interface types. Analyzing the mean values revealed that microgravity significantly decreased the pointing displacement and velocity for operating the BA interface, while the PA interface was in general not affected by altered gravity, although it showed much greater displacement than BA. Such large displacements in pointing can interfere with fine-tuning at the end of its goal-directed motion and thus may explain why the PA interface led to a significantly decreased pointing accuracy than the BA interface. However, the lower velocity revealed by the BA interface may also indicate why microgravity affected the pointing accuracy for this interface, because a lower velocity generally destabilizes a motion and, thus, this can be the reason for reduced precise pointing. Dividing the velocity profile into a ballistic and a corrective part revealed that microgravity caused a general prolongation of the ballistic phase and thus shortening the correction phases, which is responsible for adjusting the motion at its end and reducing errors in aiming. Thus, the shorter correction phases could justify the significant higher error rates yielded from both interface types, especially in microgravity.

In conclusion, this motion analysis supplied answers to the research question Q-2, which asks how changes in gravity affects visuomotor coordination. It provided some indications why the inside-coded hand reference led to a reduced accuracy under microgravity and why microgravity caused an increased error rate for both interface types. But there is no clear indication why the outside-coded interface led to significantly higher error rates under hypergravity. This analysis also provided indications why the world reference led to more imprecise pointing and slower movements than the hand reference.

8.6 Summary

To support control and symbolic-input tasks for payload operations within a future AR environment, it is important to identify the most suitable condition for interface placement, especially under microgravity. Hence, in this chapter I have presented research, investigating the spatial placement modality of interactive AR interfaces in consideration of studying the effect of altered gravity on its resulting hand-eye coordination. Therefore, three placement conditions were identified, each varying in its spatial reference related to humans' body, its handling mode and the presence of haptic cues. For comparison I set up one outside-coded interface that requires aimed pointing movements towards external space (world-referenced), and two inside-coded interfaces that were referenced to humans' body and head. The body-referenced interface constituted a hand-held interface that demands bimanual coordination, while by the head-reference the interface was aligned to HMD's field of view that did not provide haptic cues for direct object selection. To find out this placement reference that contributes most to the visuomotor performance and the least workload in altered gravity, in particular in microgravity, I designed two experiments that were conducted under parabolic flight conditions. In a within-subject design, twelve participants (six per experiment) performed a visuomotor task under short-term hyper- and microgravity, as well as under normogravity on ground and in flight. Besides performance metrics, including pointing frequency, error rate, response time and pointing accuracy, I assessed the workload by subjective experiences and physiologically by cardiac responses (only during the second experiment) analyzing the stress index acquired from the heart rate variability.

Besides verifying the feasibility of operating AR interfaces under altered gravity conditions, the study revealed several findings. First, visuomotor coordination for direct AR selection benefited from interfaces with the presence of haptic cues. This finding was regardless of the gravity load. Whether under normogravity or under altered gravity, the visuomotor performance and self-reported workload have deteriorated significantly by the absence of haptic cues. Secondly, under normogravity conditions on earth's surface it did not matter whether the interface for direct AR selection was placed outside or inside the human body frame of reference, affecting visuomotor performance and self-reported workload in the same way. Thirdly, under conditions of altered gravity, the visuomotor performance and subjective workload differed significantly between the outside-coded world-referenced interface and the inside-coded handheld one. Compared to the outside-coded world reference, the performance and self-reported workload were significantly improved by pointing towards the inside-coded interface, although it has required bimanual handling. Fourthly, aimed pointing towards the body-referenced interface showed an increase in physiological strain. This may have been caused by the interlimb coordination during the bimanual handling mode. Fifthly and lastly, a decrease in task performance under altered gravity clearly indicates that hyper- and microgravity has significantly affected the visuomotor coordination, not depending upon whether the world- or the hand-referenced bimanual interface has been used. The findings related to the world and hand references were partially supported by taking a closer look at the underlying pointing movements. It showed that the pointing targets were generally overshot, while microgravity caused a significant increase of undershoots with the handheld interface, which always resulted in a higher number of mislocalized target hits. Furthermore this analysis has shown that under microgravity in general it took longer times to bring the pointing finger to the target vicinity, which was caused by prolonged ballistic phases. This in turn has shortened the corrective part needed for fine-tuning the pointing movement to compensate errors made during the ballistic phase. This, for example, may explain the significant higher error rates resulted from both interface types. Summing up and

informing the design in interface placement, the results of this research suggest that interfaces for control and symbolic-input tasks should always provide tactile sensation by haptic cues, regardless of the present gravity load. For AR systems operated on the ground it does not matter whether such interfaces are placed within a world or a body reference, in this case handheld. In contrast, to inform the design for ISS payload operations in microgravity, it is suggested that such interfaces should be placed inside of the human body frame of reference, for example, hand-held with bimanual handling.

In cases where a bimanual operation of interfaces for control and symbolic tasks is not desired, it would be reasonable that such interfaces are either coded in the same reference than the main application, which implies a world reference, or are decoupled from the application, like the handheld interface, but requiring only one hand for operation, as provided by a head reference. Thus, subsequent work, presented in the next chapter, is deduced from this research and intended to compensate the shortcomings of the world- and head-referenced interfaces. Therefore the associated visuomotor coordination needs to be improved and optimized respectively. To overcome problems raised by the world reference, I have introduced a gravity-adopted resizing approach (see section 9.1), which uses the information about the gravity load to change physical characteristics of the interface. In contrast to world-referenced interfaces, head-referenced interfaces do not need to be integrated into the task environment and benefit from the fact that such interfaces would always be easily accessible. Although inside-coded targets contribute to an improved visuomotor coordination, the findings have shown that the lack of tactile sensation deteriorates pointing towards head-referenced interfaces, especially in microgravity. No matter whether not providing physical short-term regeneration or the repeated localization of the virtual interface plane, the position sense of the pointing arm is weakened. Thus, it is conceivable that the stimulation of isometric forces can boost the motor planning and control of arm movements. An approach to overcome this problem is presented in section 9.2, investigating the effect of increased sensorimotoric load on visuomotor coordination during aimed pointing towards non-haptic interfaces.

Chapter 9

Countermeasures: Enhancement of Direct AR Target Selection by Means of Sensorimotoric Load

As I have shown in the previous chapter, the performance of AR direct object selection using outside-coded interfaces and nonhaptic interfaces that are coded inside the human body frame of reference are impaired under short-term altered gravity. Therefore adequate countermeasures are required. This chapter aims to investigate potential approaches to improve and optimize the associated visuomotor coordination and thus to enhance the usability of such interfaces. Two methods were developed and applied in different ways, using increased sensorimotoric load. In the case of improving the outside-coded world-referenced interface, the characteristics of the operated interface are changed by the present gravity load, while the second approach manipulates the target selection by the generation of isometric force that could be meaningful during the operation of nonhaptic interfaces, which are provided by an inside-coded head-reference. Parts of this chapter have previously been published in Markov-Vetter et al. (2014, 2015, 2016).

9.1 Gravity-Adapted Target and Interface Resizing under Simulated Hypergravity

Changes in gravity cause sensorimotor disruptions in human motor coordination and eye movements that interfere astronauts' work and thus can also deteriorate their task performance during payload operations. In the last chapter, the impact of altered gravity on direct AR object selection for control and symbolic-input tasks was investigated. Therefore two experiments under parabolic flight conditions were performed to find out which interface alignment is the most efficient one and preserves human's visuomotor coordination under altered gravity, especially under microgravity. I compared three alignment conditions of the virtual pointing interface, whereby each condition had its own interrelation between the human body frame of reference and the support of haptic cues. The results showed that aimed pointing movements for direct AR selection under altered gravity benefits from targets with haptic cues and targets that are coded inside the human body frame of reference (e.g., attached to hand), while targets placed outside the body frame deteriorate the pointing performance. However, given that the future main application of an AR supported guidance system, like the MARSOP system (see Chapter 6), is predominantly coded outside of the user's body frame, the maintenance of user performance in object selection requires the introduction of appropriate countermeasures. Supposing that an outside coded AR interface that the astronaut has to operate in space is affected in the same way as the astronaut by gravity changes, it is conceivable that the hand-eye coordination will be improved. Whether detecting labels and annotations by gaze control, or pointing movements during symbolic tasks, an adequate transformation of the AR interface could overcome the human sensorimotoric disruptions under gravity changes. As stated by Fitts' Law, timing effects of the selection performance can be affected by targets' transformation in size and distance (Fitts, 1954). Thereby a supportive approach could be the dynamical transformation of the

AR interface with respect to the active gravity. Before conducting expensive experiments under simulated weightlessness conditions (e.g., parabolic flight), a proof-of-concept study under simulated hypergravity (+Gz) conditions was initially performed. Therefore I developed a gravity-adapted strategy for interface resizing that was studied by AR selection using a direct touch interface under increased gravity conditions. The resizing approach not only affects the size of a pointing target (or label), but also affects the position, i.e. in cases of more than one target also the distance between them is affected.

9.1.1 The Force-Based Sizing Approach

For improving the performance of aimed pointing movements towards virtual targets under altered gravity conditions, I developed a force-based approach for dynamic target transformation. Force-based approaches are typically used for automated positioning of labels and annotations, e.g. in 3D information visualization (Pick et al., 2010; Hartmann et al., 2004). Depending on the active gravity load, a corresponding force affecting target's size and position was calculated. The approach for target resizing and -positioning was derived from the elastic behaviour of soft bodies, which are proportional deformed to the applied gravity load G_{sim}, similar to Hooke's law (Eq. 1). Therefore, I calculated the axial (Eq. 2) and transversal (Eq. 3) strain of the target using empirical values for the modulus of elasticity E and Poisson's ratio v. Thereby, I distinguished between two techniques of target sizing – sizing by compression (SC, Eq. 4) and sizing by elongation (SE, Eq. 5). For evaluation purposes their output was compared with the unmodified sizing technique (SU) as baseline condition that did not affect the targets. For initially experimentation I limited the evaluated parameters by automated target resizing without the transversal strain Δw, but applied the axial strain Δh proportionally to target's height and width. Figure 9.1 shows the resulted sizing techniques investigated in the POC study. The gravity-based changes were also applied to the complete interface, i.e. to targets' position, that resulted in a larger target distance with the SE technique and in smaller distances with the compressed SC technique. While the SC sizing technique provides smaller targets and benefits from smaller target distances, the SE technique offers larger target size at larger distance. To my knowledge, gravity-adapted target sizing was not reported until now. For designing the normal sized targets (SU) I followed the recommended ergonomic size range for push buttons (Department of Defense, 1999) and used a squared target of 15 mm width and height.

$$\sigma = E\varepsilon \quad \text{with} \quad \varepsilon = \frac{\Delta h}{h} \quad \text{and} \quad \sigma = \frac{F}{A} \tag{1}$$

$$\Delta h = \frac{F * h}{A * F} \quad \text{with} \quad A = h^2 \quad \text{and} \quad F = G_{sim} \tag{2}$$

$$\Delta w = -w * v\frac{\Delta h}{h} \tag{3}$$

$$h_{SC} = h - \Delta h\,, \qquad w_{SC} = w - \Delta w \tag{4}$$

$$h_{SE} = h + \Delta h\,, \qquad w_{SE} = w + \Delta w \tag{5}$$

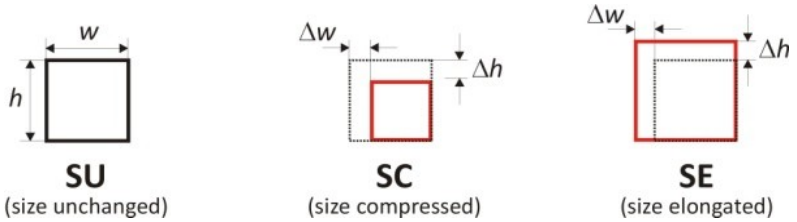


Figure 9.1: Resulted resizing methods applied to the pointing targets and the overall interface.

9.1.2 Research Objective

Focused on sensorimotor hand-eye coordination this research aimed to investigate the effect of gravity-adapted interface resizing on performance during an AR visuomotor task presented by a head-mounted display under increased gravity conditions. In accordance with the active gravity load, I expected that the variations of the pointing performance (e.g., response time, speed, accuracy; etc.) are correlating with the AR interface that was affected by the resizing approach and led to the following question: Does gravity-adapted target resizing affect the performance and workload of direct AR pointing under altered hypergravity conditions? [Q1]

The resulted sizing conditions using the gravity-adapted sizing approach interrelate to the characteristics of Fitts' law that predicts longer movement times at greater distances, as well as at smaller targets (Fitts, 1954). While the elongated method (SE) provides greater targets at larger distances, the compressed method (SC) provides smaller targets at shorter distances. Even though elongated sizing will cause the largest distance, I expected that fast pointing movements under increased gravity conditions benefit from greater targets. Therefore it was hypothesized that elongated resizing (SE) mostly decreases movement times under increased gravity load, because it provides greater targets [H1]. With respect to the physiological workload, an increase in physical effort for increased target distances (SE) was expected. Related to this, I hypothesized that compressed resizing (SC) mostly decreases the workload under increased gravity load, because it provides the shortest target distances [H2]. To answer this question and to test the hypotheses a trade-off study was conducted that was divided into two parts using different simulations of hypergravity. Firstly, a case study was performed, where +Gz load was induced by a long-arm human centrifuge (LAHC-Study). Secondly, an experiment under normogravity and simulated +Gz loads by additional arm weighting (Guardiera et al., 2008) was conducted (Weight-Study). Using a visuomotor task, I aimed to investigate the impact of increased gravity loads on size and distance of a given target interface evaluated on direct object selection by the performance and physiological workload. For evaluating the performance common measures, such as the frequency of correct and incorrect pointing, the accuracy, the response time and the pointing speed, were used. Assessing the physiological workload by HRV recording has only been applied during the Weight-Study. The LAHC and Weight studies did not only differ in the way of hypergravity simulation and workload assessment, they were also distinct in their experimental task. For the LAHC-Study, the experiment task that has already been used during the parabolic flight studies was applied. Because during the Weight-Study additionally the task performance was evaluated related to Fitts' law, I considered the international standard for pointing devices (ISO/DIS 9241-9, 2000) and this time the multi-directional tapping task (MacKenzie, 1992) was applied, which consisted of eight targets arranged in a ring. There have been only few studies applying Fitts' law on evaluation of AR interaction (Rohs et al., 2011), or on head-mounted Virtual Reality pointing (Kohli et al., 2012).

9.1.3 Case Study by Long-Arm Human Centrifuge *(LAHC Study)*

To proof the concept of the gravity-adapted approach initially, I was allowed to perform a user study under +Gz load induced by a long-arm human centrifuge (LAHC, see Figure 9.2), owned by the DLR in Cologne. Human centrifuges enable research in medicine and human physiology during altered +Gz load and are also used to train pilots and astronauts. The study was performed with one participant. The male participant (51 years old, space engineer) is very experienced under altered +Gz load (human centrifuge, parabolic flight) and familiarized with the used AR pointing system and task.

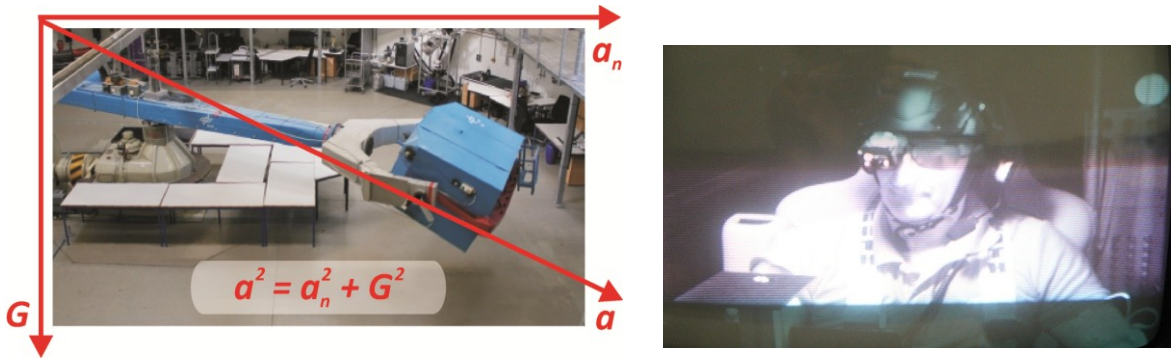


Figure 9.2: LAHC-Study: The used long-arm human centrifuge (5 m radius) with centrifugal acceleration a_n The cabin is swinging out during the rotation with resulted acceleration a in line with subject's long body axis.

9.1.3.1 Apparatus

As was already used for the parabolic flight study presented in the last section, the monocular optical see-through dataGlass2/a from Shimadzu was used (see Appendix B). The HMD was connected to the data processing unit (Lenovo ThinkPad T420s, 2.8 GHz CPU, NVIDIA Quadro NVS 4200M), which was installed under the participant's seat in the centrifuge cabin. To realize pointing with haptic feedback a panel that was installed in front of the participant was used and equipped with a multi-marker configuration. For the pointing purpose a single marker was attached to the participant's fingertip at the dominant hand. The pose data were captured with a mean frame rate of 38.74 fps (SD = 10.05) by the optical sensor at constant artificial light conditions.

9.1.3.2 Experiment Task

In response to visual stimuli the participant should point towards virtual targets under altered +Gz loads while wearing an optical see-through head-mounted display (OST HMD). Pointing in response to visual stimuli was done based on the visuomotor task used for parabolic flight studies (see section 8.2.2). By using a soft AR keyboard with squared keys of 15mm width and height (Department of Defence, 1999), the participant was requested to enter prescribed random pseudo-letters on a virtual keyboard (see Figure 9.3, right). Entering letters onto the keyboard was determined by collision tests of a virtual ray ranging from the origin of the fingertip marker to the top of the index finger. The requested letter was signalled in green, while hitting a correct key was highlighted in red. Because the

Figure 9.3: LAHC-Study: Participant sitting in the LAHC cabin, wearing the OST-HMD and pointing toward the panel *(left)*. The soft-AR keyboard *(right)*.

data processing unit was installed in the cabin of the centrifuge, the participant started the experiment with a virtual start button displayed above the keyboard and hidden afterwards.

9.1.3.3 Experiment Design and Procedure

As shown in Figure 9.4, the study was conducted during three experimental sessions, on three successive days. For the visuomotor task performance four target pools were specified that were counterbalanced presented per sizing technique. A target pool was defined as a pre-randomized series of keys. The completion time for pointing towards the keys of one target pool was predefined by 25 seconds. Within one centrifugation the participant performed the task for two sizing techniques that resulted in a total pointing time of 200 seconds per +Gz load. To adjust the duration of a key pool and the Gz loads, the first day was used for pilot testing. Because it was quite exhausting for the participant to perform arm movements in series for 200 s, I decide to reduce the operation time of a target pool to 20 s (in total 160 s per centrifugation). Pilot testing was done under 1.5*g* and 2*g*.

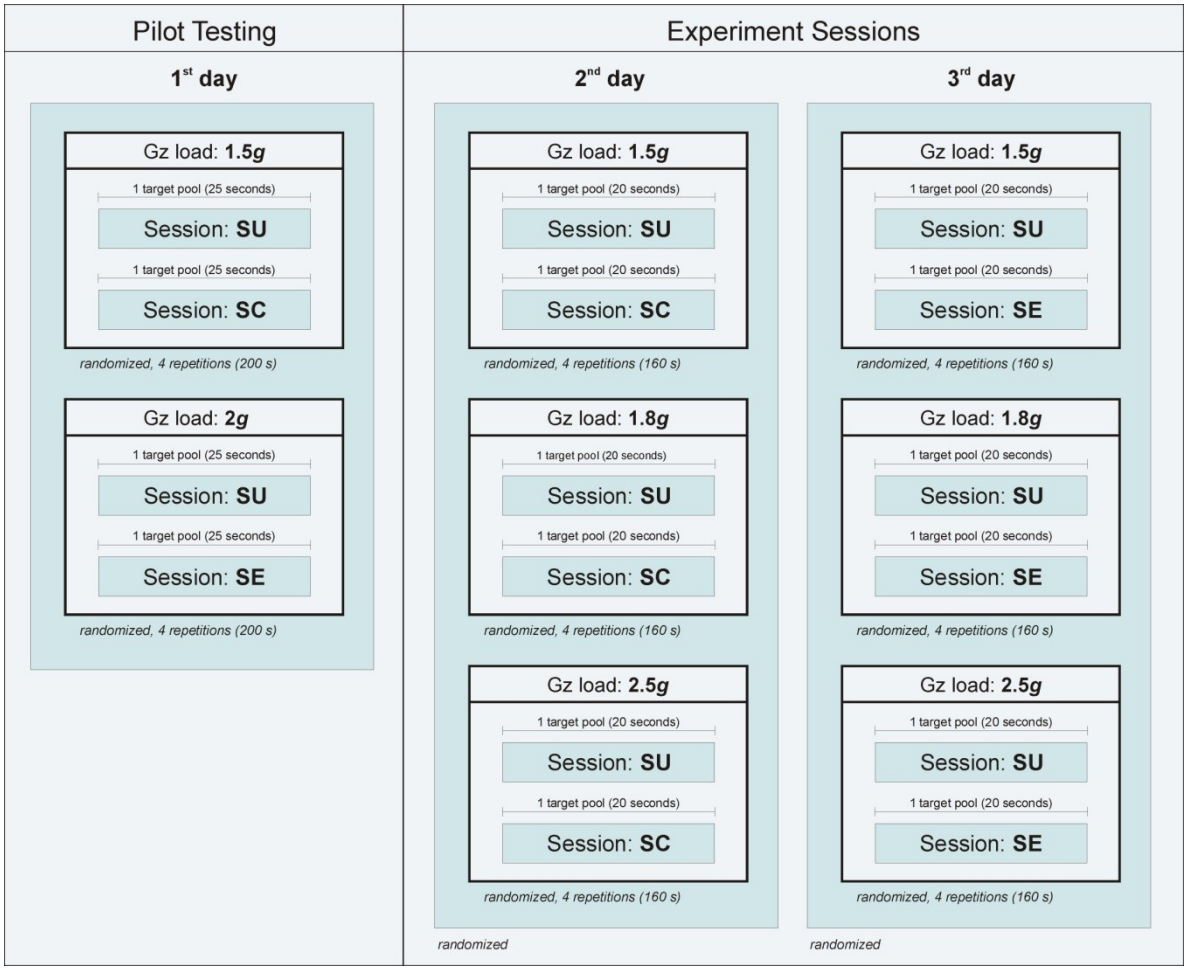


Figure 9.4: LAHC-Study: The experiment procedure performed across three days, using the first day for pilot testing, and the othersday for conducting the experiment randomized for the sizing conditions and Gz loads.

For better differentiation between the Gz loads I decided to perform the experiment under 1.5*g*, 1.8*g* and 2.3*g*. To avoid transition effects between the target pools and the method changes, the first and the last signaled key were not recorded. For physiological regeneration and to limit learning effects there was a 10 minutes break between the changes of the +Gz loads. For the experiment sessions of

the second day, within one centrifugation the participant performed the task under one +Gz load using the unchanged method (SU) and one of the methods with force-based target sizing (SC, SE). Thereby the sizing technique was changed after one target pool. The experiment was performed for SU and SC on the second day and for SU and SE on the third day under 1.5*g*, 1.8*g* and 2.5*g*. Thereby the sizing conditions and the Gz loads were systematically counterbalanced.

9.1.3.4 Results

To compare the sizing conditions I analyzed the frequencies of correct and false target hits, the percentage error rate, the response time and the pointing speed. Thereby a false target hit constitutes that the participant has pointed toward a wrong key. Because the number of resulted target hits was variable by a predefined completion time, I consider the percentage error rate calculated by dividing the total number of triggered targets by the total number of false target hits. The response time mirrors the time elapsed between the visual stimulus onset and motoric response onset, while the speed was calculated by the response time and the Euclidean distance between the centre of the last target and the centre of the present target. For statistical analysis I conducted multiple pairwise comparisons between the sizing methods across all Gz loads and at the same stage of Gz load using SAS® PROC MIXED for repeatedly measured data and SAS® PROC GLIMMIX for objective count data (see section 3.4.2, *Data Analyzing),* whereby the significance level ($\alpha = 5\ \%$) was guaranteed by the Simulate method. To compare the percentage error rate between the sizing conditions, the data values were transformed by the natural logarithm. Aware that the experiment was conducted by only one participant, this case study should prove whether the resizing approach generally affects the performance under altered hypergravity conditions.

Resizing	Correct Hits [ms]	Error [%]	Response time [ms]	Speed [mm/ms]
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Gz = 1.5*g*				
SU	12.63 ± 1.77	6.44 ± 5.87	1367.25 ± 217.33	0.063 ± 0.013
SC	10.75 ± 0.96	13.14 ± 9.05	1483.18 ± 260.69	0.040 ± 0.008
SE	**13.25** ± 0.96	**0.00** ± 0.00	**1341.46** ± 81.74	**0.075** ± 0.005
Gz = 1.8*g*				
SU	13.13 ± 1.46	2.03 ± 3.79	**1300.01** ± 166.84	0.064 ± 0.007
SC	11.50 ± 0.58	7.28 ± 10.11	1338.24 ± 181.74	0.044 ± 0.006
SE	**13.25** ± 0.96	**1.67** ± 3.33	1312.80 ± 133.50	**0.077** ± 0.008
Gz = 2.5*g*				
SU	**12.14** ± 1.07	**4.49** ± 6.11	1322.43 ± 256.35	0.062 ± 0.007
SC	11.00 ± 0.00	19.41 ± 3.49	**1239.52** ± 168.98	0.048 ± 0.007
SE	11.75 ± 0.96	14.02 ± 16.23	1389.13 ± 166.74	**0.074** ± 0.012

Table 9.1: LAHC-Study: Measures of correct target hits, error rate, response time and speed.

The participant completed 46 target pools (with 20 seconds per pool) across all Gz loads (1.5*g*, 1.8*g* and 2.5*g*) and accomplished 566 correct target hits in total with a mean of 12.30 (SD = 1.39) per target pool with 42 false target hits with a mean of 0.91 (SD = 1.28) per target pool. Only data sets where the number of false target hits was zero and the response time was less than 3000 ms were considered for analyzing the time that the participant required to respond to a visual stimulus. This resulted in 460 valid trials used for analyzing the response time and the resulted pointing speed. Table 9.1 presents

measures of the central tendency and the variability of the dependent variables for the studied sizing techniques per Gz load. As also shown in Figure 9.5 (left), pointing towards elongated targets (SE) resulted in average mostly correct target hits under the gravity levels 1.5*g* and 1.8*g*, while the compressed sizing technique (SC) led to the lowest number of correct target hits under all Gz loads. But comparing the sizing conditions (see Table 9.2) only revealed significant differences across all Gz loads, with a higher mean frequency of correct hits with SU (M = 12.65, SD = 1.47) and with SE (M = 12.75, SD = 1.14) compared to SC (M = 11.09, SD = 0.70). The same applied to the percentage error rate that revealed the lowest error rate with SE under 1.5*g* and 1.8*g* and always the highest with SC, but did not show significant differences.

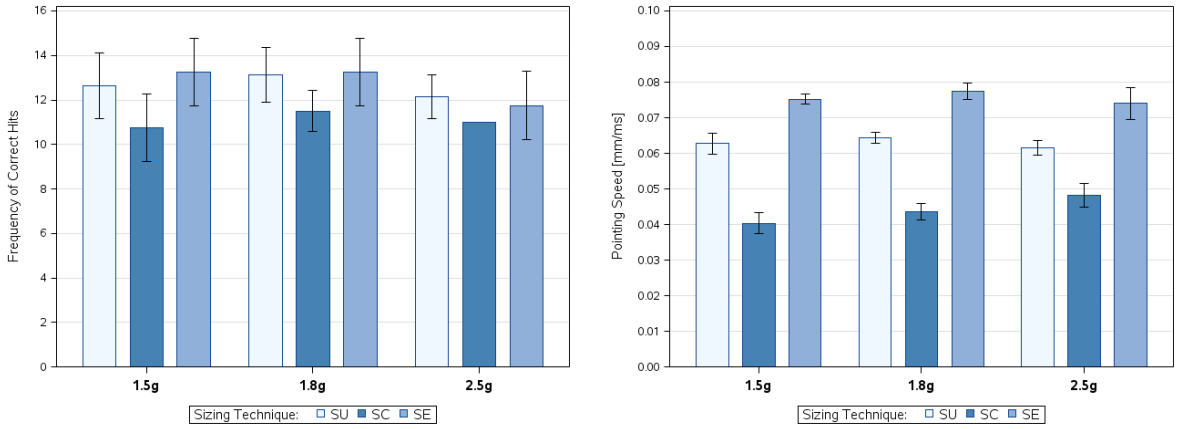


Figure 9.5: LAHC-Study: Mean frequency of correct target hits and pointing speed (with CI=95 %) of the sizing methods grouped by the Gz loads.

Because the elongated method provides larger and the compressed method shorter distances I expected longer response times with the elongated method (SE) and shorter response times with the compressed method (SC). But the performance did not show meaningful variations of the mean response times. Therefore it is more reasonable to analyze the pointing speed, because the size and distance of the target varies with the presented sizing method. The variations of the mean pointing speed (see Figure 9.5, right) showed most prominent variations for the elongated method (SE) with the highest speed and for the compressed method (SC) with the lowest speed under all Gz loads. The comparison of the means (see Table 9.2) resulted in significant differences between the sizing methods on the same level of gravity load and between the sizing methods across all Gz loads with significant highest speed under the SE conditions (M = 0.076 mm/ms, SD = 0.008) followed by the SU condition (M = 0.063 mm/ms, SD = 0.010), and with the slowest speed under the SC condition (M = 0.044 mm/ms, SD = 0.007).

In conclusion, the case study showed that gravity-adapted target resizing and positioning significantly impacts aimed pointing performance under increased Gz loads and showed a tendency to improve the pointing performance using elongated targets (SE), particularly taking into account the significant increase in the pointing speed. That means, it seems that pointing under increased gravity benefits from greater targets at larger distances. Because only one participant could be tested on the long-arm centrifuges, I looked for alternative approaches to simulate hypergravity.

Dependent Variable	+Gz	Resizing		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Correct Target Hits	-	SC	SU	43	-3.43	0.0014	0.0041	*
	-	SC	SE	43	-3.20	0.0026	0.0065	*
Reponse Time [ms]	1.5	SU	SC	166	-3.10	0.0023	0.0058	*
	1.5	SC	SE	166	3.02	0.0029	0.0075	*
	2.5	SU	SC	110	2.99	0.0035	0.0093	*
	2.5	SC	SE	110	-3.33	0.0012	0.0034	*
Pointing Speed [mm/ms]	-	SU	SC	457	17.51	<.0001	<.0001	*
	-	SU	SE	457	-12.97	<.0001	<.0001	*
	-	SC	SE	457	-26.07	<.0001	<.0001	*
	1.5	SU	SC	166	10.64	<.0001	<.0001	*
	1.5	SU	SE	166	-6.98	<.0001	<.0001	*
	1.5	SC	SE	166	-15.19	<.0001	<.0001	*
	1.8	SU	SC	175	13.46	<.0001	<.0001	*
	1.8	SU	SE	175	-9.67	<.0001	<.0001	*
	1.8	SC	SE	175	-19.73	<.0001	<.0001	*
	2.5	SU	SC	110	6.54	<.0001	<.0001	*
	2.5	SU	SE	110	-6.45	<.0001	<.0001	*
	2.5	SC	SE	110	-11.05	<.0001	<.0001	*

Table 9.2: LAHC-Study: Significant differences of performance between the sizing methods on the same level of Gz load using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate) and PROC MIXED (adjust=simulate).

9.1.4 User Study by Arm Weightings *(Weight Study)*

To verify the observed effect of the case study using the LAHC, a subsequent experiment was performed under normogravity condition. For simulation the +Gz loads, corresponding weightings were balanced attached to the participant's dominant forearm (see Figure 9.6). The extended arm weights (see Table 9.3) were calculated (Eq. 6) for each participant as follows:

$$\boldsymbol{m_{add} = (G_{sim} - G) * \frac{m_{body}}{100} * 5.38\,\%} \tag{6}$$

with G_{sim} for the simulated gravity force, m_{body} for the body weight of the participant and 5.38 % as averaged percentage arm weights introduced by Clauser et al. (1969).

Participant	$\boldsymbol{m_{body}}$ [kg]	$\boldsymbol{m_{arm}}$ [kg]	$\boldsymbol{m_{add}}$ [kg]		
			1.5*g*	**2*g***	**2.3*g***
S1	80.0	4.3	2.2	4.3	5.6
S2	78.0	4.2	2.1	4.2	5.5
S3	75.0	4.0	2.0	4.0	5.3
S4	80.0	4.3	2.2	4.3	5.6
S5	65.0	3.5	1.8	3.5	4.6
S6	69.0	3.7	1.9	3.7	4.8
S7	60.0	3.3	1.7	3.3	4.3
S8	78.0	4.2	2.1	4.2	5.5

Table 9.3: Weight-Study: Weights of the body and arm, and the resulted added weights.

Figure 9.6: Weight-Study: Participants wearing the arm weightings to simulate different hypergravity loads.

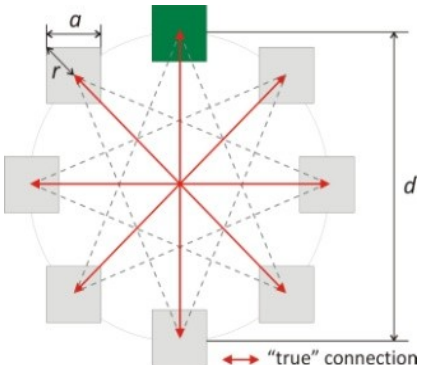
The LAHC-Study has shown that the performance under 2.5g was strongly influenced by physical demands. Therefore I decided to change the used Gz loads, in a way that the user study by arm weightings was performed under 1.5g, 2g and 2.3g. In contrast to the LAHC-Study, I investigated additionally the effect of gravity-adapted resizing on the physiological workload that was assessed by cardiac responses using the heart rate variability (see section 3.4.3, *Physiological Workload*), as it was done in previous studies presented in Chapter 7 and Chapter 8. While in these studies I analyzed the Stress Index (SI) and the LF/HF ratio, this time I used the R-R distance, a parameter from the time domain, which is the interval between two consecutive heartbeats identified by two peaks of an ECG wave. Such R-R interval is also referred to as NN-interval, so that the standard deviation for a R-R interval is also denoted as SDNN, which, for example, was used by Tümler et al. (2008).

9.1.4.1 Apparatus

The same HMD setup as for the LAHC-Study (see section 9.1.3.2) was used. All participants also performed an eye-sensor calibration (Kato and Billinghurst, 1999) immediately before the experiment. To perform the task of pointing towards outside coded targets, the participant stood in front of a wall with 50 cm distance. Depending on participant's body height the multi-pattern was individually aligned in the horizontal position, thus the target area was at the participant's eye level. The optical sensor captured the pose data with a mean frame rate of 38.52 fps (SD = 12.54). To record the HRV, the participant was equipped with a wireless eMotion HRV sensor from Mega Electronics, whereby the electrodes were placed to participant's thorax

9.1.4.2 Experiment Task

To evaluate the speed-accuracy trade-off related to Fitts' law I decided to use an appropriate task and designed a multi-directional pointing task as proposed by the ISO/DIS 9241-9 standard (2000). Therefore eight squared targets with a default size of $a = 15.0\ mm$ (see Figure 9.7) were used. The targets were arranged in a circle with a default diameter of $d = 82.5\ mm$. Like the LAHC task, the participants should point towards the targets in response to visual stimuli. For evaluation purposes by Fitts' law I defined "true" target connections of 0°, 45° and 90° that implied the same target distance and involve vertical, oblique and horizontal arm movements. The remaining target connections were used for pointing transition only.



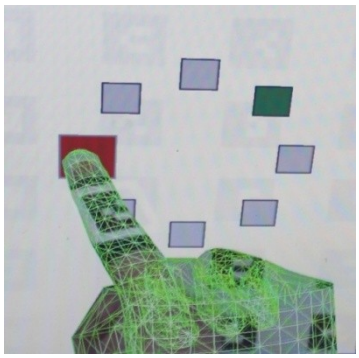


Figure 9.7: The multi-directional pointing task used during the Weight-Study.

9.1.4.3 Participants

Participants were 6 male and 2 female aged between 24 and 51 years (20-31 years: 4 participants, 40-51 years: 4 participants, M = 37.25, SD = 10.55). Seven participants have had experiences with AR interfaces in terms of participation in previous studies, while one participant was a novice. They came from backgrounds in biology, physiology, aerospace and medicine. All participants had a right-dominant arm that was used for the pointing task.

9.1.4.4 Experiment Design and Procedure

The study followed a repeated measure design with two independent variables containing three levels for gravity-based resizing (SU, SC, SE) and three levels for Gz load (1.5*g*, 2*g*, 2.3*g*). Thereby the SU level constituted the baseline condition. In a within-subject design, each participant performed the test series for all resizing methods under all gravity loads, resulting in a factorial design of 3 x 3 (see Figure 9.8). The repetition rate for each method amounted to five target pools per Gz load. Thereby a target pool was specified as a predefined series of randomized target connections for the multi-directional pointing tasks. Pointing towards the targets of one target pool should be completed by the participants in 25 seconds and constituted one test series. Overall each participant performed 45 test series. Because I did not compare the gravity loads, neither on the same level of sizing methods, nor across the sizing methods, the multi-directional task was performed in a fixed order of Gz loads (2.3*g*, 1.5*g*, 2*g*). But I used systematic variations of the presentation order of the sizing methods per Gz load. To avoid transition effects between pool changes the first and the last signaled targets performance were not recorded. Between changes of the Gz load and the sizing techniques, the participant had a five minute break for physiological regeneration. To be familiar with the pointing

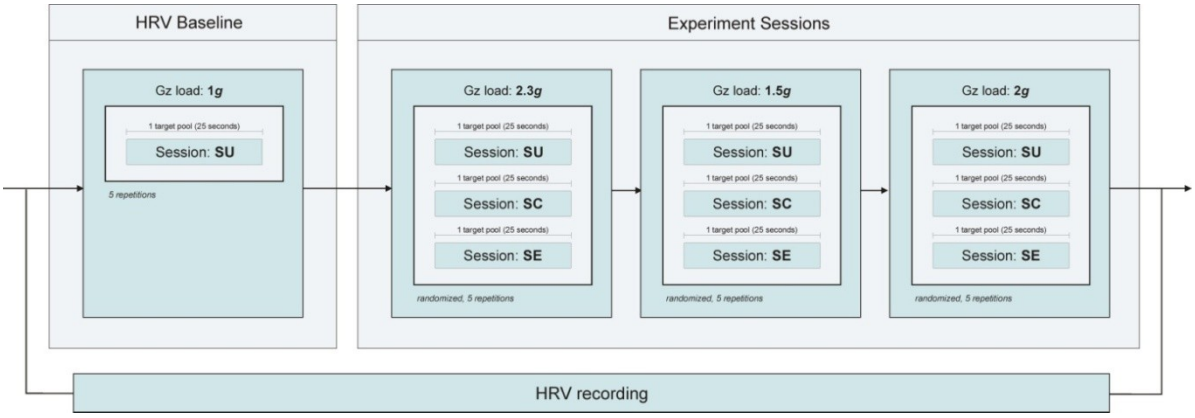


Figure 9.8: The experiment procedure of the Weight-Study that each participant conducted.

task and to check the integrity of the tracking operation, the participants undertook a short training session before starting the first condition. Before conducting the experiment sessions each participant performed the experiment without added arm weightings under the SU condition as baseline for the workload assessment by HRV. This condition was not used for evaluating the performance.

9.1.4.5 Results

Table 9.4 shows the resulted target sizes *a* with its surrounding radius (Eq. 7) and targets' distances *d* calculated by the force-based resizing approach using the active Gz load. The distance reflects the pointing range between two "true" target connections. While pointing towards normal sized targets (SU) always resulted in same target sizes and distances, the elongated sizing (SE) resulted in increased sizes and distances on increased Gz loads and contrary for the compressed sizing (SC).

$$r_s = \frac{a\sqrt{2}}{2} \tag{7}$$

Resizing	+Gz	$\boldsymbol{a}$ [mm]	$\boldsymbol{r_S}$ [mm]	$\boldsymbol{d}$ [mm]
SU	-	15.00	10.61	82.50
SC	1.5	11.67	8.25	64.17
	2.0	10.56	7.45	58.06
	2.3	9.89	6.99	54.39
SE	1.5	18.33	12.96	100.83
	2.0	19.44	13.75	106.94
	2.3	20.11	14.22	110.61

Table 9.4: Weight-Study: Resulted target size *a*, radius r_s and distance *d*.

As for the LAHC-Study the pointing performance was assessed by the frequency of correct target hits, the percentage error rate, the response time and the pointing speed depending on the resulted target distances. To compare the percentage error rate between the sizing conditions, the data values were transformed by the natural logarithm. In addition I evaluated the Euclidean distance between the target's center and the final intersection point and the percentage accuracy depending on the resulted target sizes. The sizing conditions were multiple pairwise compared for each dependent variable on same level of Gz load and across all Gz loads using SAS® PROC MIXED for repeatedly measured data and SAS® PROC GLIMMIX for objective count data (see section 3.4.2, *Data Analyzing),* whereby the significance level ($\alpha = 5\,\%$) was guaranteed by the Simulate method. Appendix I additionally provides descriptive statistics for the outcomes showing measures of the central tendency and variability for the above mentioned dependent variables. Additionally, the speed-accuracy trade off between the sizing methods according to Fitts' law was evaluated. Thus, the throughput (TP) and the resulted movement models predicting the movement times of the studied sizing conditions are presented. For analyzing the physiological workload, the R-R distance per sizing methods per Gz loads is presented and discussed.

9.1.4.5.1 Task Performance

Pointing Frequency

The data revealed that the participants pointed towards 6708 targets in total in a correct way with a mean frequency of correct target hits of 19.39 (SD = 3.37) per target pool and pointed towards 102

targets in a wrong way with a mean frequency of false target hits of 0.27 (SD = 0.66) per target pool. While Figure 9.9 shows the mean frequency of correct target hits, Figure 9.10 presents the mean percentage error rate of the sizing conditions across all Gz loads and per Gz load. With respect to the frequency of correct target hits the comparison of the sizing conditions on same level of Gz load revealed no significant differences, but comparing the sizing conditions over all Gz loads showed the highest frequency of correct target hits for pointing towards elongated targets SE (M = 20.00, SD = 2.53) that significantly differed (see Table 9.5) from SU (M = 19.05, SD = 3.52) and SC (M = 18.82, SD = 4.09). With respect to the log transformed error rate the data only yielded significant differences (see Table 9.5) under 1.5g with a significant lower error rate with SU compared to SC, and with a significant lower error rate under SE compared to SC.

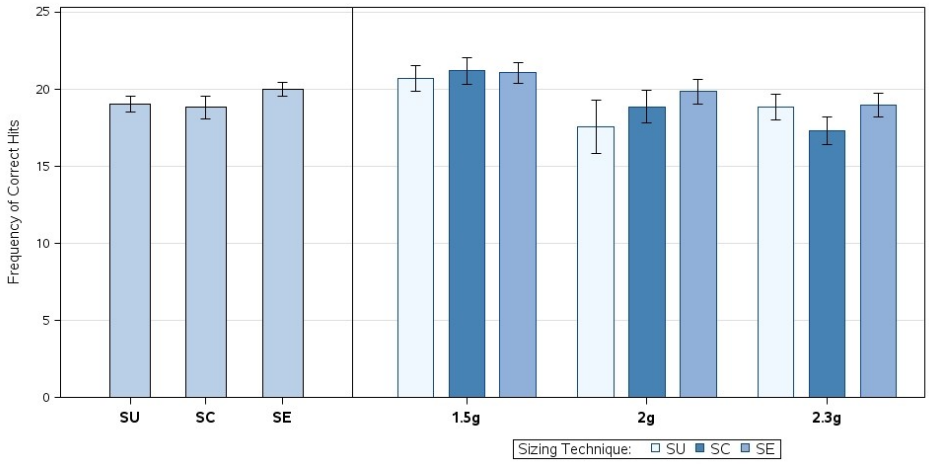


Figure 9.9: Weight-Study: Mean frequency of correct target hits (with CI=95%) of the sizing methods: *(left)* across all Gz loads and *(right)* grouped by Gz load.

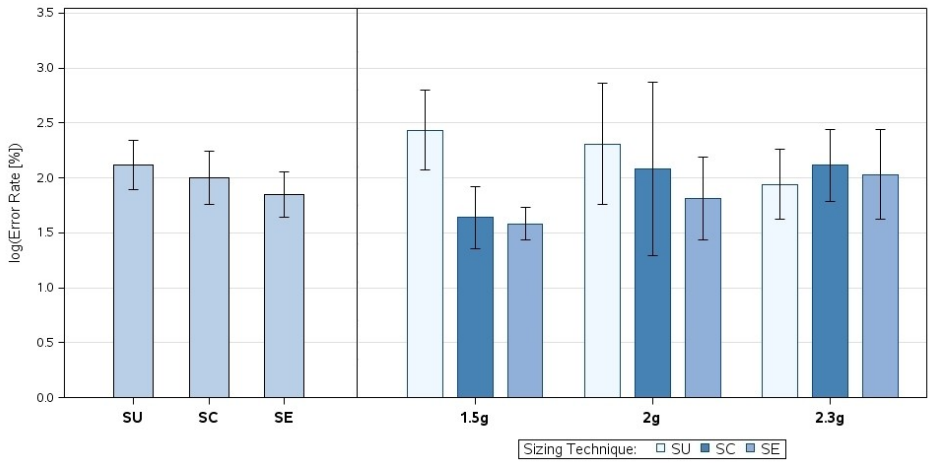


Figure 9.10: Weight-Study: Mean log(error rate[%]) (with CI=95%) of the sizing methods: *(left)* across all Gz loads, and *(right)* grouped by Gz load.

Dependent Variable	+Gz	Resizing		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Correct Target Hits	-	SE	SU	338	2.47	0.0140	0.0374	*
	-	SC	SE	338	-2.80	0.0053	0.0140	*
log(Error Rate [%])	1.5	SU	SC	10	5.50	0.0003	0.0007	*
	1.5	SU	SE	10	6.03	0.0001	0.0003	*

Table 9.5: Weight-Study: Significant differences of correct target hits and error rate between the sizing methods across all Gz loads and on same level of Gz load using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate) and PROC MIXED (adjust=simulate).

Response Time and Pointing Speed

For analyzing time effects I only considered target hits with "true" target connections resulting in the same pointing distance per Gz load and sizing technique, as well as data sets where the number of false target hits was zero and the response time was less than 3000 ms. This resulted in 2055 valid trials. Overall the participants pointed with a mean response time of 1038.60 ms (SD = 179.26). The mean response times of the sizing conditions across all Gz loads and per Gz load are presented in Figure 9.11 and show that in principle the response times increased as the gravity load increased. But the mean response times did not show prominent variations between the sizing conditions across all Gz loads and per Gz load. This is contrary to my expectation of significant slower response times at larger target distances (SE) and significant faster response times at shorter distances (SC). Because targets' size and distance vary with the used sizing technique and the Gz load, analyzing the pointing speed was more meaningful than the response time. The pointing speed was calculated by the distance between the targets (see Table 9.4) divided by the response time. Overall the participants pointed with a mean speed of 0.083 mm/ms (SD = 0.024). The mean pointing speed by the sizing techniques across all gravity loads and per Gz load is presented in Figure 9.12. Prominent mean variations yielded from the elongated sizing technique (SE) with highest speed under all Gz loads, while using compressed targets (SC) always revealed lowest speeds. Comparing the sizing technique (see Table 9.6), yielded significant differences ($p < .0001$) between all conditions across all Gz loads and grouped per Gz load. The result related to the pointing speed mean that hypothesis H1 can be accepted.

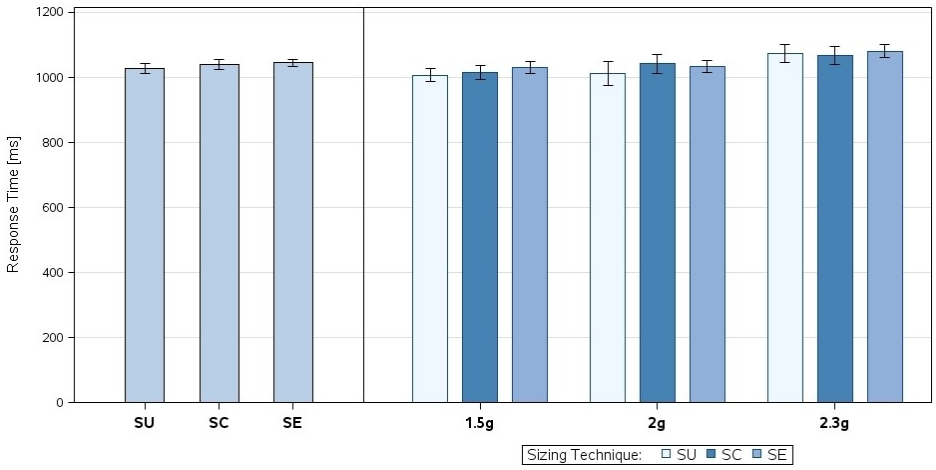


Figure 9.11: Weight-Study: Mean response time to visual stimuli (with CI=95%) of the sizing methods: *(left)* across all Gz loads and *(right)* grouped by Gz load.

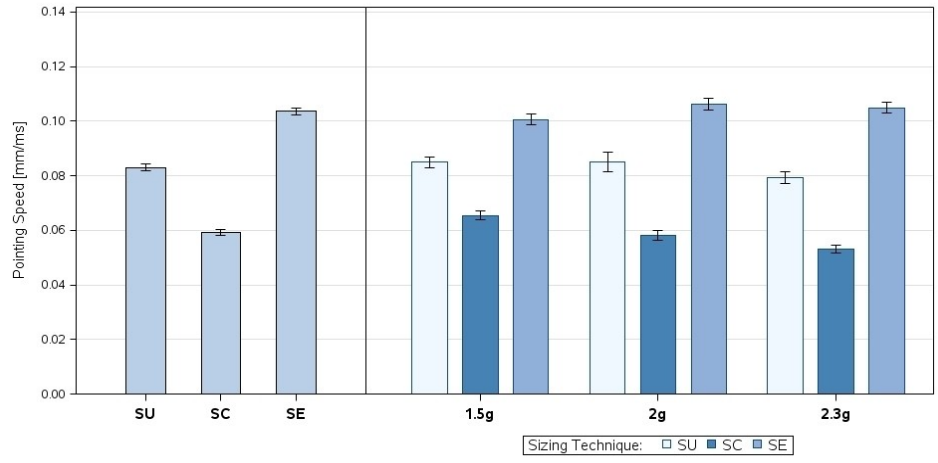


Figure 9.12: Weight-Study: Mean pointing speed (with CI=95%) of the sizing methods: *(left)* across all Gz loads and *(right)* grouped by Gz load.

Dependent Variable	+Gz	Resizing		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
Pointing Speed [mm/ms]	-	SU	SC	2046	26.61	<.0001	<.0001	*
	-	SU	SE	2046	-23.77	<.0001	<.0001	*
	-	SC	SE	2046	-55.48	<.0001	<.0001	*
	1.5	SU	SC	2046	14.90	<.0001	<.0001	*
	1.5	SU	SE	2046	-12.78	<.0001	<.0001	*
	1.5	SC	SE	2046	-27.67	<.0001	<.0001	*
	2.0	SU	SC	2046	14.72	<.0001	<.0001	*
	2.0	SU	SE	2046	-12.03	<.0001	<.0001	*
	2.0	SC	SE	2046	-32.00	<.0001	<.0001	*
	2.3	SU	SC	2046	16.96	<.0001	<.0001	*
	2.3	SU	SE	2046	-16.82	<.0001	<.0001	*
	2.3	SC	SE	2046	-36.17	<.0001	<.0001	*

Table 9.6: Weight-Study: Significant differences of the pointing speed between the sizing methods across all Gz loads and on same level of Gz load using SAS® PROC GLIMMIX (dist=poisson, link=log, adjust=simulate) and PROC MIXED (adjust=simulate).

Pointing Accuracy

The pointing accuracy reflects the precision of target pointing and was measured by the Euclidean distance d_{ED} and the surrounding radius r_s of the targets (see Table 9.4). For evaluating the pointing accuracy I initially analyzed the Euclidean distance relative to the resulted target sizes using 6708 correct target hits. Thereby the Euclidean distance d_{ED} was the distance between the centre of the target and the intersection point within the target. As presented in Figure 9.13, the variations of the mean distances show a proportional ratio between the distance and target's size, i.e. the pointing distance was greater with the increment of target's size and vice versa. Statistical analyzing of the Euclidean distance confirmed this observation by significant differences ($p < .0001$) between the sizing techniques across all Gz loads and per Gz load (see Table 9.7). The test revealed that pointing towards SC targets resulted in the significantly shortest distances and towards SE targets in significant largest pointing distances.

Contrary to the Euclidean distance, the percentage accuracy (Eq. 8) mirrors the percentage ratio of the distance d_{ED} to the target size expressed by the radius r_s.

$$accuracy = 100 - \left(\frac{d_{ED}}{r_s} * 100\right) \tag{8}$$

The participants pointed with an overall mean percentage accuracy of 56.4 %. The variations of the mean accuracy are presented in Figure 9.14. Per sizing technique over all Gz loads data revealed that relative to target's size, participants pointed most precisely using the elongated method (SE) with 58.7 % accuracy (SU: 56.7 %, SC: 53.2 %). Statistical analyzing (see Table 9.7) revealed that pointing towards elongated targets (SE) enabled a significant improvement ($p < .0001$) over all Gz loads compared to pointing towards compressed targets (SC). On the same Gz level, SE yielded significant increased accuracy under 2*g* than SE, and under 2.3*g* than SC and SU.

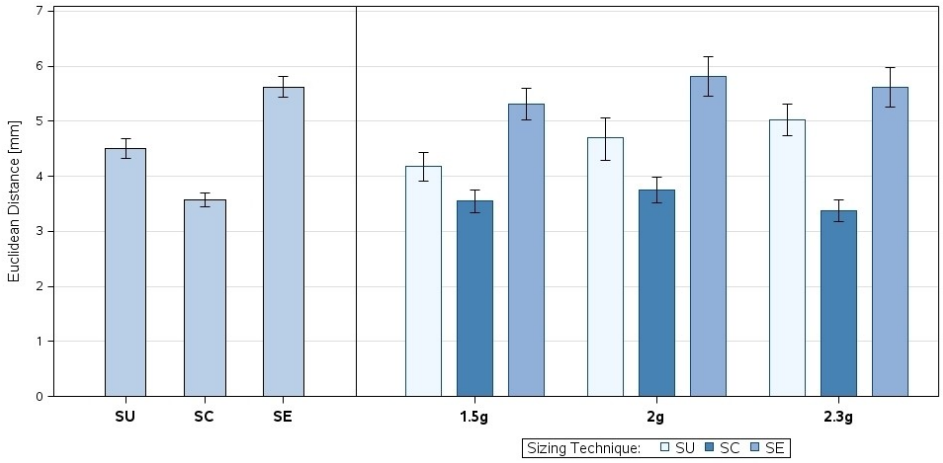


Figure 9.13: Weight-Study: Mean Euclidean distance (with CI=95%) of the sizing methods: *(left)* across all Gz loads and *(right)* grouped by Gz load.

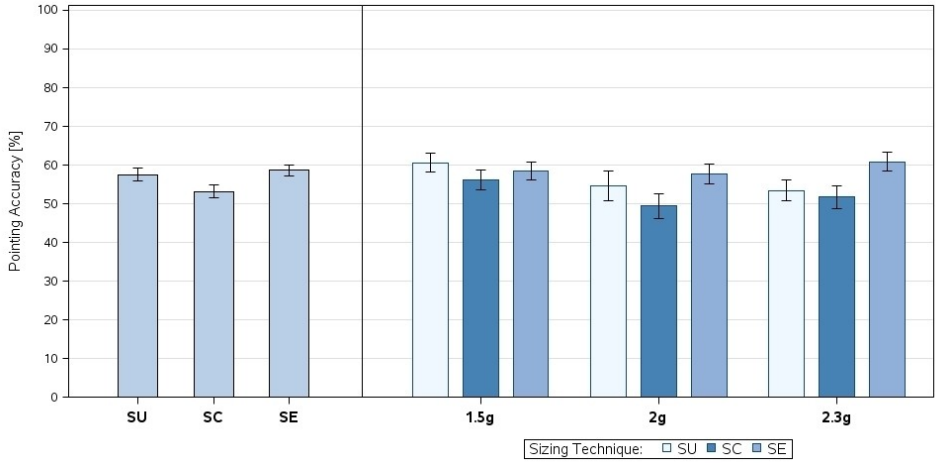


Figure 9.14: Weight-Study: Mean pointing accuracy (with CI=95%) of the sizing methods: *(left)* across all Gz loads and *(right)* grouped by Gz load.

Dependent Variable	+Gz	Resizing		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
log(accuracy [%])	-	SC	SE	6705	5.25	<.0001	<.0001	*
	2.0	SC	SE	1862	-6.13	0.0059	0.0178	*
	2.3	SU	SE	2152	-4.22	0.0024	0.0072	*
	2.3	SC	SE	2152	-8.95	0.0145	0.0436	*
Euclidean Distance [mm]	-	SU	SC	6705	16.73	<.0001	<.0001	*
	-	SU	SE	6705	-15.07	<.0001	<.0001	*
	-	SC	SE	6705	-33.52	<.0001	<.0001	*
	1.5	SU	SC	2685	8.83	<.0001	<.0001	*
	1.5	SU	SE	2685	-8.68	<.0001	<.0001	*
	1.5	SC	SE	2685	-17.84	<.0001	<.0001	*
	2.0	SU	SC	1862	8.69	<.0001	<.0001	*
	2.0	SU	SE	1862	-7.78	<.0001	<.0001	*
	2.0	SC	SE	1862	-18.36	<.0001	<.0001	*
	2.3	SU	SC	2152	11.79	<.0001	<.0001	*
	2.3	SU	SE	2152	-8.88	<.0001	<.0001	*
	2.3	SC	SE	2152	-21.77	<.0001	<.0001	*

Table 9.7: Weight-Study: Significant differences of the log(accuracy) and Euclidean distance between the sizing methods across all Gz loads and on the same level of Gz load using SAS® PROC MIXED (adjust=simulate).

9.1.4.5.2 Speed-Accuracy-Trade-Off (Fitts' Law)

In designing Human-Computer-Interfaces the assessment of ergonomics is mainly determined by Fitts' model of movement time (Eq. 7) (Fitts, 1954) that a human needs to point at a target of a given size and distance. Fitts' law predicts longer movement times at larger distances, as well as at smaller targets. The sizing approach interrelates these characteristics to each other, whereby the elongated method (SE) provides larger targets at larger distances, while the compressed method (SC) results in smaller targets at smaller distances. I used Fitts' law to evaluate the speed-accuracy trade-off of the studied sizing techniques related to direct pointing affected by added arm weightings. The metric for comparing the performance is the Throughput *TP* (Eq. 10), in bits per second (bps) calculated by the Index of Difficulty *ID* and mean movement time *MT* (Eq. 9) as time to hit a target (in milliseconds) with *a* for the intercept and *b* for the slope of measured mean response time by the target width *W*. The *ID* measures the tasks difficulty in bits using target size and distance. Because squared targets were used, the *ID* was calculated only by the targets' width. For computing the *ID* (Eq. 11) I used the Welford formulation (Welford, 1960). To reflect the observed pointing performance of the participants, the effective target width W_e (Eq. 12) (MacKenzie, 1992; Welford, 1960) as the central 96 % of the spatial distribution with SD_x as standard deviation of the mean pointing accuracy was used.

$$MT = a + b\, ID_e \tag{9}$$

$$TP = \frac{ID_e}{MT} \tag{10}$$

$$ID_e = \log_2\left(\frac{A}{W_e} + 0.5\right) \tag{11}$$

$$W_e = 4.133 * SD_x \tag{12}$$

Table 9.8 shows the resulting Fitts' parameter for the three sizing methods per +Gz load. Because the target size and distance increase with an increase in gravity, the SE method resulted in the most difficult targets with the highest ID_e under 2*g* and 2.3*g*, but also in the highest throughput (*TP*). The compressed sizing method (SC) yielded the highest index of difficulty under 1.5*g*, while under 2*g* and 2.3*g* yielded the most simple targets, but also the lowest *TP*. Pointing towards normal sized targets (SU) yielded an increased ID_e, as well as a growing throughput with the increment of gravity.

Resizing	+Gz	**A** [mm]	**W** [mm]	**W_e** [mm]	**MT** [ms]	**ID_e** [bits]	**TP** [bps]
SU	1.5	82.50	15.00	14.47	1010.61	**2.63**	**2.61**
	2.0	82.50	15.00	11.04	1052.83	**2.99**	**2.85**
	2.3	82.50	15.00	11.78	1071.96	**2.91**	**2.71**
SC	1.5	64.17	11.67	8.47	1021.53	**3.02**	**2.96**
	2.0	58.06	10.56	11.12	1032.42	**2.52**	**2.44**
	2.3	54.39	9.89	9.09	1085.23	**2.69**	**2.49**
SE	1.5	100.83	18.33	13.68	1029.36	**2.98**	**2.89**
	2.0	106.94	19.44	13.39	1035.30	**3.09**	**2.98**
	2.3	110.61	20.11	13.18	1089.69	**3.15**	**2.89**

Table 9.8: Weight-Study: The Fitts' resulted parameters: targets' distance (A), target width (W), effective target width (W_e), mean measured movement time (MT), effective Index of Difficulty (ID_e), and Throughput (TP).

The resulting Pearson's correlation coefficient *r* and the regression equations of Fitts' movements' model for the sizing conditions are presented in Table 9.9. While the movement time and the index of difficulty were very strongly positively correlated ($r > 8.0$) for the SU and SE conditions, a moderate

negative correlation was revealed by the SC condition. The Fitts' model of movement time of the SU (Eq. 13) and SE (Eq. 15) sizing conditions provides good descriptions of the observed pointing behaviour and predicted an increase in movement time with an increased target difficulty with the fastest increase under the SE condition. In contrast, the model of the compressed sizing technique SC (Eq. 14) resulted in a model with the highest intercept and lowest negative slope, i.e. that the movement time will slow decrease with an increase in the Index of Difficulty.

Resizing	**r**	**Fitts' model of movement time** (ms)	
SU	0.856	$MT = 638 + 142\ ID_e$	(13)
SC	-0.321	$MT = 1163 - 43\ ID_e$	(14)
SE	0.842	$MT = 81 + 315\ ID_e$	(15)

Table 9.9: Weight-Study: Pearson's correlation coefficient *r* between *MT* and ID_e and linear regression equation of Fitts' model of *MT* per sizing method.

9.1.4.5.3 Physiological Workload

Analyzing the R-R distance obtained by HRV shows the impact to the cardiovascular system on a certain workload. Larger workload causes a larger impact in the cardiovascular system and therefore causes a higher heart frequency and subsequently a shorter R-R interval between the heartbeats. The cardiovascular parameters were assessed during all phases of the experiment. The 1*g* SU output was used as reference measurement and showed the lowest impact on the cardiovascular system. Since the physiological workload respectively grows with the increment of gravity, the R-R distance decreased during the experiment under 2*g* and even more under 2.3*g* (see Table 9.10).

Gz	**SU**	**SC**	**SE**
	Mean ± SD	Mean ± SD	Mean ± SD
1.0	723.86 ± 156.88	-	-
1.5	674.24 ± 114.76	**680.45** ±120.29	665.09 ± 102.56
2.0	642.11 ± 119.43	**649.57** ±117.03	625.28 ± 88.07
2.3	641.97 ± 105.77	**648.31** ±119.25	645.24 ± 126.40

Table 9.10: Weight-Study: The mean R-R distance (HRV) per sizing method across all participants.

The data revealed the lowest R-R for all sizing conditions under high Gz load since the weight attached to the participants' arm constituted the major part of the workload. While the SE conditions yielded the lowest values for the R-R distances under all Gz loads, the SC condition always resulted in the highest values for the R-R distance. This suggests that pointing towards elongated targets (SE, largest target distances) mostly increased the workload, while pointing towards compressed targets (SC, shortest target distances) decreased the workload the most. Statistical analysis showed a significant main effect of the Gz loads on the R-R distance ($F_{2,4} = 27.69$, $p < .05$), but did not show effects by the sizing methods. This result indicates that hypothesis H2 can be conditionally accepted, i.e. that compressed resizing (SC) decreases the most the workload under increased gravity load, although the difference to the other sizing conditions (SE, SU) was not statistically significant.

9.1.4.6 Discussion

The results of the LAHC-Study has already shown that pointing towards outside coded targets during AR selection is affected by gravity-adapted resizing under increased Gz loads and revealed an upcoming trend for the elongated sizing method (SE) by a significant increase in the pointing speed. The Weight-Study confirmed the findings of the LAHC-Study and also showed that the workload during direct AR selection towards outside coded targets is influenced by gravity-adapted resizing with significant differences between affected targets (SE, SC) and unaffected targets (SU). In this way I can affirm that gravity-adapted target resizing affects the performance and workload of direct AR pointing under altered hypergravity conditions (Q1). The results of the Weight-Study also confirmed the observed trend of an improved performance of pointing towards targets that are influenced by elongated resizing (SE) implicating greater target sizes and larger distances between the targets. An elongated target interface not only significantly increases the pointing frequency across all Gz loads; it also shows a decreasing tendency in the error rate. In contrast to elongated targets, the compressed sizing technique (SC) yields the small targets at short distances that always caused significant closer hits to the targets' centre. But relative to the target size, pointing towards elongated targets (SE) significantly provided the most precise pointing across all Gz loads than pointing towards compressed targets (SC), and even under 2.3*g* it revealed significant more accurate pointing than SE and the unchanged baseline condition (SU). With respect to the response time I was astonished that the resulted target distances did not affects the time that was needed to respond to visual stimuli. I expected the shortest times for pointing towards compressed targets with shortest distances, and the longest times for pointing towards elongated targets with the longest distances. But considering the distance covered yielded the significantly fastest pointing speed by the longest distances using the elongated sizing method (SE). Also the analysis of the speed-accuracy trade-off related to Fitts' law showed that pointing towards elongated targets (SE) yielded in an increase in throughput with increased gravity load. The HRV based workload assessment showed an effect caused by changed gravity with attached arm weights and the alternation of the workload. The mostly increased workload was caused by pointing towards elongated targets (SE) that could be due to the distance that gets larger with the increase in gravity load. In contrast, the mostly decreased workload was provided by pointing towards the compressed targets with the shortest distances.

9.1.5 Conclusion and Summary

The performance of Augmented Reality direct object selection coded outside of the human body frame of reference is impaired under short-term altered gravity. Therefore I looked for adequate countermeasures and introduced a gravity-adapted resizing approach that dynamically modify the size and position of the pointing targets related to the active gravity load. Before conducting experimentations under microgravity conditions, I started studying how gravity-adapted resizing affects visuomotor coordination under altered hypergravity conditions. Applying the resizing approach can affect the pointing interface in two different ways that either results in elongated targets or in compressed targets. This means that with an increase in gravity elongated interfaces will increase in targets' size and distance, while compressed interfaces will decrease in sizes and distances. I conducted a proof-of-concept study under simulated hypergravity conditions to investigate the influence of gravity-adapted resizing on the performance and workload of visuomotor coordination during AR selection towards outside coded interfaces. The study was divided into two experiments, where simulated hypergravity was induced, firstly by long-arm human centrifugation (LAHC) and then by added arm weightings under normogravity. The workload was assessed by cardiac response

using the heart rate variability (HRV). Summarizing the results of both studies showed that AR selection towards outside coded targets is effected by gravity-adapted resizing and proofed evidence that pointing towards outside coded target interfaces under increased gravity benefits from elongated resizing, while compressed resizing deteriorates the pointing performance. But contrary, pointing towards elongated targets increases the physiological workload, while pointing towards compressed targets provides a decrease in workload. In future work, the effect of gravity-adapted resizing on performance and workload under microgravity conditions should be investigated. Therefore the resizing approach needs to be correspondingly adjusted to the changed condition. Further research could also investigate the effect of gravity-adapted resizing on the view management of an AR supported assistant system for space operation. It is conceivable that analogous to aimed pointing movements the gravity-adapted resizing approach affects the gaze control during the detection of labels and annotations. Related to the workload assessment by HRV, future work should consider the separation of physical and cognitive workload. This can be done by additional evaluation of the muscular activity measured by electromyogram (EMG). By doing so, it could even more precisely assess the workload during pointing and targeting.

9.2 Changing Isometric Force for Nonhaptic AR Interfaces *(Haptic Study)*

Interfaces floating in the air in front of the user, always visible and accessible are incredible (see Figure 9.15), unless the user wants to interact with them and realizes that operating such interfaces is demanding and stressful, because they do not support tactile sensation provided by haptic cues. As I have shown in the first parabolic flight experiment, presented in the last chapter (see section 8.3), aimed pointing towards nonhaptic interfaces significantly slows down the visuomotor performance not only under altered gravity conditions, but also under Earth conditions, which confirms the finding from Lindeman et al. (1999). Whether the missing physical short-term regeneration or the repeated localization of the virtual interface plane, both make their own contribution to the fact that nonhaptic interfaces weaken the arm position sense, implying a lack of proprioceptive restoration, which is normally supported by tactile sensation. Because their application would be so significant, this shortcoming prompted me to reflect on possible countermeasures to make such interfaces more operable. Driven by microgravity research on visuomotor limitations (see section 2.4.4), one possible

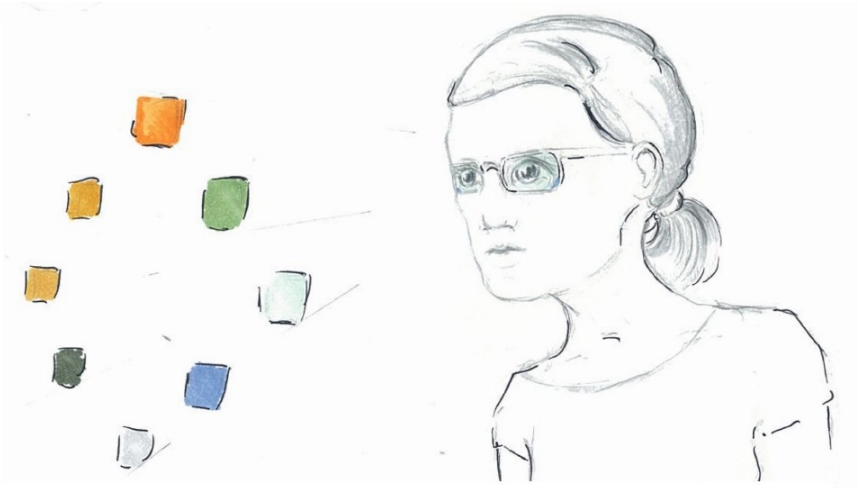


Figure 9.15: Head-referenced interfaces are not providing haptic cues, but are always easy accessible by the alignment to the HMD. (Artwork by Nina-Louisa Giebel).

approach is based on the finding that the stimulation of isometric force boosts the proprioceptive sense of aimed pointing movements and thus can compensate the lack of gravitational force (Bringoux et al., 2012). Thereby the associated motor planning was tuned by an increased sensorimotoric load, assigned to the pointing arm that supported an optimal activation and updating of the internal kinematic model. Adopting such an approach to normogravity conditions would mean that additional sensorimotoric load consequently would exaggerate the proprioceptive force level given by the nature of gravity. While in the previous research increased sensorimotoric load was used to change the characteristics of the interface respectively (see section 9.1), this time the interface remained unchanged, while increased sensorimotoric load was used to stimulate exaggerated force feeback on the pointing arm to counteract the lack of tactile sensation.

9.2.1 Research Objective

Research in user interaction normally uses isometric force to design isometric input devices, for example, by applying this information to operate a virtual spring (Lecuyer et al., 2000) or for multi-finger grasping in virtual environments (Kurillo et al., 2007). This means that such approaches generally aim to map the produced force information on characteristics of the virtual environment so that virtual objects dynamically correspond to the force. In contrast, in this experiment I increased the sensorimotic load of the pointing arm to generate exaggerated isometric force, similar it was done so in the previous Weight-study (see section 9.1.4), but not applying this information to the interface properties. By alterations of the applied force, I investigated whether the target selection can be improved for nonhaptic interfaces, aiming to answer the following research questions:

- **Q1:** Does increased sensorimotoric load speed up pointing movements for nonhaptic interfaces?

If possible effects on improved visuomotor performance could be observed:

- **Q2:** What threshold level of applied force contributes to fastest performance?
- **Q3:** Is the angle of inclination to the gravity vector differently affected?
- **Q4:** Are the findings exclusively applicable to nonhaptic interfaces?

9.2.2 Methodology

The same method, experiment task and apparatus like in the previous study using arm weightings was applied (see section 9.1.4), while the experimental design and procedure were different as it will be described in the next subsection. The nonhaptic interface was realized in the same way as the "HA" condition during the previous parabolic flight study (see section 8.1), whereby the interface was referenced to the user's head, or more exactly, aligned to the field of view of the HMD, worn by the user. As artistically represented in Figure 9.16, the individual weights were balanced attached to the pointing arm to generate the isometric force. To alternate the attached force I decided to use the following Gz loads: 1.3*g*, 1.5*g* and 1.8*g*. I still use the term Gz load instead of increased sensorimotoric load. This is not really correct, but simplifies the repeated usage. The extended arm weights were calculated in accordance with Equation 6, presented in section 9.1.4. The resulted weights that were attached to the participant's pointing arm are listed in Appendix I. Using no weight constituted the baseline condition and was designated as 1*g* level. By applying the multi-directional tapping task (see section 9.1.4.2), the participants were requested to conduct aimed pointing towards

eight squared buttons in response to randomized visual stimuli. Thereby the participants performed the experiment task in an upright posture standing in front of a wall. The task performance was analyzed by using only "true" target connections (see Figure 9.7), which enabled to assess the pointing response time and accuracy with respect to different inclination angles of the gravity vector and related pointing directions. The remaining target connections were used for pointing transition only. In order to answer the question whether the findings are exclusively applicable to nonhaptic interfaces (Q3), a control condition was used. This interface condition should also be operable with one hand, but should provide haptic cues. Thus, the interface without haptic cues was provided by an inside-coded head reference and is referred to as the "nHC" condition, while the interface supporting haptic cues corresponded to an outside-coded world reference and is also called the "HC" condition. This distinction also offered the opportunity to replicate the differences between the HA and PA condition of the first parabolic flight study (see section 8.3.4.1), confirming that nHC interface (HA condition) slows down the performance under normo- and hypergravity, while the pointing accuracy suffered with the HC interface (PA condition), but not assignable to a gravity load. To apply the interface providing haptic cues, the participant was requested to use the wall for tactile sensation, while for the nonhaptic interface condition the participant had to move back a step.

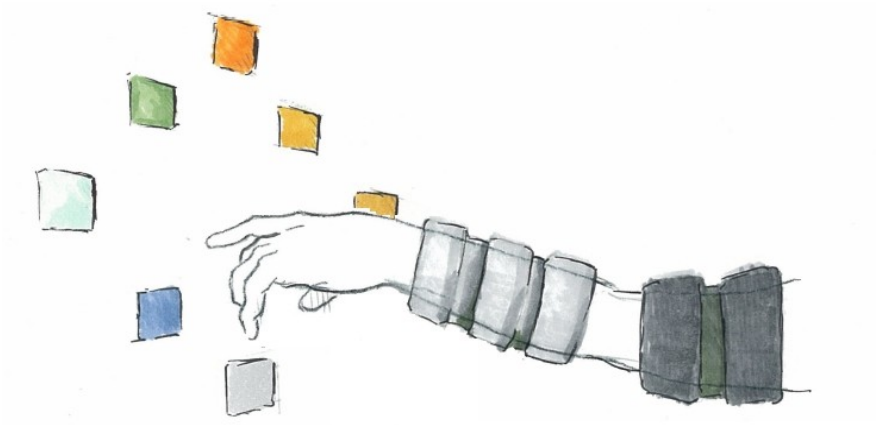


Figure 9.16: Weights fixed to the arm generating isometric force to improve aimed pointing towards nonhaptic interfaces (Artwork by Nina-Louisa Giebel).

9.2.2.1 Experiment Design & Procedure

The study consisted of two independent variables containing two levels for the haptic-feedback method (HC, nHC) and four levels for the Gz load (1*g*, 1.3*g*, 1.5*g*, 1.8*g*) with the 1*g* level as baseline condition. In a within-subject design, 29 participants performed the visuomotor task for all haptic-feedback methods under all gravity loads, resulting in a factorial design of 2 x 4 (see Figure 9.17). The repetition rate for each method amounted to three target pools per Gz load. Thereby a target pool was specified as a predefined series of 20 randomized target connections for the multi-directional pointing tasks. Thus, each participant pointed towards 60 targets per method and Gz load that resulted in 120 targets per Gz load and in 480 in total. Thereby the target pools and the interface conditions were randomized presented. Contrary to the previous study presented in this chapter, the Gz loads were also randomized. For physiological regeneration, the participant had a five minute break between changes of the Gz load and the interface condition. To be familiar with the experiment task and to check the integrity of the tracking operation, the participants undertook a short training session before starting the first experiment session.

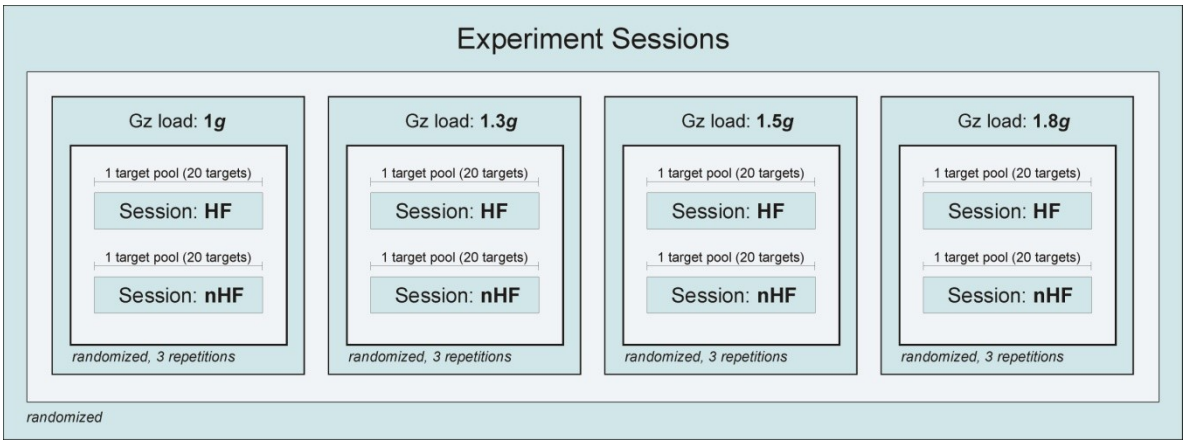


Figure 9.17: The experiment procedure of the Haptic-Study that each participant conducted.

9.2.2.2 Participants

The experiment was performed by 29 participants of whom 20 were male and 9 were female aged from 20 to 57 years (20-29 years: 16 participants; 30-39 years: 4 participants; 40-49 years: 4 participants; 50-59 years: 5 participants; M = 32.66, SD = 12.15). Thirteen participants had experiences with AR interfaces in terms of participation in previous studies, while sixteen participants were novice in dealing with AR. They came from backgrounds in biology, physiology, aerospace, sports science, medicine and computer science. All participants had a right-dominant arm that was used for the pointing task.

9.2.3 Results

To analyze the task performance outliers were removed as a function of the pointing response time, whereby 3000 ms should not be exceeded for responding to a stimulus onset. Based on observations, it can be stated that such long response times were caused by participants' physical limitations and incorrect fitting of the weights. As mentioned above, only "true" target connections were used to analyze the visuomotor performance. Thus, it was additionally possible to differentiate between horizontal (90°), oblique (45°) and vertical (0°) pointing movements with respect to the direction of the force of gravity, as can be materialized by a plumb-line as vertical reference. This resulted in 5981 valid trials used for evaluation purposes. To investigate the effect of exaggerated isometric force on visuomotor performance, the pointing response time was analyzed by comparing the baseline condition (1g) with the isometric force levels (1.3g, 1.5g, 1.8g) at the same stage of interface method. In contrast, comparing both interface conditions on same force level enabled to study the effect of haptic feedback. All multiple pairwise comparisons were performed by SAS® PROC MIXED with Simulate adjustment to keep the experimentwise error rate α of 5 %. To compare the isometric force levels, the pairwise comparisons were restricted by the PDIFF option, which requested the differences of the LS-means to the 1g control only. Appendix I additionally provides descriptive statistics for the outcomes showing measures of the central tendency and variability for the response times.

Response Time

The response time was the time elapsed between the visual triggered stimulus and motoric response onset. The participants pointed with a mean time of 1504.70 ms (SD = 447.22) under the nHC

condition, and with a mean of 1234.98 ms (SD = 333.16) under the HC condition. Figure 9.18 shows the variations of the mean per Gz load for the nHC interface (left) and the HC interface (right). It can be seen that both interface conditions revealed fastest responses under 1.3*g* (**HC:** M = 1217.40 ms; **nHC:** M = 1454.27 ms). It also shows that a further increase in the Gz load caused also an increase in response times. But it is astonishing that the baseline condition (1*g*), using no weights, yielded generally slowest pointing under the nHC condition (**1*g*:** M = 1530.22 ms; **1.5*g*:** M = 1516.25 ms; **1.8*g*:** M = 1517.42 ms), while under the HC condition pointing under 1*g* (M = 1230.60 ms) was still faster than under 1.5*g* (M = 1232.86 ms) and 1.8*g* (1259.09 ms).

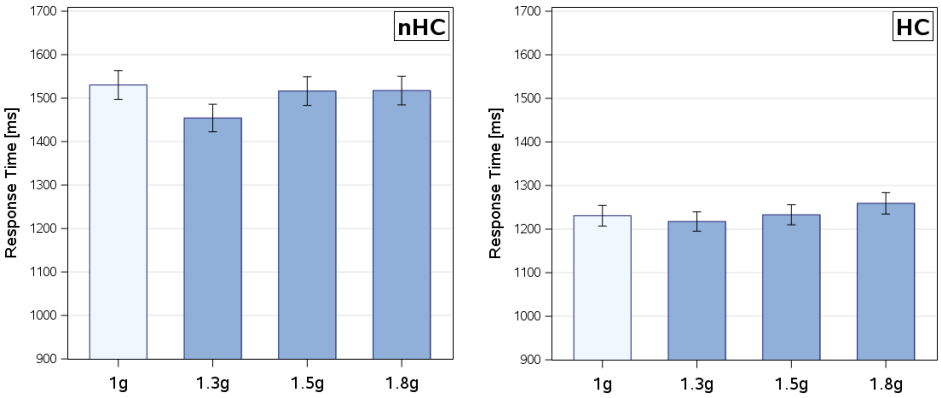


Figure 9.18: Haptic-Study - Mean response time (with CI=95%) per Gz load: *(left)* for the nonhaptic interface (nHC), and *(right)* for the haptic interface (HC).

However, the results of statistical analysis revealed significant difference only for the non-haptic interface (nHC) between the 1*g* and 1.3*g* level, while the response times resulted from the haptic interface (HC) did not significantly differ between the Gz loads (see Table 9.11). In contrast, comparing the haptic-feedback methods at the same level of Gz load verified the previous finding from the experimentation under parabolic flight conditions, showing that the support of haptic cues contributed to significantly faster pointing than using an interface not providing tactile sensation, regardless of the attached arm load (see Table 9.12).

Because the isometric force level is exaggerating the gravitational force given on earth's surface, the consideration of specific levels of the inclination angles to the gravity vector enabled to investigate how the increased sensorimotoric loads affects horizontal (90°), oblique (45°) and vertical (0°) pointing movements in response times respectively. Figure 9.19 shows the related mean value plots for both interface conditions. Based on visual inspection of the HC performance (Figure 9.19, right) it can be noted that the response time under increased Gz loads suffered most from horizontal movements under increased Gz loads, while pointing along the gravity vector (vertical) resulted in fastest movements under 1.3*g* and 1.5*g*. It can also be stated that the performances under 1*g* and 1.8*g* Gz were most stable , while the performance under 1.3*g* and 1.5*g* slowed down more and more with an increasing angle of inclination. Nevertheless, statistical analysis did not reveal significant differences between the Gz loads per inclination angle under the HC condition. In contrast, visual inspection of the performance under the nHC condition (Figure 9.19, left) showed that the performance, using no weights (1*g*), slowed down with a decreasing angle of inclination and suffered most from vertical movements. Quite the opposite can be applied to the 1.3*g* load. Thereby the performance slowed down with increasing angles of inclination and induced generally faster performance along the gravity vector, which significantly differed from pointing under 1*g* (see Table 9.11). As for the HC condition, the performance under 1.8*g* was most stable.

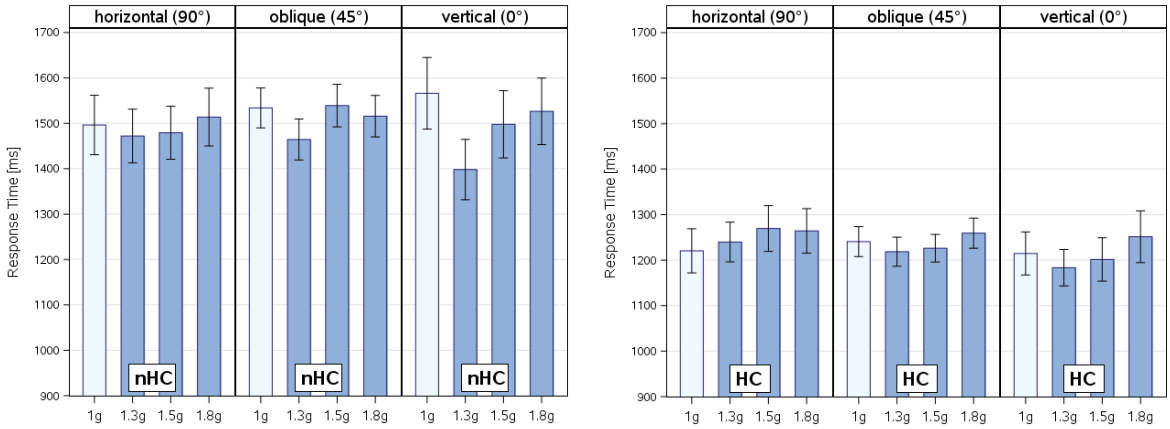


Figure 9.19: Haptic-Study - Mean response time (with CI=95%) of the interface methods paneled by the inclination angles to the gravity vector: *(left)* for the nHC interface, and *(right)* for the HC interface.

Feedback	Gz load		Inclination angle	*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
nHC	1.3*g*	1g	-	2867	-3.85	0.0012	0.0035	*
nHC	1.5*g*	1g	-	2867	0.53	0.5509	0.8827	
nHC	1.8*g*	1g	-	2867	0.18	0.5869	0.9076	
nHC	1.3*g*	1g	vertical (0°)	2859	-3.07	0.0022	0.0167	*
HC	1.3*g*	1g	-	3106	-0.78	0.4339	0.7739	
HC	1.5*g*	1g	-	3106	0.13	0.8937	0.9983	
HC	1.8*g*	1g	-	3106	1.69	0.0915	0.2185	

Table 9.11: Haptic-Study - Comparison of the response time between 1*g* and increased Gz load at the same stage of interface method using SAS® PROC MIXED (pdiff=control('1g'), adjust=simulate).

Gz load	Feedback		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*	
-	nHC	HC	5979	26.58	<.0001	<.0001	*
1*g*	nHC	HC	5973	14.89	<.0001	<.0001	*
1.3*g*	nHC	HC	5973	11.66	<.0001	<.0001	*
1.5*g*	nHC	HC	5973	13.98	<.0001	<.0001	*
1.8*g*	nHC	HC	5973	12.67	<.0001	<.0001	*

Table 9.12: Haptic-Study - Comparison of the response time between the interface methods at the same stage of Gz load using SAS® PROC MIXED (adjust=simulate).

9.2.4 Discussion and Summary

With this experiment I intended to investigate whether exaggerating the isometric force production influences the target selection by goal-directed pointing towards nonhaptic interfaces. Therefore individually calculated weights to participant's pointing arm were balanced attached. To identify effects of altered isometric force on visuomotor performance, I compared the performance without weights with the performance using different weights (0.3, 0.5 and 0.8 times of the arm weight). Resulting from a within-subject study conducted with 29 participants, the data showed that increased isometric force speeds up the visuomotor performance using a nonhaptic interface (Q1). Thereby, the fastest pointing was supported by using 0.3 times of the arm's weight (Q2), which differed significantly from the performance using no weights. Investigating the performance with respect to the inclination angles of the gravity vector revealed a similar effect for vertical pointing movements (Q3). The participants were additionally requested to conduct the same experiment, which implied

pointing towards targets, supporting tactile sensation by haptic cues. The results showed that the increased isometric force did not affect associated pointing times. Thus, it can stated that the findings are exclusively applicable to nonhaptic interfaces (Q4). Finally, it was affirmed that haptic feedback is required for improved visuomotor coordination in general. Overall, the results indicated that the idea of increasing the isometric force for operating nonhaptic interfaces is a promising approach. Future work should improve the way of how exaggerated isometric force will be generated. Using the complete arm is not user-friendly and comfortable. I can imagine that a small device around the wrist could overcome this problem.

Chapter 10

Conclusion

"Scientific knowledge is a body of statements of varying degrees of certainty - some most unsure, some nearly sure, none absolutely certain."

(Richard P. Feynman, *The Value of Science*, *What Do You Care What Other People Think?,* 1988)

In this chapter the findings of the research presented in the last chapters are summarized to support the verification of the initial thesis statements related to their fulfillments. This enables to ensure whether the meta-objective of this thesis have been met. Based on the lessons learned, a set of guidelines on Human Factors is presented, supporting the design of AR interfaces for space payload operations, but also providing principles with a high generality to their application. Finally, future directions are discussed, setting up the prerequisites to continue the work presented in this thesis.

10.1 Summary of Findings

The research presented in this thesis was intended to ease astronauts' intravehicular space activities. Besides operating the ISS, astronauts have to accomplish a series of maintenance and mission specific tasks at various International Standard Payload Racks (ISPRs), providing hardware and tools to carry out scientific and technological experiments in microgravity (see section 2.3). Such tasks are guided by standardized procedures, providing step-by-step instructions and supplementary resources, which are usually viewed by the SSC laptop laterally fixed at the rack construction. This forces a frequent switching between perceiving and processing an instruction, and its execution, whereby an effortful visual search is going ahead. As revealed by analyzing the problem space (see section 5.1.1), these shortcomings can lead to failure of attention and working memory. As a consequence, the performance quality of payload operations can suffer from an increased workload level and thus bears the risk of astronaut errors. Using an integrated design, which embeds all relevant information of an instruction into one field of view, would not only enable to maintain the focus of activity and disburden the working memory, it would also enable to cue visual attention during the task localization. In conformity with its characterized features, it is obvious that Augmented Reality has the greatest potential to improve the work of astronauts during intravehicular payload operations (see section 5.1.2). As a complement to existing technological-driven approaches (see section 5.1.3), the work conducted in this thesis was directed towards a user-driven design for AR-assisted space payload operations. This not only supported the identification of challenges tackled in this thesis (see section 5.2), it generally enabled to compensate the lack of Human Factors, which in the best case ensures an efficient system design, satisfying the user's needs from the very beginning. Overall, the research requested by the challenges implied to carry out six experiments. With the exception of the proof-of-concept study, these experiments corresponded to trade-off studies, investigating relevant interaction issues. In the following, the findings of these experiments are summarized in reference to their underlying chapters of this thesis.

Not research in its strict sense, but necessary to conduct the first experiment, the work presented in section 6.1 yielded a high-fidelity prototype intended to provide Mobile AR for Space Operations (MARSOP), which enables, on the one hand, the pragmatic authoring of AR-based payload

procedures and, on the other hand, a viewer to guide astronauts through payload procedures by means of AR support. By introducing the MARSOP system, I could not only close the fundamental gap left by the related space flight project, it also enabled me to clarify whether the concept of using AR can contribute to payload tasks by fulfilling the demand of Human Factors directives. Thus, this research tackled the first challenge (see section 5.2.1), investigating the effect of mobile AR on payload tasks. Thereby the main challenge was given by the claim to carry out usability testing in the field, calling for a real payload environment, real payload tasks and participants, who are representative of the intended user population. Only in this way I could prove whether AR works and is accepted from domain experts. This research on field testing to prove the MARSOP concept (see section 6.2) revealed following findings:

(1) This research has shown that the use of AR for instruction gathering is capable to fulfill the requirements of ISPR payload operations.

(2) This research has shown that AR reduces sequence errors and ensures the sequential mode of payload operations, and thus provides a higher level of performance quality than the common guidance method.

(3) This research has proven that the subjective workload during those tasks, which vary in their instructions and so requiring constant learning, benefit more from AR than from the common guidance method.

(4) This research indicated a high level of satisfaction and acceptance from domain application experts for AR instead than for the current guidance interface.

After it was initially confirmed that AR has the potential to replace the common guidance interface, I identified two essential challenges in respect to the task and environmental conditions from heuristic evaluation. These challenges were addressed, on the one hand, to the visual search needed for task localization at the payload and, on the other hand, to the spatial device placement of interfaces used for control and symbolic input tasks. Considering the aspect on task localization, its research constituted the second challenge tackled in this thesis. This challenge was focused on the spatial orientation of target cues given by the resources of a payload instruction (section 5.2.2). Its objective was to find out the most suitable spatial reference frame for payload operations, which can be affected by egocentric and exocentric coding of the target cues. Thereby, the cognitive process of spatial updating is not only affected by the spatial orientation of the display, in which the target cues are presented, it is also influenced by the spatial orientation of the visualization within this display. Considering these criteria, three conditions for target cueing were identified: an exocentric display condition, which also implies an exocentric visualization, and two egocentric display conditions, providing either exocentric or egocentric visualization. Because related research on AR task localization (see section 4.1.2) revealed ambiguous results and thus there is no adequate guideline, it was needed to solve the trade-off between egocentric and exocentric target cueing for payload tasks. To answer the question where task cues are displayed most efficiently, it was important to evaluate the effort spent on task localization in an adequate way, as a function of the resulting workload level. By reason of frequent neglect of assessing the workload in AR usability testing, the main contribution of this research was to raise the awareness for the importance of assessing the workload in its full dimension. Thus, I evaluated and discussed the efficiency of task localization by primary and secondary task performance, subjective rating as well as by cardiovascular responses, analyzing the heart rate variability (HRV). Thereby, the task localization performance was also analyzed with respect to the initial search of the off-screen target and the subsequent in-view detection. This research on AR task localization (see Chapter 7), also referred to as the "ARGuide" study, revealed following findings:

(5) This research has proven that target cues presented within an egocentric display accelerate the visuomotor performance of the overall task localization, compared to exocentric target cueing, as provided for payload operations by the SSC-based external display.

(6) This research has proven that the pointing accuracy is not affected by the spatial orientation of the target cues.

(7) This research has proven that an egocentric presentation of target cues within an egocentric display accelerates the visuomotor performance during the in-view detection of the target object.

(8) This research has proven that head movements are reduced by an exocentric area map presented within an egocentric display.

(9) This research has proven that an exocentric area map presented within an egocentric display leads to the lowest effort spent on concurrent task integration and performance during the initial search.

(10) This research has shown that an exocentric area map presented within an egocentric display provided the most relaxed state.

(11) This research has proven that the subjectively rated workload benefits more from an egocentric presentation of target cues within an egocentric display than from those presented within an exocentric display.

(12) This research has shown the importance of multidimensional workload assessment for AR usability testing.

(13) This research has indicated that target cues presented by an exocentric area map within an egocentric display provoke egocentric spatial updating during the task localization process.

In the course of the research on AR task localization I also considered the investigation of the relevance of human's individual predisposition in spatial processing (see Chapter 7). This additionally revealed the following findings:

(14) This research has shown that human individual predisposition for visual spatial processing needs to be considered in AR research, which is addressed to user's spatial navigation skills.

(15) This research has shown that users following an egocentric strategy are performing better with an egocentric display and presentation scheme than those one following an egocentric strategy.

(16) This research has proven that users following an exocentric strategy consume less processing resources during task localization than those one following an egocentric strategy.

While the above presented findings are only valid for normogravity conditions, the research on spatial device placement of interfaces, used for control and symbolic input tasks, was addressed to the environmental restriction given by the absence of gravity. Thus, the third challenge tackled in this thesis, was faced by the question how influential gravity is for the placement of AR input devices (see section 5.2.3). Following the related AR research under normogravity conditions (see section 4.1.3), this involved to consider a world, body and head reference to place such interfaces, but there was no study undertaken under modified conditions of gravity. In contrast, microgravity research comprehensively investigated goal-directed pointing, but the related findings cannot answer the question on device placement. Thus, the third challenge, tackled in this thesis, was addressed to investigate the effect of altered gravity conditions on the quality of visuomotor coordination revealed from a world, hand and head referenced AR interface. To solve the trade-off between these three

placement conditions, I conducted two experiments under parabolic flight conditions in the scope of the "3DPick" study. Being aware that such experiment environment provides only knowledge about short-term hyper- and microgravity, it nevertheless offers the opportunity to check whether an effect in principle occurs. This research on AR device placement under altered gravity (see Chapter 8) revealed following findings:

(17) This research has shown that AR is operable under short-term hyper- and microgravity.

(18) This research has verified that haptic feedback is superior to non haptic feedback for goal-directed pointing tasks under normogravity.

(19) This research has proven that haptic feedback is superior to non haptic feedback for goal-directed pointing tasks under short-term hyper- and microgravity.

(20) This research has proven that visuomotor performance under normogravity does not differ between AR interfaces referenced to the world and those referenced to the hand.

(21) This research has proven that visuomotor coordination under short-term hyper- and microgravity benefit most from handheld AR interfaces.

(22) This research has proven that hyper- and microgravity affects visuomotor performance on world-, hand and head-referenced AR interfaces, but in different ways.

While all findings presented up to here were revealed from challenges identified from previous heuristic evaluation, other challenging questions have arisen by the research on AR device placement under altered gravity conditions. Reasoned by its outcomes, it was needed to enhance hand-eye coordination resulted from a world- and head-referenced AR interface. Therefore I conducted two countermeasure studies, both considering, but in different ways, the idea of applying increased sensorimotoric load to improve their underlying visuomotor performance (see Chapter 9), and therefore these studies were referred to as the "GSim" studies. Taking this into account, to encounter deficits by the world reference, sensorimotoric load was applied to the interface characteristics. Therefore, I introduced a resizing approach, which enables the adaption of the present gravity load to the size of the pointing targets and the distance between them. This approach implied two different resizing conditions, resulting in either compressed or elongated interfaces, whereby both interrelate to the characteristics of Fitts' law, which predicts longer movement times at greater distances, as well as at smaller targets. To proof the concept of this resizing approach, I conducted a trade-off study under simulated hypergravity conditions, comparing the visuomotor performance resulted from unchanged, compressed and elongated AR interfaces, which was supplemented by analyzing their speed-accuracy trade-offs modeled after Fitts' law. This research on gravity-adapted target and interface resizing (see section 9.1) revealed following findings, which are only valid under increased gravity conditions:

(23) This research has proven that gravity-adapted target resizing affects the performance and workload of direct pointing towards world-referenced AR interfaces.

(24) This research has proven that elongated resizing (greater target size, longer target distance) improves the visuomotor performance with world-referenced AR interfaces.

(25) This research has proven that elongated resizing provides the highest throughput while handling world-referenced AR interfaces.

(26) This research has indicated that compressed resizing (smaller target size, shorter target distance) supports the least physiological workload while handling world-referenced AR interfaces.

In contrast to the approach of enhancing world-referenced AR interfaces by applying gravity force to the interface, the concept to optimize head-referenced interfaces has envisaged to apply increased sensorimotoric load to the pointing arm. Because a head reference implies the absence of haptic cues, the additionally load should compensate the lack of tactile sensation by amplifying the proprioceptive sense, which in turn lead to an exaggerated isometric force level of the pointing arm. To initial study its effect on goal-directed pointing under normogravity conditions, I conducted a usability test, comparing the visuomotor performance revealed under different increased load conditions, to the normal unloaded condition. In case of an occurring effect, it would be meaningful to ensure that this approach counteracts the missing tactile sensation exclusively. Thus, the study was also performed for a world-referenced interface, which provided haptic feedback. This research on changing isometric force while operating nonhaptic interface (see section 9.2) revealed following findings:

(27) This research has verified previous findings, stating that haptic feedback is superior to non haptic feedback for goal-directed pointing tasks under hypergravity condition.

(28) This research has proven that increased sensorimotoric load attached to the pointing arm improves visuomotor performance under normogravity conditions while handling a nonhaptic head-referenced interface.

(29) This research has proven that using 0.3 times of the arm weight leads to faster pointing while handling a nonhaptic head-referenced interface under normogravity conditions.

(30) This research has shown that increased sensorimotoric load attached to the pointing arm does not affect visuomotor performance on a haptic interface, referenced to world objects.

Overall, the research conducted in the scope of this thesis yielded a lot of findings, expanding the knowledge on how can AR contribute to payload operations and what needs to be considered when operating AR input devices under altered gravity conditions, as well as how deficits caused by wrong placement condition can be compensated or at least mitigated. To get an impression about the extent and time needed to conduct this research, Figure 10.1 shows its assigned studies placed in a timeline with competing works. It also shows the project in progress, continuing the work on the "ARGuide" project. The "µARGuide" study is intended to investigate the effect of short-term hyper- and microgravity on the spatial orientation of target cues used for task localization. Meanwhile of writing this thesis, Microsoft Cooperation and NASA has started they project Sidekick[78], which is also positioned at the timeline.

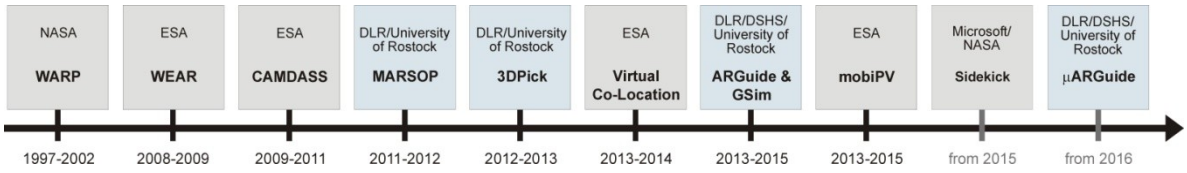


Figure 10.1: Timeline positioning of the studies conducted in the scope of in this thesis.

[78] With the project Sidekick, which has started in 2015, Microsoft Cooperation and NASA intend to support astronauts' work by AR, either remotely by experts or by AR instruction manuals (Metz, 2015). It is also planned to use AR for inventory management. Thereby the central focus is on Microsoft's HoloLens, which has been introduced in January 2015.

10.2 Revisiting Thesis Statements

As stated in section 1.4, the meta-objective of the research conducted in the course of this thesis was:

> To improve the work of astronauts during intravehicular payload operations by means of advanced user interfaces using Augmented Reality.

To verify this objective, several hypotheses were drafted, which will be revisited here. In consideration of the findings (1-30) summarized in last section, I will discuss each thesis statement related to its fulfillment in the following:

H1: Augmented Reality is capable to fulfill the requirements of payload operations.

The MARSOP system presented in section 6.1 resulted in a fully operational AR prototype that enables easy authoring and viewing of AR-based PODF procedures. Summative evaluation by field testing, as presented in section 6.2, has shown that AR is applicable and fulfills payload tasks in the same way as usual (1). Although there were complaints about the quality of the used HMD, AR was accepted from application domain experts and predominantly preferred (4). Thus, this hypothesis can be confirmed.

H2: Augmented Reality enhances the performance of payload operations.

By means of Biolab procedures and resulted from field testing, it was shown that AR increases the performance quality of payload operations. It not only enables a homogeneous workflow, it also ensures the sequential mode of operation, decreasing sequence errors (2). Subsequent research on task localization has also shown that AR accelerates the localization process in general (5) and especially during the in-view detection of the target object (7). Although the former one is not only valid for an egocentric presentation scheme, it could also be verified for an exocentric presentation scheme, like an area map, but still presented by an egocentric AR display. Thus, this hypothesis can be confirmed too.

H3: Augmented Reality decreases the workload during payload operations.

While the findings of field testing supports the assumption that the subjective workload during payload operations benefits from AR (3), subsequent research on task localization has revealed ambiguous workload findings, which vary as a function of the method applied for assessment. Although it could also be shown that self-reported workload benefits more from an egocentric AR target cueing than from cueing as provided by the common external display (11), other workload dimensions benefit from an exocentric presentation of target cue, like an area map, presented within an egocentric AR display. So it could be shown that such a presentation scheme contributes more to a better dual-task integration (9) and an optimal level of arousal assessed by HRV (10). Hence, this hypothesis can be only partially confirmed, although an exocentric presentation of target cues within an egocentric display can be assigned to AR, following the definition from Milgram and Fumio (1994),. Moreover, these findings have corroborated the importance of exploiting the full dimensions of workload assessment (12).

H4: Augmented Reality is operable under altered gravity conditions.

This hypothesis can be confirmed by findings revealed from research conducted under parabolic flight conditions (17). It has shown that AR is not only operable under short-term hypergravity, but also, and more important, under short-term microgravity. Thus, it can be assumed that AR will also be operable under long-term microgravity, as will be requested by AR assisted payload tasks.

H5: Handheld Augmented Reality input devices enhance the quality of visuomotor coordination under altered gravity conditions.

While the pointing performance on world- and hand-referenced interfaces did not differ under normogravity conditions (20), equivalent experimentation under parabolic flight conditions has shown that goal-directed pointing under short-term hyper- and microgravity benefits more from hand-referenced interfaces (21), although their handling is affected by modified conditions of gravity as just by handling a world- and a head-referenced interface (22). Therefore, this hypothesis can be confirmed. Even if the hand is part of the body, it cannot be assumed that using other parts of the body contributes in the same way, as was shown for the forearm by Harrison et al. (2011). It should also be mentioned that the confirmation of this hypothesis is restricted to pointing movements and can differ for other input modalities, implying, for example, sliding movements (Lindeman et al., 1999).

H6: Tactile sensation enhances the quality of visuomotor coordination while operating Augmented Reality input devices under altered gravity conditions.

First, it should be mentioned that previous findings from Lindeman et al. (1999) could be confirmed, showing that haptic feedback is required to operate AR input devices under normogravity conditions (18, 27). The findings revealed from parabolic flight tests, confirm the same under altered gravity conditions. By comparing with a world- and hand-referenced interface, it could be shown that a head reference, not supporting tactile sensation, deteriorate the pointing performance under short-term hypergravity, but even more under short-term microgravity (19). Thus, this hypothesis can be confirmed.

To support answering the meta-objective of this thesis, it can be stated that all hypotheses could be generally confirmed by the outcomes of empirical tests of relevant research. As a result, the meta-objective can be considered as achieved and, thus, it can be verified that the application of AR as advanced user interface for intravehicular payload operations is a promising approach to improve astronauts' task performance.

10.3 Guidelines and Principles

To support a user-centered design process for AR assisted payload tasks in an efficient way, this section presents a set of guidelines derived from the findings revealed in the scope of this thesis. Because some of the guidelines provide a high level of generality to its application, they even comply with principles.

Besides the fact that the traditional external display, as provided by the SSC laptop, should be replaced by an egocentric display, such as a HMD, the first guidelines specify issues on the spatial orientation of target cue visualization, considering the usage of non-conform cues and 3D-registered AR cues within an egocentric display. To ensure the best visuomotor performance, but also an optimal workload level, it is recommended to use a hybrid visualization method for efficient task localization. Assuming an egocentric display, the visuomotor performance during the initial search to localize the off-screen target area is not affected by the spatial orientation of the visualization scheme, while the workload level benefits more from an exocentric static visualization. Things look different for the subsequent in-view target detection. Thereby a 3D-registered AR visualization of the target cues contributes to enhanced visuomotor performance and is more inclined to support a better workload level. Thus the following guidelines are specified for efficient task localization for payload operations:

(1) To localize an off-screen target area by AR, use a non-conform visualization of the target cues, like a display-fixed area map.

(2) For in-view detection of the target object by AR, use a 3D-registered AR visualization of the target cues, like highlighted geometry.

Additionally, if payload tasks are complex or tedious, it is important that the operator provides sufficient attentional resources. This can be ensured by considering the individual predisposition in spatial navigation of the operator. Because it seems that exocentric operators spend less attentional resources during task localization in general and independent of the used visualization scheme for the target cues, it is suggested to meet the following guideline:

(3) Be aware that operators having an egocentric navigation strategy may require more attentional resources for task localization.

The next one is modeled after this guideline, but rather complies with a principle for AR usability testing. In the case of designing alternative approaches for target cueing or more general for navigation purposes, corresponding usability testing should consider such individual differences in spatial navigation, and thus it is recommended:

(4) When performing usability tests on AR navigational processes, extend the study design to analyze differences between egocentric and exocentric users.

Therefore the Turning Study[79] related to Goeke et al. (2013) can be used to identify which navigation strategy is given to the participants by nature. More generally and not only suited for usability testing related to issues of task localization, the following is also recommended:

(5) When performing usability tests on AR interfacing, assess the workload level multidimensionally, using primary task measures, dual-task integration, biofeedback and subjectively ratings.

[79] Turning Study. URL: http://www.navigationexperiments.com/TurningStudy.html, last visit: 28.09.2016

Thereby, it is important to use the right methods, as described in detail by (Gawron, 2008) or in particular for physiological measures by Wierwille (1979). The guidelines (1)-(3) are still restricted to normogravity conditions and need to be verified for microgravity. They are also restricted to payload operations. However, the following guidelines are referring to the placement of AR input devices, which are decoupled from payload operations, and thus they provide a higher level of generality to their application. Different than before, these principles are not only considering their usage under normogravity, but also under altered gravity conditions. Assuming an AR interface which is operated by direct pointing, it is suggested to meet the following three principles:

(6) When providing AR input devices under normogravity, hypergravity or microgravity conditions, ensure that tactile sensation is supported by haptic cues.

Because head-referenced interfaces usually provide no haptic cues, it is not recommended to use such interface reference, regardless of the present gravitational condition. Rather it is suggested to use a world- or hand-referenced interface. Thus, the following principles suggest which reference type is most suited in accordance with the gravitational condition:

(7) When placing AR input devices under normogravity conditions, use a world or a hand reference.

(8) When placing AR input device under hypergravity and microgravity conditions, use a hand reference.

This does not mean that a world-referenced interface is not operable under hyper- and microgravity conditions, but to ensure the best hand-eye coordination, it is more appropriate to use a hand-referenced one. If the usage of a world-referenced interface is nevertheless desired, for example, under hypergravity, it is necessary to adapt the interface characteristics to the present gravity load. Therefore the following recommendation is proposed:

(9) When providing a world-referenced input device under hypergravity conditions, enlarge the targets and expand the distance between the pointing targets, corresponding to the present gravity load.

Something similar is required to use a head-referenced interface, which usually not provides tactile sensation by haptic cues. To compensate the loss of haptic feedback, it is suggested to change the isometric force level of the pointing arm by attached weights. Assuming normogravity conditions, the following is suggested:

(10) When providing a nonhaptic input device under normogravity conditions, increase the isometric force level of the pointing arm by 0.3 times.

Aware that there is still room for supplementations, following this set of guidelines and principles in conjunction with existing design rules for usability (see section 3.2), an interface designer should allow to deploy a good conceptual model for AR assisted space payload operations from a Human Factors perspective or to continue the process started in this thesis.

10.4 Subsequent Work and Future Directions

As finally mentioned in the last section, the work done in the scope of this thesis still leaves room for improvement and supplementation, raising further issues addressed to research on AR-assisted payload operations. Thereby future directions proposed in this section range from research on perceptual-cognitive changes, over compensating inter-individual differences in spatial updating, to enhancing the operation of nonhaptic interfaces. But before presenting such future directions, it is necessary to discuss what subsequent work needs to be done to complete the initial MARSOP prototype presented in section 6.1, as well as what issues are still relevant to, for example, the content presentation, which can be valorized by enhancing the presentation of situated labels provided by the MARSOP Viewer. Just as important as completing or improving the MARSOP prototype, is the adaptation of the PODF Standard to the needs of AR content visualization, which will also be discussed here.

10.4.1 Completion of the MARSOP Prototype

Although the MARSOP system, as presented in section 6.1, is able to easily transfer a PODF procedure into the AR space and was finally accepted by application domain experts, it should be noted that the resulting high-fidelity prototype is not yet fully completed. To finalize the iterative design process, started in this thesis, it is required to enrich the initial prototype by integrating functionalities related to task localization and the placement of input devices, but considering their related guidelines, as presented in last section. While the research on device placement has already considered the effect of altered gravity, equivalent microgravity research for task localization is missing, but currently in progress. Although each design iteration requests separate usability testing in course of the MARSOP system, it was not taken into account because of the required environmental conditions, but it would also be inefficient to conduct further ground-based experiments without considering weightlessness. Thus, after conducting microgravity research on task localization, subsequent work should be addressed to verify the usability of an upgraded version of the initial MARSOP prototype aboard the ISS. As was done by the proof-of-concept study (see section 6.2), a results-driven summative evaluation should give information about the beneficial effects of using AR yielded by astronauts, while the shortcomings of the MARSOP study, as described in section 6.2.7, should be resolved. Therefore, the presence of the HoloLens[80] aboard the ISS can compensate the deficit of using an inadequate display and could substantially increase the performance quality. Another point of criticism was the non-uniform and partially missing acquisition of performance data regarding the traditionally guidance condition. To make it more comparable with the performance, resulting from the MARSOP system, it would be necessary to support the common guidance condition with AR equipment. As was learned from the research on task localization, the field test should also consider assessing the workload in its full range.

Just as important as the issues mentioned before are technical aspects, only slightly touched in the course of this thesis. This not only includes the usage of an adequate HMD, but also a stable 3D pose tracking, as the most important issue for AR. The model-based approach, as was used in this work, needs to be optimized. Besides providing a more detailed virtual model of the payload, it is conceivable, for example, to dynamically load different subsets of keyframes, depending on the task

[80] It should be noted here that in the scope of the Sidekick project, Microsoft together with NASA has conducted first tests of the HoloLens under altered gravity condition onboard the Weightless Wonder C9 jet in June 2015. Since December 2015, the Hololens is available onboard the ISS and tested by astronauts.

area in question. This would enable to reduce the number of needed keyframes, which in turn could increase the tracking performance. As was already discussed in section 6.2.2, the approach of model-based KLT tracking fails to comply with the requirements as set out by near distance application due to the lack of sufficient 2D features, causing frequent initialization phases. Thus, it be should tested whether using a hybrid approach, combining model-based tracking with an inertial system or object-based SLAM (e.g., Selvatici et al., 2008, Salas-Moreno et al, 2013) can better meet the tracking requirements.

10.4.2 Enhancement of Situated Labels

In section 6.1.4.1, it was outlined how the MARSOP Viewer visualizes the content provided by an AR ODF procedure. Besides screen-stabilized visualization, the viewer also supports seamless embedding of situated visualizations, which are information that is not only spatially but also semantically associated with the payload surrounding and the task requested by a procedure step. Besides 3D auxiliary objects and highlighted geometry, the MARSOP Viewer supports labeling of the payload's components, which are currently presented only in a static manner. This means that the label is anchored to the specified 3D point of the referred object, but repositioned by a unitary offset, whereby both the label and the referred object are linked with a straight connection line. Consequently, it bears the risk of a lack of clarity in regard to occluding relevant payload components and, in cases of multiple labeling, of overlapping labels or crossing connection lines. Thus, the viewer needs to be extended by an strategy for optimal label setting in consideration of the criteria for label placement, which, amongst others, request for a close proximity between the label and the referred object, a non-overlapping and non-crossing of the labels and connection lines, but also for a minimal length of each connection line (Schmalstieg and Höllerer, 2016).

10.4.3 Adaptation of the PODF Standard

In cases of adopting AR to payload operations, another issue of directing future activities calls for an appropriate adjustment of the PODF standard to the claims of an AR surrounding. This means that general standards, used to compile a PODF instruction, like embedded symbols and signs, need to be replaced by adequate presentation schemes to provide screen or world-stabilized content visualization. It should also be revised how useful it is to continue using written text-based instruction. On the one hand, showing text on top of the visual display is not coherent with the meaning of AR and wastes unnecessary space, which can be used more effectively, for example, for the visualization of a 2D area map, indicating the target location. On the other hand, it may induce a resource conflict, because reading an instruction claims the same perception processor than viewing related image resources or observing the task area.

10.4.4 Investigation of Perceptual-Cognitive Changes

One critical point in applying AR to instructional operations is the risk of perceptual narrowing, also referred to as attentional tunneling, especially during longer task sessions. As was indicated by the findings resulted from the research on task localization (see Chapter 7) and already discussed by Grubert et al. (2010) for picking processes, AR may cause an increased level of stress or arousal, which in cases of overarousal can induce the attentional tunnel effect. Thus, it is important to verify that such effect is not induced during AR-assisted payload operations. In cases where AR is tunneling, it should be tested whether decreasing the degree of registration can counteract this

problem, because the findings resulted from the research on task localization have shown that using an 2D area map on top the visual display can provide a lower level of arousal. Closely linked to the problem of attentional tunneling, a more general question has driven my interest while writing this thesis. Besides the fact that AR is clearly increasing our perception, supporting fast bottom-up processing, it is more important how AR is influencing our cognition. So for example, it is needed to examine whether the application of AR, especially those supporting task performance, is capable to induce learning or to stimulate mental processes that will be of essential importance for off-nominal situations, where improvisation and creativity are requested. This is similar to Sweller's imagination effect (2008), which ensures better task performance of experts when the operation, covered by an instruction, can be mentally simulated. This presumes that the guidance fading effect was considered, ensuring less assistance as the level of expertise increases.

10.4.5 Compensation of Differences in Spatial Updating

Besides attentional tunneling, as discussed before, the research on task localization also showed that operators having an egocentric strategy for spatial updating consume more attentional resources than those having an exocentric strategy, independent of the method used for target cueing. Thus, using egocentric operators, especially for long payload tasks, bears the risk of an increased workload level, which in turn can decrease the performance quality. This frames the next challenge for future research. By doing so, first it is important to replicate this finding. Then, in case of confirmation, it is needed to find out what affects egocentric users, to subsequently develop a method to compensate such inter-individual differences, and, thus, to reduce the effort spent by egocentric users.

10.4.6 Enhancement of Operating Nonhaptic Interfaces

The last issue discussed here, is directed towards enhancing nonhaptic interfaces, which are interfaces floating in the air in front of the user's eye to enable, for example, symbolic input by direct touch selection. Although haptic feedback is required to stabilize the proprioceptive sensation of pointng movements, the application of such interfaces would, nevertheless, be meaningful, because they are always accessible at a short distance without the necessity for using or providing surfaces. Thus, the related research conducted in this thesis mainly served the purposes of stimulating efforts spent on improving the handling of nonhaptic interfaces. Because it might be reasonably assumed that the operation of such interfaces is disadvantaged by the lack of proprioception, the most obvious approach is to increase the sense of force to substitute the missing sensory information. Taken such approach into account, the results of the research presented in section 9.2, showed that pointing movements towards non-haptic interfaces can be speeded up by raising the isometric force level of the pointing arm. Besides its verification for other operating modes, like sliding movements, future work could be targeted to improve the way the increased sensorimotoric load is generated. Using the complete arm to attach weights, as was done here, is not user-friendly and comfortable. I can imagine that a small device around the wrist could overcome this problem. Furthermore it should be tested whether an individual weight is required, or it will be sufficient to use a standard weight. With respect to the problem space considered in this thesis, it would also be useful to operate such interfaces in weightlessness. Because weights are not suited in such an environment, the arm positioning sense needs to be boosted in a different way, such as using a rubber band.

Appendix A

NASA Raw Task Load Index (RTLX)

Mentale Anforderung *(Mental Demand)* Wie geistig anspruchsvoll war die Aufgabe

(How mentally demanding was the task?)

sehr niedrig / *very low* — sehr hoch / *very high*

Physische Anforderung *(Physical Demand)* Wie körperlich anstrengend war die Aufgabe?

(How physically demanding was the task?)

sehr niedrig / *very low* — sehr hoch / *very high*

Zeitliche Anforderung *(Temporal Demand)* Wie eilig oder hastig war das Tempo der Aufgabe?

(How hurried or rushed was the pace of the task?)

sehr niedrig / *very low* — sehr hoch / *very high*

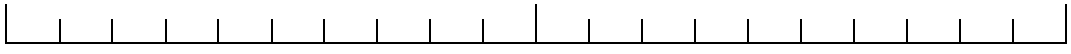
Durchführung *(Performance)* Wie erfolgreich waren Sie beim Erreichen worum Sie gebeten wurden?

(How successful were you in accomplishing what you were asked to do?)

Ausgezeichnet / *Perfect* — erfolglos / *failure*

Arbeitsaufwand *(Effort)* Wie schwer mussten Sie arbeiten, um Ihr Anforderungsniveau zu schaffen?

(How hard did you have to work to accomplish your level of performance?)



sehr niedrig / *very low* — sehr hoch / *very high*

Frustration *(Frustration)* Wie unsicher, entmutigt, irritiert, angespannt und verärgert waren Sie?

(How insecure, discouraged, irritated, stressed, and annoyed were you?)

sehr niedrig / *very low* — sehr hoch / *very high*

Appendix B

MARSOP Prototype: Example Files, AR Displays, File Formats

This appendix is related to the description of the high-fidelity MARSOP prototype presented in section 6.1. Besides listing the head-mounted displays available for the work conducted in this thesis, this appendix also presents various examples of XML files used by the MARSOP constructor and viewer component. This includes the XML specification of the keyframes for the markerless model-based tracking and a traditional ODF procedure required by the MARSOP Constructor, while the viewer needs only an AR ODF procedures resulting from the constructor. This section also provides the core elements and their belonging attributes of the AR Procedure File (ARPF) format.

Markerless Model-based Tracking provided by the MARSOP system

Extract from a XML file resulted from the offline keyframe generator: (from Millberg, 2012)

```
<!DOCTYPE Keyframe>
<Keyframe>
    <Image width="320" height="240" name="01.jpg"/>
    <KeypointInformation descriptorDim="64" detector="SURF" descriptor="SURF"
        numberOfKeypoints="97"/>
    <Keypoint>
        <Descriptor x="153.773" y="84.7297" size="8.35512" octave="0" response="0.759986"
            class_id="0" angle="-2.9824">
                0.00122089  0.00271965  -0.00268632 0.00450122  -0.00882315 0.0226538
                -0.00664186 0.0253064   0.00808463  0.0295203   0.00661055  0.0174793
                -0.00057708 0.0039023   0.00100097  0.00189519  0.0149665   0.0320002
                0.00573297  0.0437752   0.302176    0.452031    -0.102812   0.296322
                0.029667    0.0989913   -0.0746891  0.190518    -0.0149685  0.0265143
                -0.00243514 0.0135818   -0.027037   0.0419613   -0.00105078 0.0985082
                0.180175    0.490271    0.0845008   0.310001    0.073062    0.146853
                -0.0116057  0.346151    -0.005131   0.019575    -0.00425628 0.0303196
                -0.00632919 0.00673322  -0.00164923 0.00699219  0.0281832   0.0444242
                -0.0171835  0.0559437   -0.00146417 0.0167543   0.00167506  0.0582077
                -0.00197842 0.00210129  -0.00119001 0.00343525
        </Descriptor>
        <Correspondences x="-24.8724" y="136.405" z="137.501"/>
    </Keypoint>
    ...
</Keyframe>
```

Available Head-Mounted Displays

The table below shows the head-mounted displays (HMD) that were available for the work activities conducted in the scope of this thesis. Both HMD approaches, video see-through and optical see-through, types of AR displays were available in different designs (see Figure 6.6).

Head-Mounted Diplay (HMD)	VST/OST	FOV	Display	Optical Sensor	Applied during
ARvision-3D (Trivisio)					
	binocular video see-through	42° diagonal (4:3), 34° (horiz), 25° (vert)	two SVGA AMLCD, 24 bit color, 800x600 pixels	integrated stereo camera setup, 752x480 pixels, 60 fps	tests, demon-strations
WRAP 920AR (Vuzix)					
	binocular video see-through	31° diagonal	two LCDs, color, 24 bit color, 752x480 pixel	integrated stereo camera setup, 640x480 pixels, 60 fps	MARSOP study, ARGuide study
dataGlass2/a (Shimadzu)					
	monocular optical see-through	31° diagonal	semi-trans-parent LCD, 800x600 pixels	mounted Microsoft LifeCam HD-5000 webcam, 1280x720 pixels, 30 fps	3D Pick study, Weight study, Haptic Study
STAR 1200XLD (Vuzix)					
	binocular optical see-through	35° diagonal	two WVGA LCD, 24 bit color, 852x480 pixels	up to 1280x720 pixels	tests

AR Authoring provided by the MARSOP Prototype

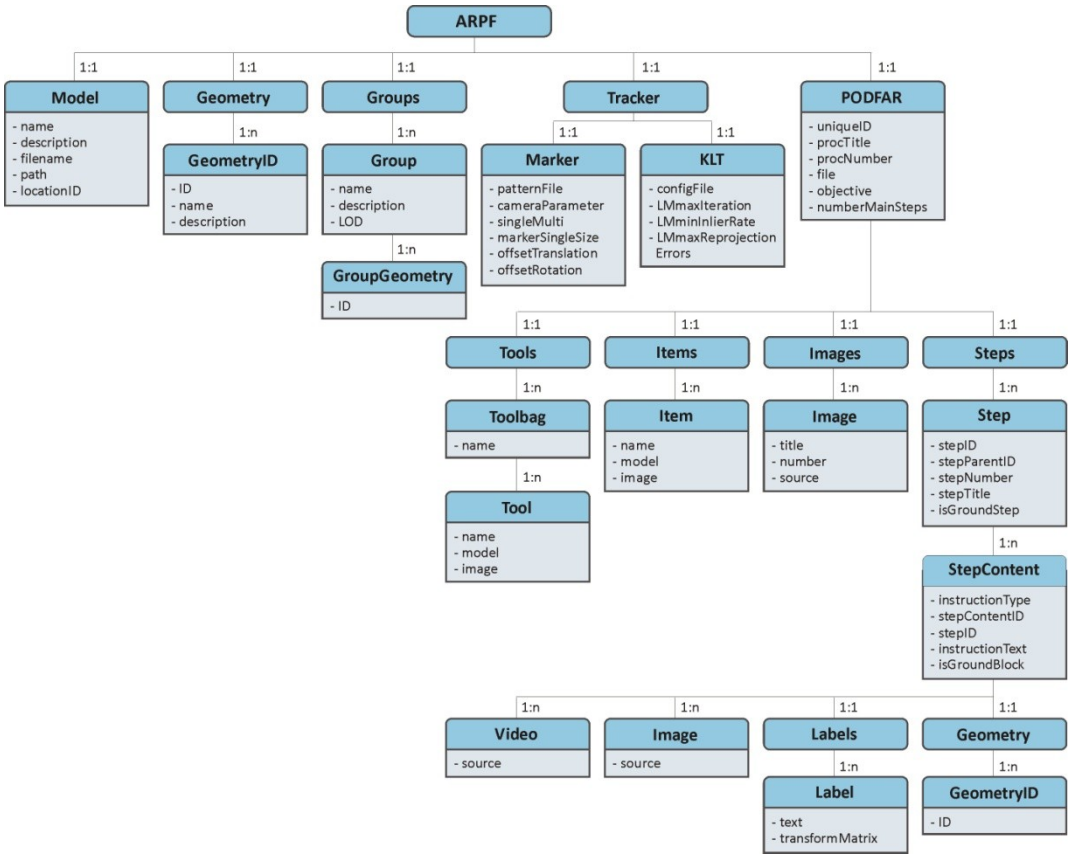
Example of a conventional ODF procedure converted to XML format:

```
− <ChecklistProcedure xsi:noNamespaceSchemaLocation="checklist.xsd" maxItemId="4005" primaryLang="en">
    <SchemaVersion>2.0</SchemaVersion>
    <AuthoringTool>NASA/ODF PAT 2.0.2; XMetaL 6.0.2.70</AuthoringTool>
  − <MetaData procType="Nominal" status="FIN" author="PW">
      <Date>22 JUL 2011</Date>
      <UniqueId>T1_E_BLB_INC_2510</UniqueId>
      <ProcCode>Biolab Rack / Inc. 28 - ALL / FIN-1 / HC</ProcCode>
    </MetaData>
  − <ProcTitle>
      <ProcNumber>2.510</ProcNumber>
      <Text>BIOLAB EXPERIMENT CONTAINER INSTALLATION</Text>
    </ProcTitle>
  − <ProcedureObjective lang="en">
    − <Para>
        <Text>This is a generic procedure for installing experiment containers (EC) to the Biolab Centrifuge Rotors inside the Incubator. The procedure
          can be coupled with PODF 2.511 (BIOLAB EXPERIMENT CONTAINER REMOVAL) if necessary. </Text>
      </Para>
    </ProcedureObjective>
  − <ItemizedList>
    − <ListTitle><Text>CREW</Text></ListTitle>
    − <Para><Text>One</Text></Para>
    </ItemizedList>
  − <ItemizedList>
      <ListTitle><Text>DURATION</Text></ListTitle>
      <Para><Text>15 minutes</Text></Para>
    </ItemizedList>
  − <ItemizedList>
      <ListTitle><Text>TOOLS</Text></ListTitle>
    − <ItemizedList>
        <ListTitle><Text>Columbus Tool Kit</Text></ListTitle>
        <Para><Text>Tool Bag 2:</Text></Para>
        <Para><Para><Text>Flat Screw Driver, 4 mm</Text></Para></Para>
      </ItemizedList>
    </ItemizedList>
  − <ItemizedList>
      <ListTitle><Text>ITEMS</Text></ListTitle>
      <Para><Text>Mini-Maglite1 P/N 528-20084-5</Text></Para>
    </ItemizedList>
  − <ItemizedList>
      <ListTitle><Text>REFERENCED PROCEDURES</Text></ListTitle>
      <Para><Text>ESA SODF; JOINT SYS-PL: NOMINAL; 2.201 BIOLAB SMOKE DETECTION MONITORING CONTROL</Text></Para>
    </ItemizedList>
  − <Step stepId="i3921">
    − <StepTitle>
        <StepNumber>1</StepNumber>
        <Text>GETTING EC FROM TCU</Text>
      </StepTitle>
    − <Step stepId="i3936">
      − <StepTitle>
          <LocationInfo><Location><Text>COL1A2_G2</Text></Location></LocationInfo>
          <StepNumber>1.1</StepNumber>
        − <Instruction>
            <ClearText><Text>Open completely TCU 2 Door (press latch on inner right side of dorr lever)</Text></ClearText>
          </Instruction>
        </StepTitle>
      </Step>
    + <Step stepId="i3937"></Step>
    </Step>
    <Step stepId="i1">
    − <StepTitle>
        <StepNumber>2</StepNumber>
        <Text>SAFETY INSPECTION OF INCUBATOR WORKING VOLUME</Text>
      </StepTitle>
    − <StepContent itemId="i2">
      − <Image display="inline">
          <ImageReference source="T1_E_BLB_INC_2510_files/E_BLB_INC_2510.jpg" width="640" height="459"
          alt="E_BLB_INC_3571_files/EC_Exch_Door.jpg"/>
        − <ImageTitle>
            <ImageNumber>1</ImageNumber>
            <Text>Inc EC Exchange Door Mechanical Latch and Inspection Window</Text>
          </ImageTitle>
        </Image>
      </StepContent>
    + <Step stepId="i4"></Step>
    + <Step stepId="i11"></Step>
    </Step>
  + <Step stepId="i3890"></Step>
  </Step>
+ <Step stepId="i31"></Step>
</ChecklistProcedure>
```

Example of an XML-based ARPF, providing an AR ODF procedure resulted from the MARSOP Constructor and executable by the MARSOP Viewer:

```
- <ARPF>
    <Model description="Facility for Biological Experiments on the ISS" objectives="Multi-user facility for cells, tissues, small plants and animals"
    path="data\3D_models\3ds" locationID="COL-A2" name="Biolab" filename="test4.3ds"/>
    <Geometry/>
  - <Groups>
    - <group description="-20 degree to +8 degree storage for samples and chemicals used in the automated process" name="AAS">
        <groupGeometry ID="Geode876"/>
        <groupGeometry ID="Geode819"/>
        <!-- further geometry IDs -->
      </group>
    + <group description="-20 degree to 8 degree storage for samples and chemicals used in the automated process" name="Automatic Storage (AS 1)"></group>
    + <group description="-20 degree to 8 degree storage for samples and chemicals used in the automated process" name="Automatic Storage (AS 2)"></group>
      <!-- further groups -->
    </Groups>
  - <Tracker>
      <Marker PatternFile="data/marker/multi/biolab.dat" MarkerOffsetX="10" CameraParameter="data/cam_param/camera_para.dat" MarkerOffsetY="20"
      MarkerOffsetZ="30" MarkerSingleSize="80" SingleMulti="multi"/>
    </Tracker>
  - <PODFAR Objective="To disengage TCUs 1 and 2 Launch Fixations during commissioning before first switch on of Biolab." NumberMainSteps="4"
    UniqueID="T2_E_BLB_TCU_1103" ProcTitle="BIOLAB TCUS LAUNCH FIXATIONS DISENGAGEMENT" File="T2_E_BLB_TCU_1103.xml"
    ProcNumber="1.103">
    - <ToolsColumbus>
      - <Toolbag name="Tool Bag 1">
          <tool model="" image="data/images/toolsItems/rachet.jpg" name="Ratchet Wrench 1/4""/>
          <tool model="" image="data/images/toolsItems/extension.jpg" name="102 mm Extension 1/4" Drive"/>
          <tool model="" image="data/images/toolsItems/torque.jpg" name="Torque Wrench 4-20 Nm"/>
        </Toolbag>
      - <Toolbag name="Tool Bag 3">
          <tool model="" image="data/images/toolsItems/M5.jpg" name="M5 (4 mm) Hex Head Driver 1/4" Drive"/>
          <tool model="" image="data/images/toolsItems/M6.jpg" name="M6 (5 mm) Hex Head Driver 1/4" Drive"/>
        </Toolbag>
      </ToolsColumbus>
    - <Images>
        <image title="TCU Launch Brackets In Launch Configuration" number="1" id="1" source="data/images/T2_E_BLB_TCU_1103_files
        /E_BLB_TCU_1103_f1.jpg"/>
        <image title="TCU Launch Brackets Stowage Locations After Removal" number="2" id="2" source="data/images/T2_E_BLB_TCU_1103_files
        /E_BLB_TCU_1103_f2.jpg"/>
      </Images>
    - <Steps>
      - <step stepParentID="0" stepID="1" isGroundStep="0" stepNumber="1" stepTitle="VERIFYING POWER OFF">
        - <stepContent InstructionType="VerifyCallout" stepContentID="1" locationInfo="COL1A2_G1" stepID="1" isGroundBlock="0"
          InstructionText="Verify Biolab Rack POWER sw - OFF">
          - <label text="POWER">
              <trans m30="-506.5237576184945" m20="0" m31="-134.9656247120984" m10="0" m21="0" m32="-562.8912636630748" m00="1" m11="1"
              m22="1" m33="1" m01="0" m12="0" m23="0" m02="0" m13="0" m03="0"/>
            </label>
          - <geometry>
              <geometryID ID="Geod981"/>
              <geometryID ID="Geod964"/>
            </geometry>
          </stepContent>
        </step>
      - <step stepParentID="0" stepID="2" isGroundStep="0" stepNumber="2" stepTitle=" REMOVING BRACKETS OF TCU DOOR LATCH SIDE ">
        - <stepContent InstructionType="ClearText" stepContentID="1" stepID="2" isGroundBlock="0" InstructionText="Unscrew and remove TCU launch
          bracket 1 (upper left) [four captive screws, Ratchet Wrench 1/4"; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive (4 mm
          for right screws and 5 mm for left screws)].">
            <video source="data/videos/T2_videos/T2_2_1.mpg"/>
            <image source="data/images/T2_E_BLB_TCU_1103_files/E_BLB_TCU_1103_f1.jpg"/>
          - <label text="Launch loaction of Bracket 1">
              <trans m30="-358.3452644872588" m20="0" m31="-116.4654657600371" m10="0" m21="0" m32="-46.57864071964957" m00="1" m11="1"
              m22="1" m33="1" m01="0" m12="0" m23="0" m02="0" m13="0" m03="0"/>
            </label>
          </stepContent>
          <stepContent imageID="2" stepContentID="2" stepID="2" isGroundBlock="0"/>
        + <stepContent InstructionType="ClearText" stepContentID="3" stepID="2" isGroundBlock="0" InstructionText="Install bracket 1 to stowage
          location on TCU door (check screw size for correct orientation, 5 mm screws to upper position). [M5 (4 mm) Hex Head Driver 1/4" Drive] and [M6 (5
          mm) Hex Head Driver 1/4" Drive]"></stepContent>
        + <stepContent InstructionType="ClearText" stepContentID="4" stepID="2" isGroundBlock="0" InstructionText="Unscrew and remove TCU launch
          bracket 2 (upper left) [four captive screws, Ratchet Wrench 1/4"; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive (4 mm
          for right screws and 5 mm for left screws)]."></stepContent>
          <stepContent imageID="2" stepContentID="5" stepID="2" isGroundBlock="0"/>
        + <stepContent InstructionType="ClearText" stepContentID="6" stepID="2" isGroundBlock="0" InstructionText="Install bracket 2 to stowage
          location on TCU door (check screw size for correct orientation, 5 mm screws to upper position). [M5 (4 mm) Hex Head Driver 1/4" Drive] and [M6 (5
          mm) Hex Head Driver 1/4" Drive]"></stepContent>
        </step>
      + <step stepParentID="0" stepID="3" isGroundStep="0" stepNumber="3" stepTitle="REMOVING BRACKETS OF DOOR HINGE SIDE"></step>
      + <step stepParentID="0" stepID="4" isGroundStep="0" stepNumber="4" stepTitle="CLEANUP"></step>
      </Steps>
    </PODFAR>
  </ARPF>
```

The core elements, the belonging attributes and cardinalities of the AR Procedure File (ARPF) format used to exchange digital content between the MARSOP Constructor and the Viewer:



Appendix C

MARSOP Study: Tasks, Questionnaire and Skewed Distributions

This appendix is related to the proof-of-concept study presented in section 6.2 and shows the procedures used for the experiment tasks, their provided resources and the post-questionnaire that the participants had to fill in. It also listed the comments of the participants about their likes and dislikes. Finally, this appendix shows box plots for skewed distributions of the completion time.

Biolab ODF instructions of the experiment task T1

2.510 BIOLAB EXPERIMENT CONTAINER INSTALLATION

OBJECTIVE:

This is a generic procedure for installing experiment containers (EC) to the Biolab Centrifuge Rotors inside the Incubator. The procedure can be coupled with PODF 2.511 (BIOLAB EXPERIMENT CONTAINER REMOVAL) if necessary.

CREW

One

DURATION

15 minutes

TOOLS

Columbus Tool Kit

Tool Bag 2:

Flat Screw Driver, 4 mm

ITEMS

Mini-Maglite1 P/N 528-20084-5

Mini-Maglite2 P/N 528-20084-5

REFERENCED PROCEDURES

ESA SODF; JOINT SYS-PL: NOMINAL; 2.201 BIOLAB SMOKE DETECTION MONITORING CONTROL

1. GETTING EC FROM TCU

COL1A2_G2	1.1	Open completely TCU 2 Door (press latch on inner right side of dorr lever)
COL1A2_G2	1.2	Get marked EC from TCU 2 Tray
COL1A2_G2	1.3	Put EC on table
COL1A2_G2	1.4	Close TCU 2 Door

2. SAFETY INSPECTION OF INCUBATOR WORKING VOLUME

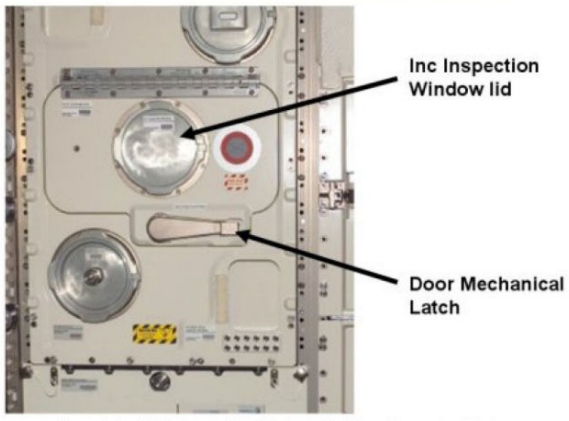


Figure 1. Inc EC Exchange Door Mechanical Latch and Inspection Window

COL1A2_E1	2.1	Inc Inspection Window lid – press latch and turn until removable (ref. to fig. 1)
COL1A2_E1	2.2	Remove and tempstow the Inc Inspection Window lid
COL1A2_E1	2.3	Verify rotors (two) – no rotation (visually)
COL1A2_E1	2.4	Reinstall Inc Inspection Window lid

3. OPENING EC EXCHANGE DOOR VIA MANUAL OVERRIDE

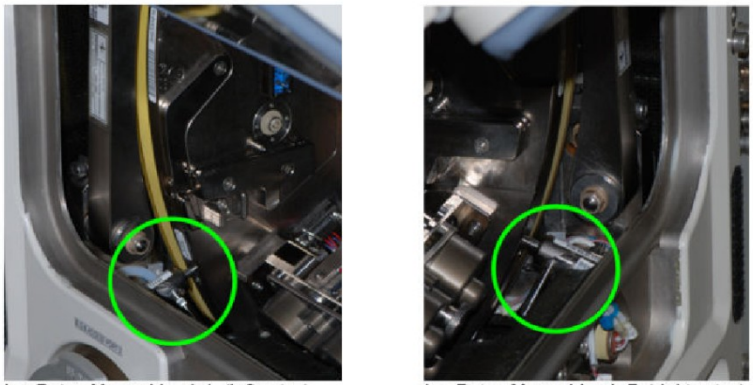
CAUTION
Do not try to rotate door lever without manual override in place to prevent damage to locking mechanism.

Figure 2. EC Exchange Door manual override

COL1A2_E1	3.1	Press Door Mechanical Latch to release door lever (ref. to fig. 1 or 2)
COL1A2_E1	3.2	Keep pushing the manual override to the right using the screwdriver and rotate door lever 90°↷ (Flat Screwdriver 4 mm, ref. to fig. 2)
COL1A2_E1	3.3	Open EC Exchange Door

4. INSTALLING THE EXPERIMENT CONTAINERS

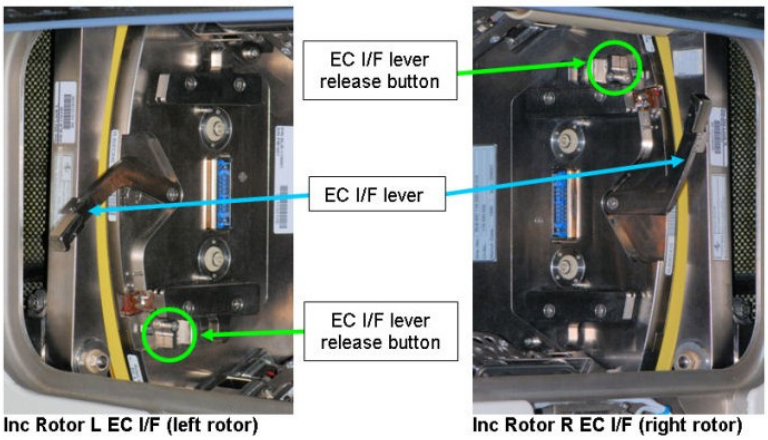
COL1A2_E1	4.1	Move rotor manually to the position EC 3



Inc Rotor Manual Lock L (left rotor) Inc Rotor Manual Lock R (right rotor)

Figure 4. Inc Rotor Manual Locks of both Rotors

COL1A2_E1	4.2	Engage Inc Rotor Manual Lock R to fix rotor (ref. to fig. 4)



Inc Rotor L EC I/F (left rotor) **Inc Rotor R EC I/F (right rotor)**

Figure 5. Inc Rotor EC I/Fs for both Rotors

COL1A2_E1	4.3	Push EC I/F lever release button to release EC I/F lever (ref. to fig. 5)
COL1A2_E1	4.4	Pull EC I/F lever away from Rotor to open position (ref. to fig. 5)
COL1A2_E1	4.5	Verify electrical connectors of EC – not damaged
COL1A2_E1	4.6	Slide EC into rotor's guiding interface (ref. to fig. 6)
COL1A2_E1	4.7	Push EC I/F lever towards rotor to close (audible click)
COL1A2_E1	4.8	✓EC I/F lever – fixed (by pulling EC I/F lever (which should not move))
COL1A2_E1	4.9	Disengage Inc Rotor Manual Lock R

5. CLOSING INC EC EXCHANGE DOOR

COL1A2_E1	5.1	Verify Inc EC Exchange Door Seal – for integrity and correct seat (visually)
COL1A2_E1	5.2	Close EC Exchange Door
COL1A2_E1	5.3	Rotate Door lever 90°↶, engage at mechanical latch (do not move latch and door lever anymore after locking the door)

6. CLEANUP
Tools and items → stow according to stowage note

Biolab ODF instructions of the experiment task T2

1.103 BIOLAB TCUS LAUNCH FIXATIONS DISENGAGEMENT

OBJECTIVE:

To disengage TCUs 1 and 2 Launch Fixations during commissioning before first switch on of Biolab.

CREW

One

DURATION:

15 minutes

TOOLS:

Columbus Tool Kit:

Tool Bag 1:

- Ratchet Wrench 1/4"
- Ratchet Tool 1/4" Drive
- 102 mm Extension 1/4" Drive
- Torque Wrench 4-20 Nm

Tool Bag 3:

- M5 (4 mm) Hex Head Driver 1/4" Drive
- M6 (5 mm) Hex Head Driver 1/4" Drive

1. VERIFYING POWER OFF

> NOTE
> Loose screws sequentially only a few turns each screw until all screws disengaged.

COL1A2_G1 Verify Biolab Rack POWER sw – OFF

2. REMOVING BRACKETS OF TCU DOOR LATCH SIDE

COL1A2_E2 Remove brackets of TCU Door latch side

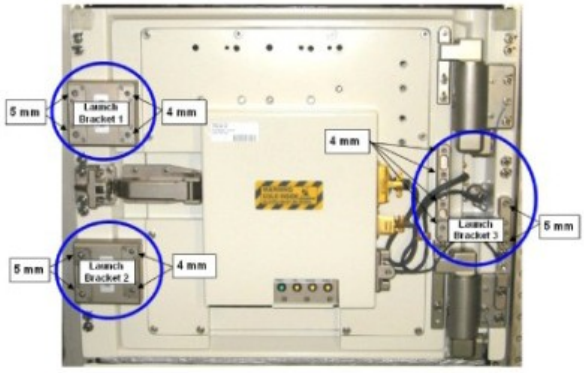


Figure 1. TCU Launch Brackets in Launch Configuration

Unscrew and remove TCU launch bracket 1 (upper left) [four captive screws, Ratchet Wrench 1/4"; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive (4 mm for right screws and 5 mm for left screws)].

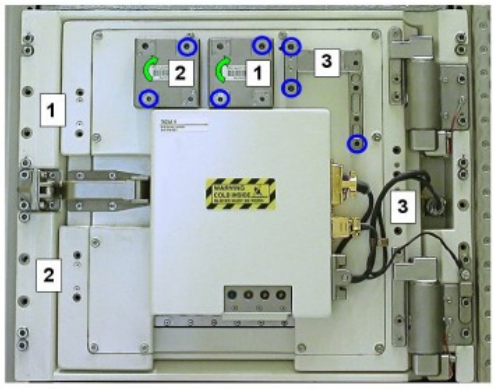


Figure 2. TCU Launch Brackets Stowage Locations After Removal

Install bracket to stowage location on TCU door (check screw size for correct orientation, 5 mm screws to upper position).

Repeat for TCU Launch Bracket 2 (lower left).

3. REMOVING BRACKETS OF DOOR HINGE SIDE

Unscrew and remove TCU Launch Bracket 3 [six captive screws, Ratchet Wrench 1/4"; 102 mm Extension 1/4" Drive; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive].

Install bracket to support location TCU door.

4. CLEANUP

Stow tools and items

Resources provided for task T1

Task T1: BIOLAB EXPERIMENT CONTAINER (EC) INSTALLATION						
Step	Instruction	Sources of information				
		image	video	label	3D object	highlighted geometry
1.	GETTING EC FROM TCU					
1.1	Open completely TCU 2 Door (press latch on inner right side of door lever)	-	1 (10 s)	1	-	4
1.2	Get marked EC from TCU 2 Tray	-	-	1	-	-
1.3	Put EC on table	-	-	-	-	-
1.4	Close TCU 2 Door	-	-	-	-	1
2.	SAFETY INSPECTION OF INCUBATOR WORKING VOLUME					
2.1	Inc Inspection Window lid → press latch and turn until removable (ref. to fig. 1)	1	1 (10 s)	1	-	1
2.2	Remove and tempstow the Inc Inspection Window lid	-	-	1	-	1
2.3	Verify rotors (two) – no rotation (visually)	-	-	-	-	1
2.4	Reinstall Inc Inspection Window lid	-	-	-	1	1
3.	OPENING EC EXCHANGE DOOR VIA MANUAL OVERRIDE					
3.1	Press Door Mechanical Latch to release door lever (ref. to fig. 2)	1	1 (5 s)	2	-	2
3.2	Keep pushing the manual override to the right using the screwdriver and rotate door lever 90° ↻ (Flat Screwdriver 4 mm, ref. to fig. 2)	-	1 (13 s)	1	1	-
3.3	Open EC Exchange Door	-	-	-	-	2
4.	INSTALLING THE EXPERIMENT CONTAINERS					
4.1	Move rotor manually to the position EC 3	-	1 (4 s)	-	-	-
4.2	Engage Inc Rotor Manual Lock R to fix rotor (ref. to fig. 4)	1	1 (4 s)	-	-	-

4.3	Push EC I/F lever release button to release EC I/F lever (ref. to fig. 5)	-	1 (5 s)	-	-	-
4.4	Pull EC I/F lever away from Rotor to open position (ref. to fig. 5)	-	1 (5 s)	-	-	-
4.5	Verify electrical connectors of EC – not damaged	-	-	-	-	-
4.6	Slide EC into rotor's guiding interface (ref. to fig. 6)	1	1 (12 s)	-	-	-
4.7	Push EC I/F lever towards rotor to close (audible click)	-	1 (9 s)	-	-	-
4.8	✓EC I/F lever – fixed (by pulling EC I/F lever (which should not move))	-	-	-	-	-
4.9	Disengage Inc Rotor Manual Lock R	-	1 (7 s)	-	-	-
5.	CLOSING INC EC EXCHANGE DOOR					
5.1	Verify Inc EC Exchange Door Seal – for integrity and correct seat (visually)	-	-	-	-	-
5.2	Close EC Exchange Door	-	-	1	-	1
5.3	Rotate Door lever 90° ↻, engage at mechanical latch (do not move latch and door lever anymore after locking the door)	-	-	1	1	1

Resources provided for task T2

Task T2: BIOLAB TCUS LAUNCH FIXATIONS DISENGAGEMENT						
Step	Instruction	Sources of information				
		image	video	label	3D object	highlighted geometry
1.	VERIFYING POWER OFF					
1.1	Verify Biolab Rack POWER sw – OFF	-	-	1	-	2
2.	REMOVING BRACKETS OF TCU DOOR LATCH SIDE					
2.1	Unscrew and remove TCU launch bracket 1 (upper left) [four captive screws, Ratchet Wrench 1/4"; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive (4 mm for right screws and 5 mm for left screws)].	1	1 (45 s)	1	-	-
2.2	Install bracket to stowage location on TCU door (check screw size for correct orientation, 5 mm screws to upper position). Repeat for TCU Launch Bracket 2.	1	1 (48 s)	1	-	1
3.	REMOVING BRACKETS OF DOOR HINGE SIDE					
3.1	Unscrew and remove TCU Launch Bracket 3 [six captive screws, Ratchet Wrench 1/4"; 102 mm Extension 1/4" Drive; M5 (4 mm) Hex Head Driver 1/4" Drive; M6 (5 mm) Hex Head Driver 1/4" Drive].	1	1 (59 s)	1	-	-
3.2	Install bracket to support location TCU door.	1	1 (57 s)	1	-	1

Post questionnaire

HMD+Tracking	*strongly disagree*				*strongly agree*
H1. I was satisfied with the HMD wearing comfort.	☐	☐	☐	☐	☐
H2. Seeing the environment by video images was confusing for me.	☐	☐	☐	☐	☐
H3. I noticed latencies.	☐	☐	☐	☐	☐
H4. The latencies were annoying for me.	☐	☐	☐	☐	☐
H5. I did not get the impression of being in the physical reality anymore.	☐	☐	☐	☐	☐

Visual Feedback	*strongly disagree*				*strongly agree*
V1. The HUD information were clearly.	☐	☐	☐	☐	☐
V2. I was satisfied with the quality of the HUD information.	☐	☐	☐	☐	☐
V3. The tracking frame feedback was helpfully for me.	☐	☐	☐	☐	☐
V4. The 3D information (labels, etc.) had confused me.	☐	☐	☐	☐	☐
V5. I used the 3D information (labels, etc.).	☐	☐	☐	☐	☐

Resources	*strongly disagree*				*strongly agree*
R1. The Resource Pad was helpfully for me.	☐	☐	☐	☐	☐
R2. I could handle the Resource Pad easily.	☐	☐	☐	☐	☐
R3. I was not satisfied while using the Resource Pad.	☐	☐	☐	☐	☐

System/Control (SUS, Brooke, 1996)	*strongly disagree*				*strongly agree*
C1. I thought the system was easy to use.	☐	☐	☐	☐	☐
C2. I found the system unnecessarily complex.	☐	☐	☐	☐	☐
C3. I found the various functions in this system were very well integrated.	☐	☐	☐	☐	☐
C4. I felt very confident using the system.	☐	☐	☐	☐	☐
C5. Controlling the system was very cumbersome.	☐	☐	☐	☐	☐

Preference

I would prefer to use the system compared to the conventional system: ☐ yes ☐ no

Miscellaneous

What did you like?

_____

_____

_____

What did you dislike?

_____

_____

_____

Comments:

_____

_____

_____

Comments of participants about the MARSOP system

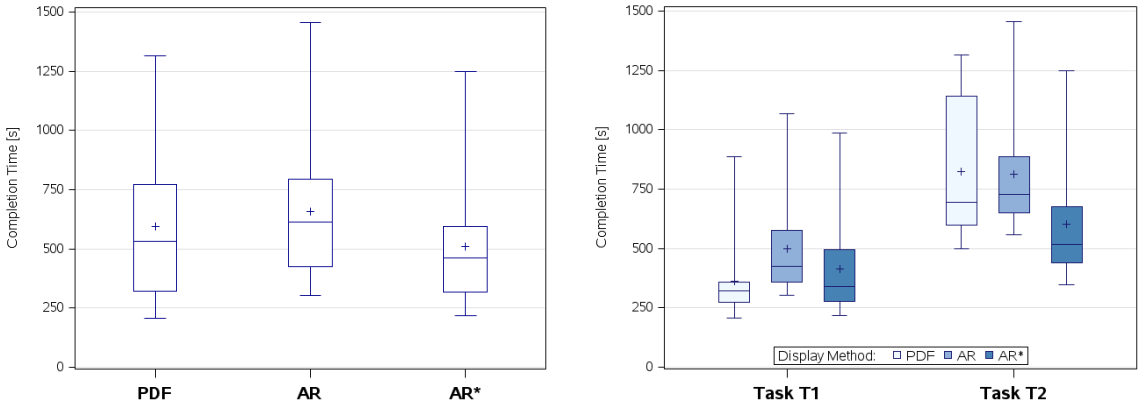
What did the participants LIKE?:
"3D information was very beneficial."
"3D labels make searching of location easy and make work faster."
"3D labels are very useful and straight forward."
"Easy searching for locations and items."
"Blending instruction text allows step-by-step-working."
"The overlap of the instructions and the real world sight."
"Using the Resource Pad and the 3D labels."
"Easy availability of the Resource Pad."
"Using the video enable faster solving the task."
"The integration of video."
"Pictures/videos for clarification."
"Audio instruction of the videos."
"Voice guidance was helpful where the video was out of the sight."
"Voice instruction are helpful."
"Very easy to navigate."
"Going forward to next step by speech command."
"Step-by-step by voice commanding (easier than with laptop)."

What did the participants DISLIKE?:
"HMD comfort, but is adaptable."
"Seeing the real world through the HMD."
"Limited field of view."
"Field of view is to narrow."
"Latency in video display."
"Feedback delay."
"Reduced contrast/color in video display."
"Quality/definition of the image on display does not always allow for recognition details "
"No control of the video resources."

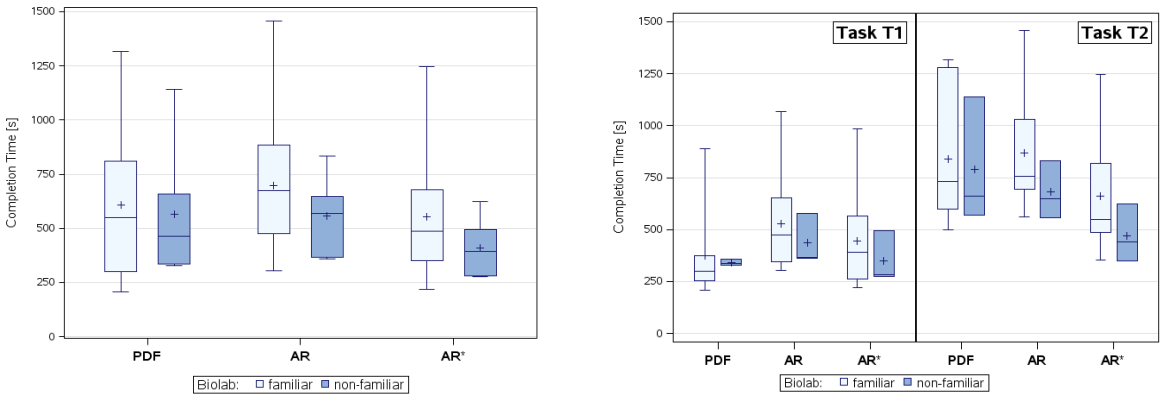
General comments:
"Very promising."
"3D labels and videos make the work easier."
"The AR system is very helpful while performing unknown tasks."
"The information in overlays is very helpful; it is the major advantage over conventional systems."
"There should be a hint, that 3D information or resources are available."
"Tracking was good, but the image quality makes the experience difficult."
"I would prefer HMDs where the environment is not bridged through a video."
"OST HMD would be better."
"Google should be adjustable in distance."
"The latency is not the problem, but an adaption to rely more on the visual perception than the joint positioning is necessary."
"Using of ear plugs suggested."

Skewed distributions of the completion time

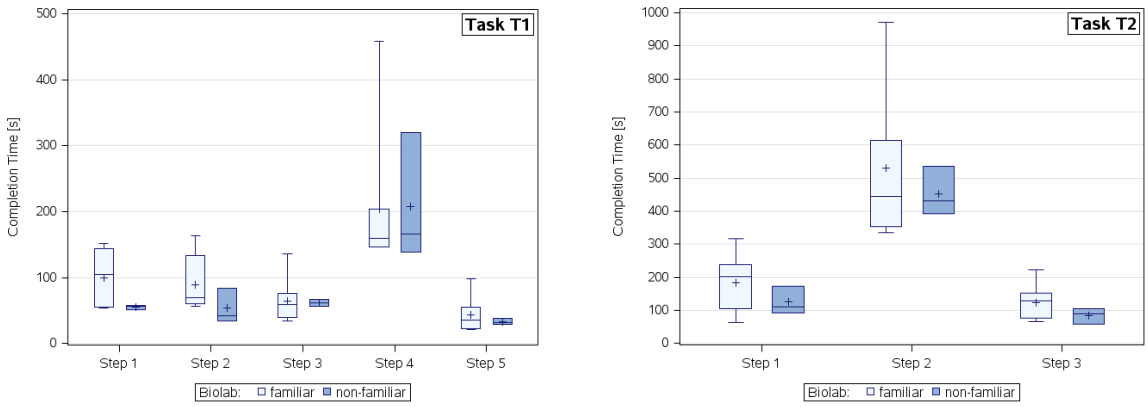
Distribution of completion time of the display method, *(left)* across both task, and *(right)* per experiment task:



Distribution of completion time of the display method by participants' Biolab experiences, *(left)* across both task, and *(right)* per experiment task:



Distribution of completion time of the AR procedure steps of the experiment tasks by participants' Biolab experiences:



Appendix D

ARGuide Study: Descriptive and Inferential Statistics

This appendix is related to the outcomes of the experiment presented in Chapter 5 and provides descriptive statistics, showing measures of the central tendency and the variability of the dependent variables. It additionally presents the measures and results of the comparison of the LF/HF ratio between the display conditions related to the physiological workload.

Visuomotor Task: Descriptive statistics per trial type, display technique (ED, HUD, AR) and spatial navigation strategy

Measures of the *session completion time* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	6	101.51	101.36	9.61	3.92	87.08	114.75
		egocentric	7	115.18	107.52	15.73	5.95	100.84	145.18
		total	**13**	**108.87**	**107.36**	**14.58**	**4.04**	**87.08**	**145.18**
	HUD	exocentric	6	99.46	98.50	6.66	2.72	90.33	107.43
		egocentric	7	92.05	92.99	10.80	4.08	71.34	105.43
		total	**13**	**95.47**	**95.96**	**9.57**	**2.66**	**71.34**	**107.43**
	AR	exocentric	6	89.95	84.94	12.67	5.17	79.77	114.28
		egocentric	7	83.68	81.80	12.55	4.74	63.82	103.21
		total	**13**	**86.58**	**83.74**	**12.50**	**3.47**	**63.82**	**114.28**
	total		39	96.97	96.09	15.22	2.44	63.82	145.18
dual-task	**ED**	exocentric	6	121.84	115.20	23.65	9.66	96.98	151.45
		egocentric	7	125.94	124.19	16.56	6.26	106.92	148.60
		total	**13**	**124.05**	**118.46**	**19.36**	**5.37**	**96.98**	**151.45**
	HUD	exocentric	6	104.79	106.28	15.39	6.28	78.26	124.54
		egocentric	7	109.68	111.38	15.34	5.80	93.39	136.37
		total	**13**	**107.43**	**108.00**	**14.92**	**4.14**	**78.26**	**136.37**
	AR	exocentric	6	103.48	100.78	19.89	8.12	81.68	132.83
		egocentric	7	91.25	87.34	10.29	3.89	79.95	106.72
		total	**13**	**96.89**	**96.13**	**16.06**	**4.46**	**79.95**	**132.83**
	total		39	109.45	106.92	19.96	3.20	78.26	151.45

Measures of the *task completion time* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	72	8.42	8.02	2.93	0.34	3.49	16.07
		egocentric	84	9.55	8.60	3.61	0.39	3.33	21.76
		total	**156**	**9.03**	**8.39**	**3.35**	**0.27**	**3.33**	**21.76**
	HUD	exocentric	72	8.25	7.53	3.03	0.36	3.69	16.80
		egocentric	84	7.63	7.44	2.30	0.25	3.78	15.65
		total	**156**	**7.91**	**7.44**	**2.67**	**0.21**	**3.69**	**16.80**
	AR	exocentric	72	7.45	6.82	2.81	0.33	3.54	15.33
		egocentric	84	6.93	6.47	2.96	0.32	3.09	21.32
		total	**156**	**7.17**	**6.71**	**2.89**	**0.23**	**3.09**	**21.32**
	total		468	8.04	7.50	3.07	0.14	3.09	21.76
dual-task	**ED**	exocentric	72	10.11	9.41	3.99	0.47	3.57	20.19
		egocentric	84	10.44	10.13	3.39	0.37	4.80	18.51
		total	**156**	**10.29**	**9.81**	**3.67**	**0.29**	**3.57**	**20.19**
	HUD	egocentric	72	8.69	8.15	3.11	0.37	3.94	18.46
		egocentric	84	9.08	8.38	3.39	0.37	4.06	19.56
		total*	**156**	**8.90**	**8.32**	**3.26**	**0.26**	**3.94**	**19.56**
	AR	exocentric	72	8.58	7.78	3.90	0.46	3.41	24.12
		egocentric	84	7.56	7.13	2.62	0.29	3.70	15.41
		total	**156**	**8.03**	**7.42**	**3.30**	**0.26**	**3.41**	**24.12**
	total		468	9.07	8.40	3.53	0.16	3.41	24.12

Measures of the *search time* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	72	5.04	4.95	1.24	0.15	2.64	9.92
		egocentric	84	5.87	5.79	1.55	0.17	2.48	12.67
		total	**156**	**5.49**	**5.49**	**1.47**	**0.12**	**2.48**	**12.67**
	HUD	exocentric	72	4.23	3.81	1.66	0.20	2.56	13.75
		egocentric	84	4.09	3.97	1.13	0.12	1.94	9.09
		total	**156**	**4.15**	**3.85**	**1.40**	**0.11**	**1.94**	**13.75**
	AR	exocentric	72	3.81	3.62	1.02	0.12	2.00	8.36
		egocentric	84	3.56	3.48	1.13	0.12	1.60	6.79
		total	**156**	**3.67**	**3.59**	**1.08**	**0.09**	**1.60**	**8.36**
	total		468	4.44	4.13	1.53	0.07	1.60	13.75
dual-task	**ED**	exocentric	72	6.38	6.02	2.13	0.25	2.33	11.95
		egocentric	84	6.52	6.21	2.00	0.22	2.43	13.28
		total	**156**	**6.45**	**6.17**	**2.05**	**0.16**	**2.33**	**13.28**
	HUD	exocentric	72	4.36	4.01	1.38	0.16	2.50	10.50
		egocentric	84	4.58	4.23	1.86	0.20	2.40	12.35
		total	**156**	**4.48**	**4.10**	**1.66**	**0.13**	**2.40**	**12.35**
	AR	exocentric	72	4.59	4.25	1.97	0.23	2.00	13.31
		egocentric	84	4.02	3.56	1.60	0.17	1.54	10.61
		total	**156**	**4.28**	**3.89**	**1.80**	**0.14**	**1.54**	**13.31**
	total		468	5.07	4.67	2.09	0.10	1.54	13.31

Measures of the *operation time* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	72	3.36	2.73	2.29	0.27	0.67	9.29
		egocentric	84	3.67	3.18	2.76	0.30	0.51	12.87
		total	**156**	**3.53**	**2.97**	**2.55**	**0.20**	**0.51**	**12.87**
	HUD	exocentric	72	4.00	3.02	2.36	0.28	1.07	11.90
		egocentric	84	3.53	2.92	1.97	0.21	0.72	7.72
		total	**156**	**3.75**	**3.00**	**2.17**	**0.17**	**0.72**	**11.90**
	AR	exocentric	72	3.63	3.76	2.50	0.29	0.83	10.43
		egocentric	84	3.36	2.95	2.39	0.26	0.84	15.07
		total	**156**	**3.49**	**3.37**	**2.44**	**0.20**	**0.83**	**15.07**
	total		468	3.59	3.07	2.39	0.11	0.51	15.07
dual-task	**ED**	exocentric	72	3.72	3.35	2.59	0.30	0.88	12.14
		egocentric	84	3.91	3.67	2.50	0.27	0.76	10.45
		total	**156**	**3.82**	**3.51**	**2.53**	**0.20**	**0.76**	**12.14**
	HUD	exocentric	72	4.32	3.67	2.65	0.31	0.82	11.82
		egocentric	84	4.49	4.29	2.47	0.27	0.90	12.18
		total	**156**	**4.41**	**3.90**	**2.54**	**0.20**	**0.82**	**12.18**
	AR	exocentric	72	3.98	3.57	2.63	0.31	0.92	11.08
		egocentric	84	3.53	3.03	2.07	0.23	0.94	9.13
		total	**156**	**3.74**	**3.11**	**2.35**	**0.19**	**0.92**	**11.08**
	total		468	3.99	3.57	2.49	0.12	0.76	12.18

Measures of the *operation time (single-pointing)* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	36	1.57	1.35	0.77	0.13	0.67	4.46
		egocentric	42	1.92	1.48	1.50	0.23	0.51	7.44
		total	**78**	**1.76**	**1.38**	**1.23**	**0.14**	**0.51**	**7.44**
	HUD	exocentric	36	2.20	1.92	0.89	0.15	1.07	5.42
		egocentric	42	1.94	1.92	0.69	0.11	0.72	3.79
		total	**78**	**2.06**	**1.92**	**0.80**	**0.09**	**0.72**	**5.42**
	AR	exocentric	36	1.72	1.38	1.13	0.19	0.83	6.71
		egocentric	42	1.61	1.42	0.81	0.13	0.84	5.10
		total	**78**	**1.66**	**1.40**	**0.96**	**0.11**	**0.83**	**6.71**
	total		234	1.83	1.56	1.02	0.07	0.51	7.44
dual-task	**ED**	exocentric	36	1.90	1.44	1.11	0.19	0.88	5.46
		egocentric	42	2.02	1.66	1.10	0.17	0.76	5.02
		total	**78**	**1.96**	**1.50**	**1.10**	**0.12**	**0.76**	**5.46**
	HUD	exocentric	36	2.32	2.16	1.02	0.17	0.82	4.50
		egocentric	42	2.55	2.19	1.13	0.17	0.90	5.89
		total	**78**	**2.44**	**2.16**	**1.08**	**0.12**	**0.82**	**5.89**
	AR	exocentric	36	1.81	1.60	0.79	0.13	0.92	4.78
		egocentric	42	1.76	1.72	0.60	0.09	0.94	2.96
		total	**78**	**1.78**	**1.61**	**0.69**	**0.08**	**0.92**	**4.78**
	total		234	2.06	1.75	1.01	0.07	0.76	5.89

Measures of the *operation time (multiple directional-pointing)* [s]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	36	5.16	4.83	1.85	0.31	2.33	9.29
		egocentric	42	5.41	4.90	2.63	0.41	2.25	12.87
		total	**78**	**5.30**	**4.86**	**2.29**	**0.26**	**2.25**	**12.87**
	HUD	exocentric	36	5.80	5.52	1.97	0.33	2.35	11.90
		egocentric	42	5.13	5.30	1.47	0.23	2.61	7.72
		total	**78**	**5.44**	**5.40**	**1.74**	**0.20**	**2.35**	**11.90**
	AR	exocentric	36	5.55	4.80	1.97	0.33	3.61	10.43
		egocentric	42	5.11	4.58	2.15	0.33	2.67	15.07
		total	**78**	**5.31**	**4.67**	**2.07**	**0.23**	**2.67**	**15.07**
	total		234	5.35	4.95	2.04	0.13	2.25	15.07
dual-task	**ED**	exocentric	36	5.55	4.92	2.34	0.39	2.70	12.14
		egocentric	42	5.80	5.18	2.03	0.31	2.43	10.45
		total	**78**	**5.68**	**5.12**	**2.17**	**0.25**	**2.43**	**12.14**
	HUD	exocentric	36	6.33	5.80	2.21	0.37	2.61	11.82
		egocentric	42	6.43	6.24	1.83	0.28	3.26	12.18
		total	**78**	**6.38**	**6.17**	**2.00**	**0.23**	**2.61**	**12.18**
	AR	exocentric	36	6.15	5.70	1.92	0.32	2.93	11.08
		egocentric	42	5.30	5.16	1.37	0.21	3.09	9.13
		total	**78**	**5.69**	**5.31**	**1.69**	**0.19**	**2.93**	**11.08**
	total		234	5.92	5.55	1.98	0.13	2.43	12.18

Measures of the *Euclidean pointing distance* [mm]:

Trial Type	Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**ED**	exocentric	72	7.91	7.49	2.68	0.32	0.42	12.63
		egocentric	84	7.35	7.48	2.40	0.26	1.74	11.98
		total	**156**	**7.61**	**7.49**	**2.54**	**0.20**	**0.42**	**12.63**
	HUD	exocentric	72	7.60	7.48	2.81	0.33	0.42	12.05
		egocentric	84	8.21	7.82	2.64	0.29	1.44	11.99
		total	**156**	**7.93**	**7.55**	**2.73**	**0.22**	**0.42**	**12.05**
	AR	exocentric	72	7.69	7.50	2.77	0.33	1.25	12.05
		egocentric	84	7.58	7.48	2.66	0.29	0.71	12.28
		total	**156**	**7.63**	**7.49**	**2.70**	**0.22**	**0.71**	**12.28**
	total		468	7.72	7.49	2.66	0.12	0.42	12.63
dual-task	**ED**	exocentric	72	7.81	7.55	2.94	0.35	0.92	12.74
		egocentric	84	7.34	7.48	3.06	0.33	0.20	11.98
		total	**156**	**7.56**	**7.48**	**3.01**	**0.24**	**0.20**	**12.74**
	HUD	exocentric	72	7.34	7.48	2.58	0.30	1.22	11.98
		egocentric	84	7.89	7.49	2.72	0.30	1.71	11.99
		total	**156**	**7.63**	**7.49**	**2.66**	**0.21**	**1.22**	**11.99**
	AR	exocentric	72	7.80	7.48	3.05	0.36	0.64	12.57
		egocentric	84	7.91	7.49	2.71	0.30	1.94	12.48
		total	**156**	**7.86**	**7.49**	**2.86**	**0.23**	**0.64**	**12.57**
	total		468	7.68	7.49	2.85	0.13	0.20	12.74

Reaction-Time Task: Descriptive statistics of response rate and time per display technique (ED, HUD, AR) and spatial navigation strategy

Measures of the *response rate* [%]:

Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
ED	exocentric	6	91.26	94.48	12.09	4.94	68.00	100.00
	egocentric	7	95.16	97.06	5.59	2.11	85.71	100.00
	total	**13**	**93.36**	**97.06**	**8.98**	**2.49**	**68.00**	**100.00**
HUD	exocentric	6	96.38	98.44	4.71	1.92	88.89	100.00
	egocentric	7	92.24	97.22	11.08	4.19	70.45	100.00
	total	**13**	**94.15**	**97.22**	**8.67**	**2.41**	**70.45**	**100.00**
AR	exocentric	6	98.99	100.00	2.47	1.01	93.94	100.00
	egocentric	7	91.77	96.43	10.00	3.78	72.73	100.00
	total	**13**	**95.10**	**100.00**	**8.16**	**2.26**	**72.73**	**100.00**
total		39	94.20	97.22	8.41	1.35	68.00	100.00

Measures of the *response time by task* [ms]:

Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
ED	exocentric	210	735.25	604.00	475.80	32.83	42.00	3863.00
	egocentric	194	993.67	804.00	627.27	45.04	41.00	3747.00
	total	**404**	**859.34**	**685.00**	**567.94**	**28.26**	**41.00**	**3863.00**
HUD	exocentric	192	693.80	565.00	407.08	29.38	81.00	3079.00
	egocentric	165	863.33	687.00	494.52	38.50	38.00	3662.00
	total	**357**	**772.15**	**641.00**	**456.87**	**24.18**	**38.00**	**3662.00**
AR	egocentric	192	666.02	565.00	333.69	24.08	42.00	2072.00
	exocentric	129	989.35	766.00	613.25	53.99	161.00	3821.00
	total	**321**	**795.96**	**645.00**	**492.04**	**27.46**	**42.00**	**3821.00**
total		1082	811.77	647.00	511.95	15.56	38.00	3863.00

Measures of the *response time by search* [ms]:

Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
ED	exocentric	135	735.93	604.00	453.36	39.02	43.00	2739.00
	egocentric	123	991.10	804.00	613.70	55.34	41.00	3747.00
	total	**258**	**857.58**	**684.00**	**549.78**	**34.23**	**41.00**	**3747.00**
HUD	exocentric	98	649.53	593.00	305.36	30.85	81.00	2014.00
	egocentric	90	841.06	685.00	467.17	49.24	38.00	2537.00
	total	**188**	**741.22**	**633.50**	**401.80**	**29.30**	**38.00**	**2537.00**
AR	exocentric	106	702.85	605.50	355.06	34.49	317.00	2072.00
	egocentric	64	1083.98	847.00	706.93	88.37	282.00	3821.00
	total	**170**	**846.34**	**649.00**	**546.74**	**41.93**	**282.00**	**3821.00**
total		616	818.96	664.50	510.12	20.55	38.00	3821.00

Measures of the *response time by operation* [ms]:

Display	Spatial Strategy	N	Mean	Median	Stddev	Stderr	Min	Max
ED	exocentric	75	734.04	603.00	516.93	59.69	42.00	3863.00
	egocentric	71	998.13	768.00	654.54	77.68	382.00	3624.00
	total	**146**	**862.47**	**685.00**	**600.61**	**49.71**	**42.00**	**3863.00**
HUD	exocentric	94	739.95	525.00	488.75	50.41	321.00	3079.00
	egocentric	75	890.05	723.00	527.40	60.90	363.00	3662.00
	total	**169**	**806.56**	**643.00**	**510.24**	**39.25**	**321.00**	**3662.00**
AR	exocentric	86	620.63	526.00	301.16	32.47	42.00	1832.00
	egocentric	65	896.17	725.00	492.38	61.07	161.00	3362.00
	total	**151**	**739.24**	**643.00**	**416.62**	**33.90**	**42.00**	**3362.00**
total		466	802.26	644.50	514.76	23.85	42.00	3863.00

Physiological Workload: Statistic on the LF/HF ratio per trial type and display technique (ED, HUD, AR)

Measures of the *LF/HF ratio*:

Trial Type	Display	N	Mean	Median	Stddev	Stderr	Min	Max
single-task	**AR**	11	-3.98	-1.86	5.95	1.79	-18.10	2.39
	ED	11	-3.71	-1.61	6.29	1.90	-17.95	2.60
	HUD	11	-3.99	-1.91	6.58	1.98	-18.49	2.64
	total	33	-3.89	-1.86	6.08	1.06	-18.49	2.64
dual-task	**AR**	11	-3.46	-2.04	6.37	1.92	-18.03	3.66
	ED	11	-2.61	-0.66	6.06	1.83	-14.72	4.72
	HUD	11	-5.31	-7.13	6.12	1.85	-18.05	2.63
	total	33	-3.79	-1.79	6.10	1.06	-18.05	4.72

Comparison of the *LF/HF ratio* between the display conditions at the same stage of trial type using SAS® PROC MIXED (adjust=simulate):

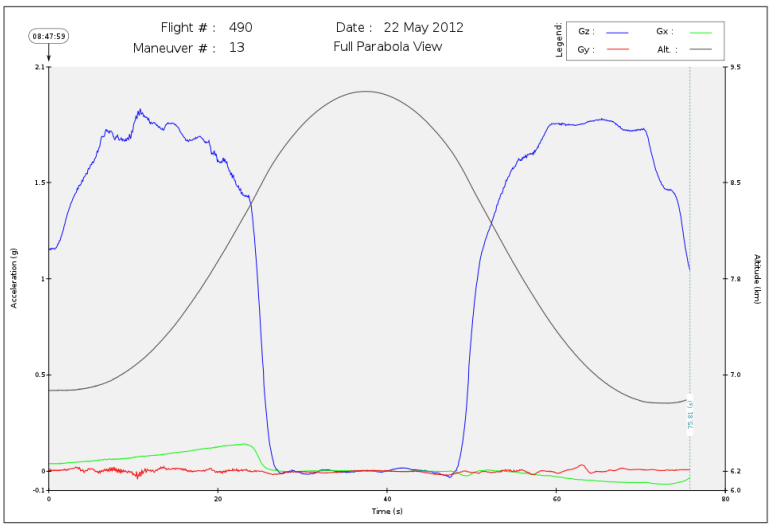
Trial Type	Display		*DF*	*t Value*	*Pr > |t|*	*Adj. p-Val*
single-task	AR	ED	30	-0.10	0.9203	0.9942
	AR	HUD	30	0.00	0.9976	1.0000
	ED	HUD	30	0.10	0.9179	0.9937
dual-task	AR	ED	30	-0.32	0.7500	0.9470
	AR	HUD	30	0.70	0.4891	0.7637
	ED	HUD	30	1.02	0.3150	0.5653

Appendix E

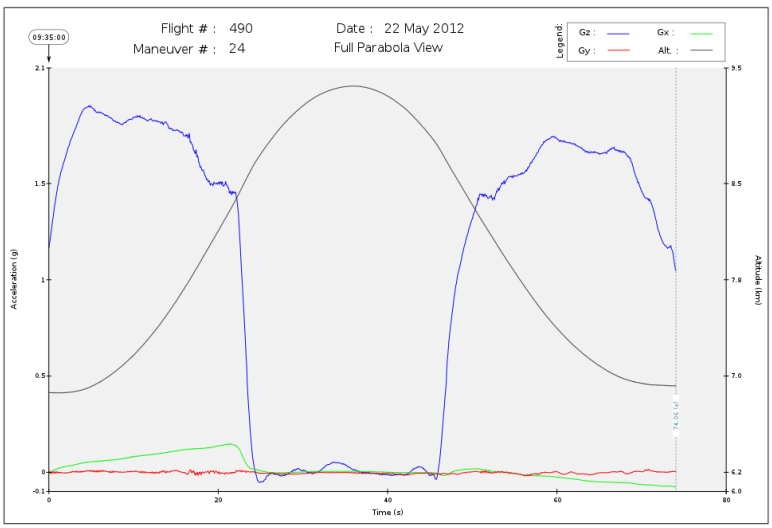
3DPick Study: Sample Parabolas and Experiment Installation Schematics

This appendix is related to the parabolic flight experiments presented in Chapter 8 and provides samples for resulted parabola sequences and the installation schematics of the flight racks installed aboard the Zero-G Airbus A300.

Samples of resulted parabola sequences during the 56th PFC showing qualitative deviations in flight performance (provided by Novespace)



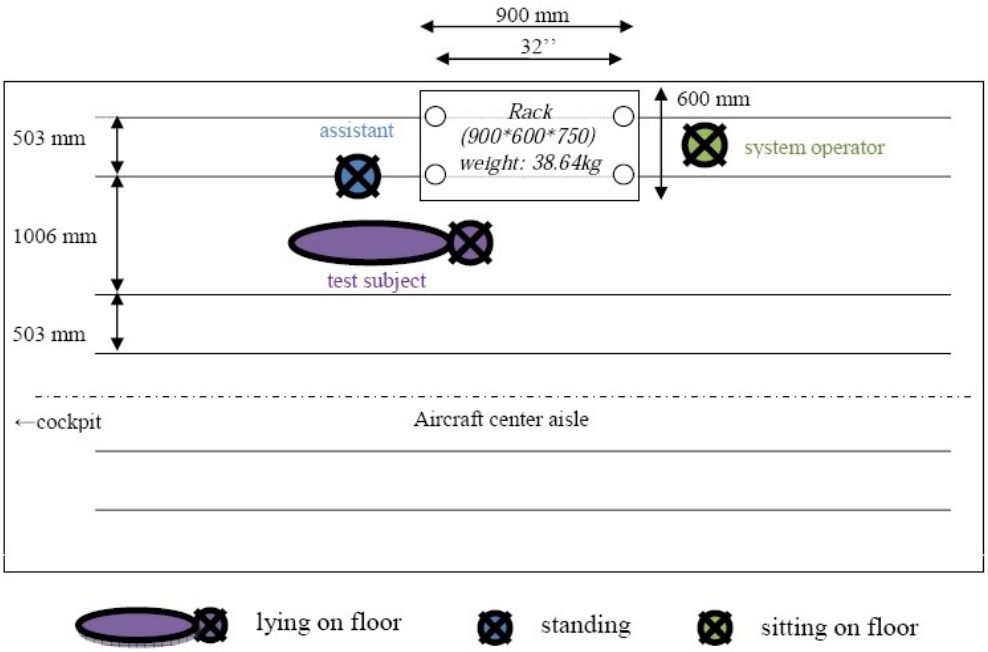
Parabola sequence of good quality (related to the microgravity phase):



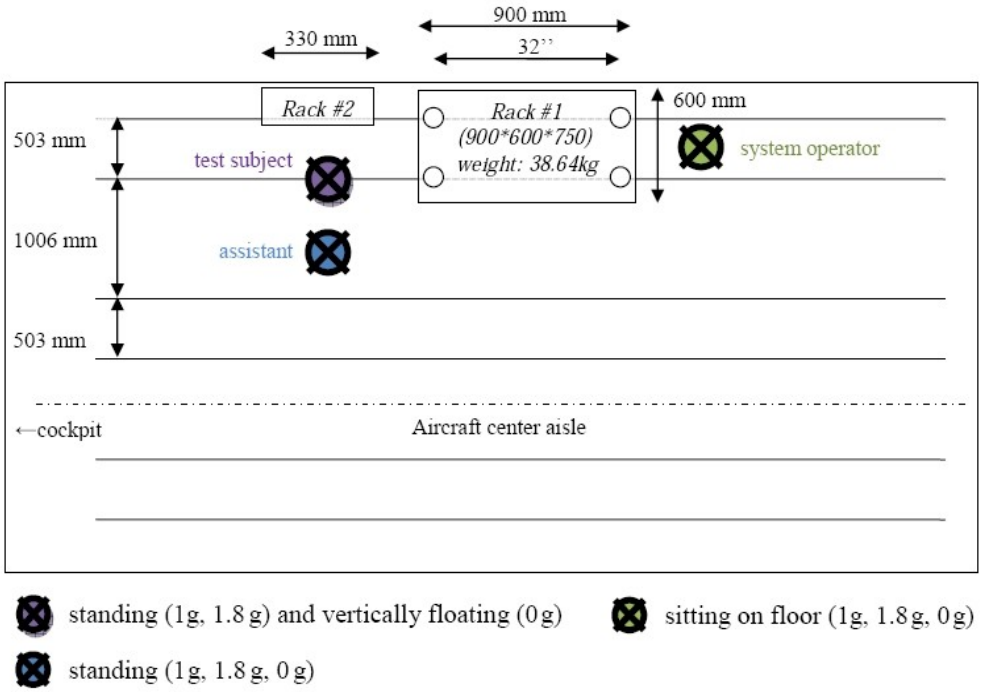
Parabola sequence of poor quality (related to the microgravity phase):

Schematics for installation the flight racks and experimental setups in the Zero-G Airbus A300

Schematic of installation of the experiment in the Zero-G Aircraft A300 during the 56th PFC:



Schematic of installation of the experiment in the Zero-G Aircraft A300 during the 58th PFC:



Appendix F

3DPick Study: Descriptive Statistics of the First Experiment (56th ESA PFC)

This appendix is related to Chapter 8.4 and provides descriptive statistics of the outcomes showing measures of the central tendency and the variability of the dependent variables. It also provides the box plots of skewed distributions and corresponding transformations by natural logarithm.

Descriptive statistics per placement condition (HA, PA, BA) and *g*-level (1*g* PRE, 1.8*g*, 0*g*, 1*g*, 1*g* POST)

Measures of *initial reaction time* [ms]:

Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
HA	1*g* PRE	25	2371.92	2125.00	1021.06	204.21	1000.00	5016.00
	1.8*g*	30	2806.67	2133.00	2279.57	416.19	875.00	12750.00
	0*g*	30	3681.20	2296.50	3402.94	621.29	359.00	13859.00
	1*g*	30	2823.30	1890.50	2339.84	427.19	1046.00	12046.00
	1*g* POST	30	2570.87	2031.50	1281.13	233.90	1187.00	5843.00
	total	145	2867.30	2110.00	2277.18	189.11	359.00	13859.00
PA	1*g* PRE	25	1970.64	1922.00	758.15	151.63	890.00	4781.00
	1.8*g*	30	1774.43	1687.50	863.19	157.60	1015.00	5640.00
	0*g*	30	2337.53	1734.50	1303.20	237.93	860.00	5578.00
	1*g*	30	1904.10	1695.50	800.23	146.10	625.00	4141.00
	1*g* POST	30	1714.71	1500.00	635.44	114.13	953.00	4141.00
	total	145	1937.70	1710.50	919.28	76.08	625.00	5640.00
BA	1*g* PRE	25	1501.24	1532.00	420.90	84.18	860.00	2672.00
	1.8*g*	30	1758.37	1453.00	985.59	179.94	688.00	4484.00
	0*g*	30	1468.20	1203.00	805.09	146.99	578.00	3516.00
	1*g*	30	1596.00	1539.00	624.79	114.07	593.00	3078.00
	1*g* POST	30	1671.84	1531.00	758.61	136.25	750.00	4640.00
	total	145	1602.98	1507.50	749.63	62.04	578.00	4640.00
total		435	2134.32	1735.00	1570.59	75.13	359.00	13859.00

Measures of *frequency of correct target hits*:

Placement	*g*-level	N	Sum	Mean	Median	Stddev	Stderr	Min	Max
HA	1*g* PRE	25	260	10.40	10.00	2.02	0.40	6.00	14.00
	1.8*g*	30	305	10.17	11.00	3.80	0.69	0.00	15.00
	0*g*	30	204	6.80	7.50	3.60	0.66	0.00	13.00
	1*g*	30	338	11.27	12.00	3.38	0.62	3.00	16.00
	1*g* POST	30	327	10.90	12.00	4.05	0.74	1.00	16.00
	total	145	1434	9.89	10.00	3.80	0.32	0.00	16.00
PA	1*g* PRE	25	344	13.76	14.00	2.33	0.47	9.00	17.00
	1.8*g*	30	410	13.67	14.50	3.33	0.61	1.00	18.00
	0*g*	30	344	11.47	12.00	3.14	0.57	5.00	17.00
	1*g*	30	442	14.73	15.00	2.84	0.52	6.00	19.00
	1*g* POST	30	466	15.03	15.00	2.65	0.48	10.00	19.00
	total	145	2006	13.74	14.00	3.13	0.26	1.00	19.00
BA	1*g* PRE	25	353	14.12	15.00	2.32	0.46	9.00	18.00
	1.8*g*	30	405	13.50	15.00	3.97	0.73	5.00	19.00
	0*g*	30	384	12.80	13.50	3.65	0.67	2.00	18.00
	1*g*	30	422	14.07	14.00	2.83	0.52	5.00	18.00
	1*g* POST	30	442	14.26	15.00	3.14	0.56	5.00	18.00
	total	145	2006	13.74	14.50	3.26	0.27	2.00	19.00
total		435	5446	12.46	13.00	3.85	0.18	0.00	19.00

Measures of *percentage error rate* [%]:

Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
HA	1*g* PRE	25	18.77	18.18	13.33	2.67	0.00	57.14
	1.8*g*	30	23.08	19.38	23.58	4.30	0.00	100.00
	0*g*	30	39.53	33.24	26.49	4.84	0.00	100.00
	1*g*	30	18.51	17.65	14.74	2.69	0.00	70.00
	1*g* POST	30	23.95	19.38	23.18	4.23	0.00	88.89
	total	145	24.97	20.00	22.26	1.85	0.00	100.00
PA	1*g* PRE	25	15.99	11.76	12.75	2.55	0.00	52.63
	1.8*g*	30	16.80	15.39	16.13	2.94	0.00	66.67
	0*g*	30	21.07	16.23	18.75	3.42	0.00	58.82
	1*g*	30	13.58	10.56	12.51	2.28	0.00	53.85
	1*g* POST	30	11.92	6.67	12.49	2.24	0.00	44.44
	total	145	15.84	11.76	14.93	1.24	0.00	66.67
BA	1*g* PRE	25	11.13	11.11	8.88	1.78	0.00	30.77
	1.8*g*	30	15.66	10.82	17.98	3.28	0.00	57.14
	0*g*	30	24.54	25.66	16.73	3.05	0.00	55.00
	1*g*	30	13.14	11.81	10.66	1.95	0.00	31.25
	1*g* POST	35	14.02	11.11	13.80	2.48	0.00	54.55
	total	145	15.84	11.76	14.75	1.22	0.00	57.14
total		435	18.87	15.00	18.13	0.87	0.00	100.00

Measures of *response time* [ms]:

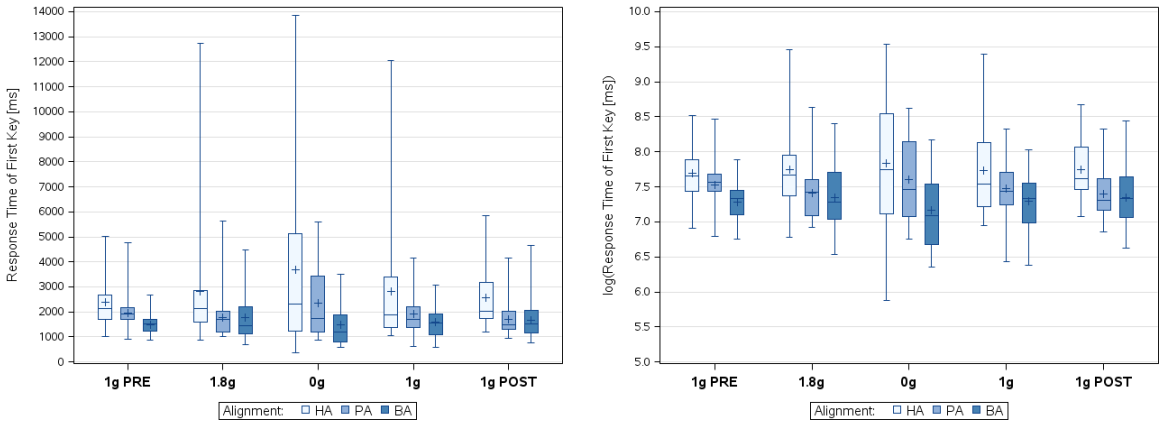
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
HA	1*g* PRE	236	1471.63	1484.00	412.26	26.84	31.00	2531.00
	1.8*g*	273	1359.59	1297.00	401.66	24.31	94.00	3453.00
	0*g*	165	1542.41	1422.00	700.20	54.51	31.00	6344.00
	1*g*	305	1344.11	1265.00	474.48	27.17	78.00	4172.00
	1*g* POST	293	1326.91	1266.00	401.32	23.45	125.00	2938.00
	total	1272	1392.86	1328.00	475.32	13.33	31.00	6344.00
PA	1*g* PRE	310	1201.56	1187.50	333.20	18.92	16.00	2125.00
	1.8*g*	373	1168.44	1125.00	504.63	26.13	78.00	9016.00
	0*g*	298	1324.24	1172.00	677.83	39.27	31.00	6016.00
	1*g*	407	1151.56	1094.00	371.27	18.40	125.00	4782.00
	1*g* POST	428	1176.74	1141.00	329.42	15.92	47.00	2640.00
	total	1816	1197.83	1140.00	453.90	10.65	16.00	9016.00
BA	1*g* PRE	334	1238.55	1188.00	446.29	24.42	31.00	4657.00
	1.8*g*	384	1155.11	1094.00	376.79	19.23	16.00	2781.00
	0*g*	306	1204.40	1094.00	603.26	34.49	15.00	8422.00
	1*g*	391	1218.80	1109.00	579.05	29.28	109.00	8578.00
	1*g* POST	413	1182.06	1141.00	339.62	16.71	47.00	2937.00
	total	1828	1198.32	1140.00	474.60	11.10	15.00	8578.00
total		4916	1248.47	1187.00	474.88	6.77	15.00	9016.00

Measures of *accuracy by Euclidean distance* [mm]:

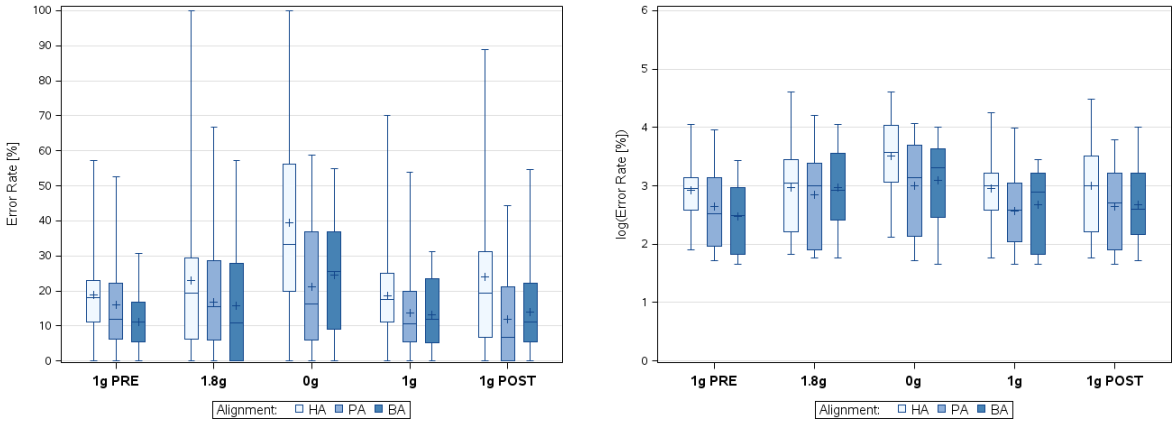
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
HA	1*g* PRE	236	5.39	5.49	1.51	0.10	0.66	7.78
	1.8*g*	273	4.96	5.15	1.67	0.11	0.63	7.79
	0*g*	165	5.22	5.54	1.63	0.13	1.10	7.71
	1*g*	305	5.06	5.14	1.57	0.09	0.27	7.79
	1*g* POST	293	5.21	5.30	1.47	0.10	0.63	7.75
	total	1272	5.15	5.30	1.57	0.05	0.27	7.79
PA	1*g* PRE	310	5.64	5.94	1.55	0.16	2.02	7.79
	1.8*g*	373	5.37	5.56	1.62	0.14	0.72	7.78
	0*g*	298	4.94	4.96	1.81	0.17	0.79	7.79
	1*g*	407	5.36	5.63	1.53	0.14	1.38	7.78
	1*g* POST	428	5.97	6.25	1.28	0.13	2.64	7.79
	total	1816	5.43	5.66	1.60	0.07	0.72	7.79
BA	1*g* PRE	334	5.43	5.69	1.46	0.10	1.24	7.75
	1.8*g*	384	5.07	5.20	1.71	0.11	0.56	7.77
	0*g*	306	4.98	5.16	1.72	0.13	1.17	7.77
	1*g*	391	5.21	5.37	1.65	0.11	0.94	7.76
	1*g* POST	413	5.78	6.13	1.32	0.10	1.88	7.75
	total	1828	5.29	5.51	1.60	0.05	0.56	7.77
total		4916	5.26	5.43	1.59	0.03	0.27	7.79

Skewed distributions and corresponding transformations by natural logarithm per placement condition (HA, PA, BA) grouped by *g*-level (1*g* PRE, 1.8*g*, 0*g*, 1*g*, 1*g* POST)

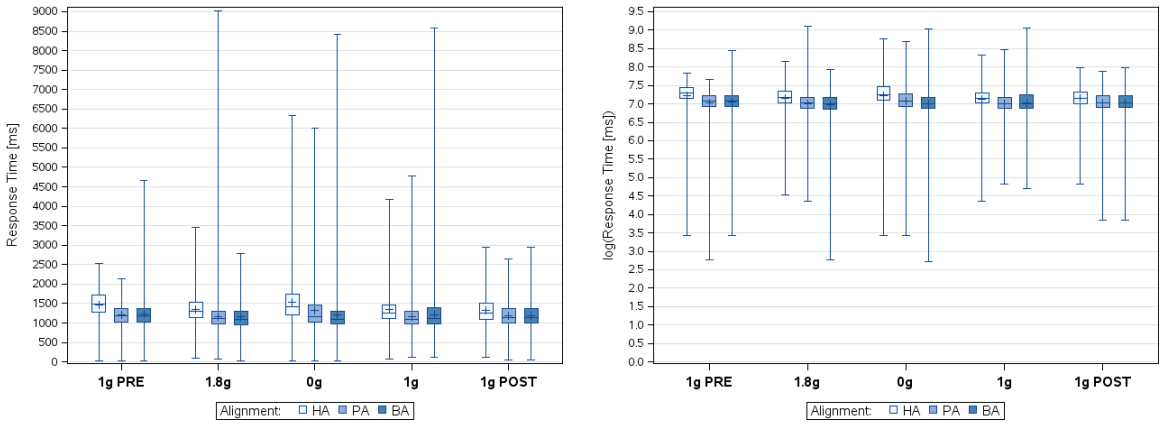
Distribution and log-transformation of *initial reaction time*:



Distribution and log-transformation of *percentage error rate*:



Distribution and log-transformation of *response time*:



Appendix G

3DPick Study: Descriptive Statistics of the Second Experiment (58th ESA PFC)

This appendix is related to Chapter 8.5 and provides descriptive statistics of the outcomes showing measures of the central tendency and the variability of the dependent variables. It also provides the box plots of skewed distributions and corresponding transformations by natural logarithm.

Descriptive statistics per placement condition (PA, BA) and *g*-level (1*g* PRE, 1.8*g*, 0*g*, 1*g*)

Measures of *initial reaction time* [ms]:

Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1*g* PRE	42	1593.81	1507.50	497.31	76.74	781.00	3503.00
	1.8*g*	42	1965.79	1665.50	920.81	142.08	1026.00	5103.00
	0*g*	42	4021.36	3855.50	1924.47	296.95	723.00	7970.00
	1*g*	42	2390.60	2062.00	1207.18	186.27	1084.00	7058.00
	total	168	2492.89	1884.00	1548.96	119.50	723.00	7970.00
BA	1*g* PRE	42	1690.19	1370.50	878.93	135.62	874.00	5038.00
	1.8*g*	42	1601.21	1262.50	935.90	144.41	885.00	5753.00
	0*g*	42	2018.31	1870.00	937.92	144.72	603.00	5470.00
	1*g*	42	1560.62	1319.50	1066.40	164.55	716.00	6803.00
	total	168	1717.58	1395.00	965.61	74.50	603.00	6803.00
total		336	2105.24	1646.00	1345.95	73.43	603.00	7970.00

Measures of *frequency of correct target hits*:

Placement	*g*-level	N	Sum	Mean	Median	Stddev	Stderr	Min	Max
PA	1*g* PRE	42	604	14.38	15.00	2.06	0.32	10.00	18.00
	1.8*g*	42	448	10.67	11.00	3.26	0.50	3.00	17.00
	0*g*	42	381	9.07	9.50	3.19	0.49	0.00	14.00
	1*g*	42	564	13.43	14.00	3.12	0.48	0.00	19.00
	total	168	1997	11.89	12.50	3.62	0.28	0.00	19.00
BA	1*g* PRE	42	592	14.10	14.00	2.58	0.40	5.00	20.00
	1.8*g*	42	564	13.43	14.00	2.92	0.45	4.00	18.00
	0*g*	42	573	13.64	14.00	2.88	0.44	8.00	19.00
	1*g*	42	664	15.81	16.00	2.39	0.37	7.00	19.00
	total	168	2393	14.24	15.00	2.84	0.22	4.00	20.00
total		336	4390	13.07	13.00	3.45	0.19	0.00	20.00

Measures of *percentage error rate* [%]:

Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1*g* PRE	42	6.06	5.72	6.68	1.03	0.00	21.43
	1.8*g*	42	21.53	20.71	14.05	2.17	0.00	60.00
	0*g*	42	32.52	30.77	17.13	2.67	0.00	69.23
	1*g*	42	9.01	6.67	9.16	1.43	0.00	47.37
	total	166	17.24	13.33	16.22	1.26	0.00	69.23
BA	1*g* PRE	42	6.15	5.56	7.42	1.14	0.00	27.78
	1.8*g*	42	7.98	6.25	8.35	1.29	0.00	29.41
	0*g*	42	15.41	12.50	10.80	1.67	0.00	40.00
	1*g*	42	4.90	5.13	5.71	0.88	0.00	21.05
	total	168	8.61	6.25	9.17	0.71	0.00	40.00
total		336	12.90	8.33	13.82	0.76	0.00	69.23

Measures of *response time* [ms]:

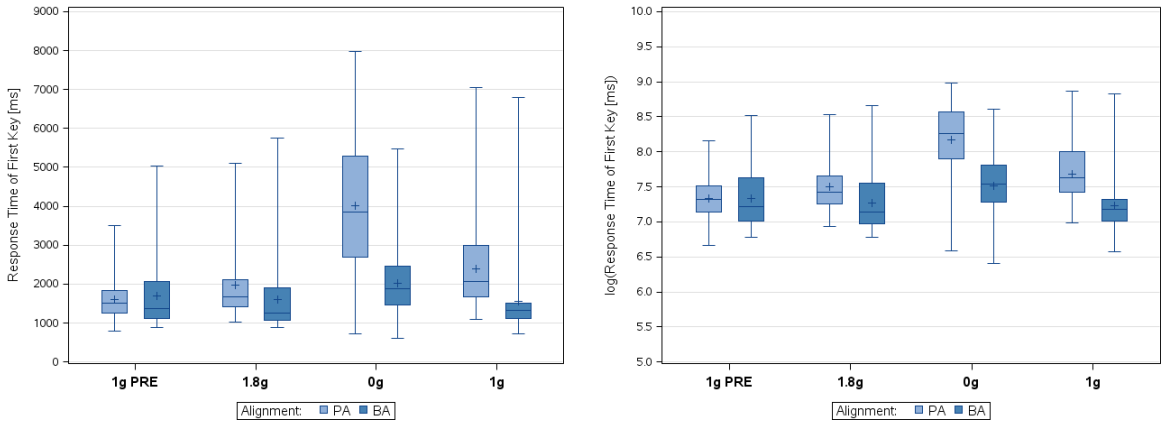
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1*g* PRE	561	1309.23	1267.00	245.31	10.36	497.00	2462.00
	1.8*g*	357	1277.55	1228.00	364.60	19.30	632.00	5577.00
	0*g*	265	1317.32	1250.00	392.07	24.08	549.00	4661.00
	1*g*	511	1254.98	1228.00	303.75	13.44	547.00	4141.00
	total	1694	1287.45	1245.00	316.64	7.69	497.00	5577.00
BA	1*g* PRE	554	1311.88	1313.00	292.28	12.42	844.00	3613.00
	1.8*g*	516	1207.58	1163.00	339.39	14.94	633.00	5163.00
	0*g*	487	1230.23	1184.00	445.70	20.20	657.00	7014.00
	1*g*	627	1184.88	1164.00	235.91	9.42	711.00	2680.00
	total	2184	1232.57	1197.00	333.76	7.14	633.00	7014.00
total		3878	1256.54	1218.50	327.48	5.26	497.00	7014.00

Measures of *accuracy by Euclidean distance* [mm]:

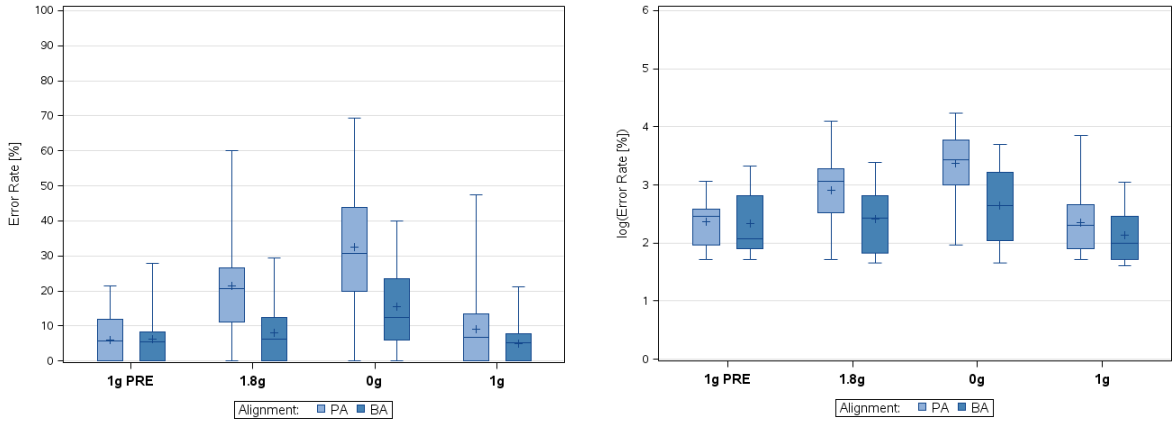
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1*g* PRE	561	5.07	5.59	2.45	0.10	0.23	7.60
	1.8*g*	357	5.38	5.58	2.10	0.11	0.37	7.60
	0*g*	265	5.75	6.30	2.03	0.12	0.25	7.60
	1*g*	511	5.52	6.05	2.11	0.09	0.23	7.60
	total	1694	5.37	5.84	2.23	0.05	0.23	7.60
BA	1*g* PRE	554	4.88	4.74	2.18	0.09	0.20	7.60
	1.8*g*	516	4.46	4.41	2.03	0.09	0.15	7.60
	0*g*	487	4.85	4.97	2.16	0.10	0.52	7.60
	1*g*	627	4.32	4.29	2.22	0.09	0.05	7.60
	total	2184	4.61	4.49	2.16	0.05	0.05	7.60
total		3878	4.95	5.08	2.22	0.04	0.05	7.60

Skewed distributions and corresponding transformations by natural logarithm per placement condition (PA, BA) grouped by *g*-level (1*g* PRE, 1.8*g*, 0*g*, 1*g*)

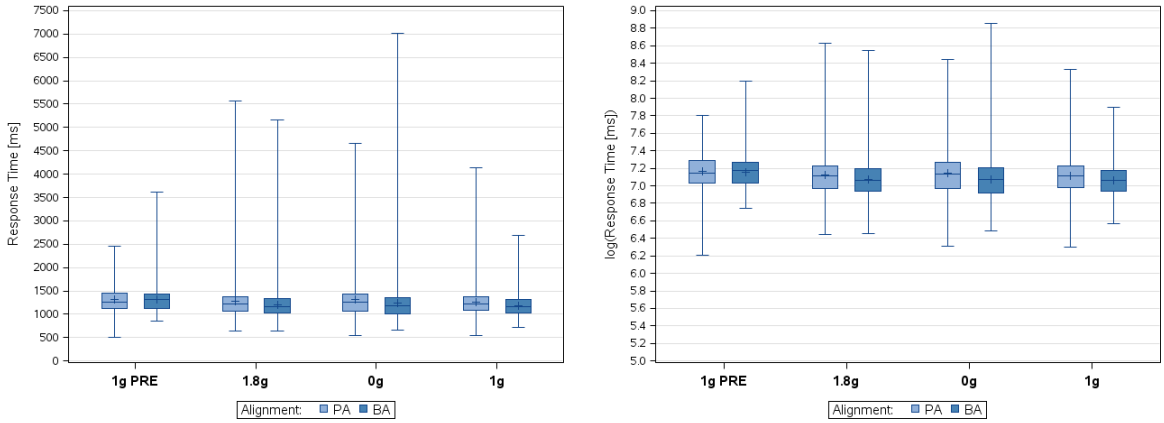
Distribution and log-transformation of *initial reaction time*:



Distribution and log-transformation of *percentage error rate*:



Distribution and log-transformation of *response time*:



Appendix H

3DPick Study: Descriptive Statistics of the Motion Analysis (58th ESA PFC)

This appendix is related to Chapter 8.6 and provides descriptive statistics of the outcomes showing measures of the central tendency and the variability of the dependent variables. It also provides the *allocations of the number of tracking paths to the time classes and time series plots showing the distribution of the pointing velocity per interface condition, both related to the two-component model.*

Mislocalization of target: Measures of overshot and undershot targets per interface placement (PA, BA) and *g*-level (0*g*, 1.8*g*, 1*g*)

Frequency and percentage of *overshot* and *undershot* targets:

Placement	Type	1*g*		0*g*		1.8*g*	
		Sum	*Percentage*	*Sum*	*Percentage*	*Sum*	*Percentage*
PA	overshoot	471	83.51	317	83.20	357	79.69
	undershoot	93	16.49	64	16.80	91	20.31
	total	564	100.00	381	100.00	448	100.00
BA	overshoot	531	79.97	409	71.38	459	81.38
	undershoot	133	20.03	164	28.62	105	18.62
	total	664	100.00	573	100.00	564	100.00

Measures of distance between the target's centre and final intersection as difference d_z [mm]:

Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1g	564	4.68	4.86	4.38	0.18	-6.49	14.54
	0g	381	4.88	5.24	4.74	0.24	-7.48	14.13
	1.8g	448	3.99	4.25	4.44	0.21	-6.52	14.53
	total	1393	4.51	4.77	4.51	0.12	-7.48	14.54
BA	1g	664	3.01	2.58	3.62	0.14	-7.27	13.55
	0g	573	2.72	2.51	4.41	0.18	-8.37	14.24
	1.8g	448	3.99	4.25	4.44	0.21	-6.52	14.53
	total	1801	2.91	2.73	3.80	0.09	-8.37	14.24
total		3194	3.61	3.44	4.20	0.07	-8.37	14.54

Two-Component Model: Allocation of number of tracking paths to time classes per interface placement (PA, BA) and *g*-level (0*g*, 1.8*g*, 1*g*)

Time class	PA			BA		
	1*g*	0*g*	1.8*g*	1*g*	0*g*	1.8*g*
0.1	1418	761	1024	1719	1330	1414
0.2	1490	821	1069	1794	1365	1479
0.3	1533	820	1067	1846	1427	1517
0.4	1533	815	1096	1885	1421	1547
0.5	1494	848	1064	1816	1343	1452
0.6	1592	816	1161	1906	1412	1519
0.7	1546	827	1093	1805	1342	1446
0.8	1544	796	1099	1842	1384	1505
0.9	1493	742	1058	1762	1304	1437
1.0	1419	664	1025	1652	1194	1350
1.1	1250	515	915	1417	1035	1152
1.2	927	401	663	984	782	763
1.3	658	299	476	602	547	529
1.4	409	194	301	354	325	320
1.5	228	160	182	176	190	176
1.6	143	125	134	87	128	85
1.7	81	84	87	52	76	54
1.8	51	59	57	40	59	39
total	18809	9747	13571	21739	16664	17784

Two-Component Model: Measures of pointing displacement and velocity per interface placement (PA, BA) and *g*-level (0*g*, 1.8*g*, 1*g*)

Measures of *displacement* [mm]:

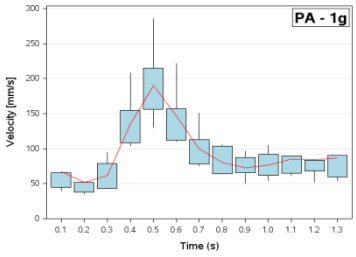
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1g	91	2.76	2.42	1.34	0.14	0.94	7.70
	0g	91	2.83	2.50	1.00	0.11	1.34	6.80
	1.8g	91	2.78	2.44	1.14	0.12	1.28	7.06
	total	273	2.79	2.44	1.17	0.07	0.94	7.70
BA	1g	84	2.39	2.11	1.20	0.13	0.57	6.18
	0g	84	1.90	1.69	0.81	0.09	0.71	4.52
	1.8g	84	2.39	2.01	1.21	0.13	0.69	6.29
	total	252	2.23	1.88	1.11	0.07	0.57	6.29

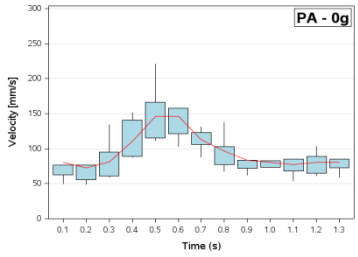
Measures of *pointing velocity* [mm/s]:

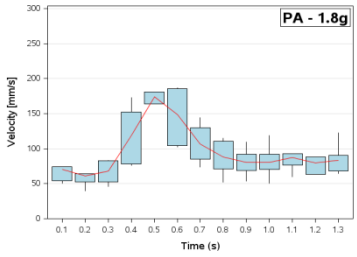
Placement	*g*-level	N	Mean	Median	Stddev	Stderr	Min	Max
PA	1g	91	95.23	79.37	48.60	5.09	35.14	285.85
	0g	91	96.20	85.04	34.51	3.62	48.71	220.85
	1.8g	91	96.04	82.77	40.23	4.22	39.35	255.53
	total	273	95.82	82.77	41.37	2.50	35.14	285.85
BA	1g	84	90.58	80.56	45.59	4.97	23.84	239.47
	0g	84	70.79	62.03	30.02	3.28	26.77	167.51
	1.8g	84	92.03	76.77	47.03	5.13	28.15	242.11
	total	252	84.47	71.33	42.56	2.68	23.84	242.11

Two-Component Model: Distribution of pointing velocity over the time classes per interface placement (PA, BA) and *g*-level (0*g*, 1.8*g*, 1*g*)

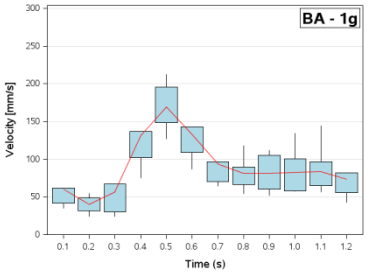
Distribution of pointing velocity over the time classes under the PA conditions:

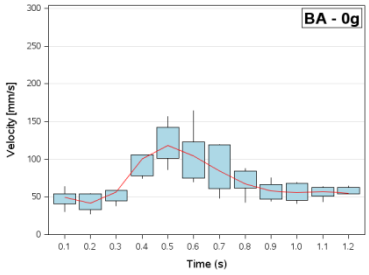


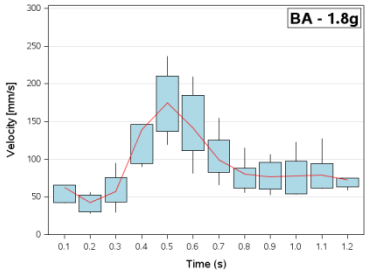




Distribution of pointing velocity over the time classes under the BA conditions:







Appendix I

Weight and Haptic Study: Descriptive Statistics

This appendix is related to the outcomes of the Weight-Study and the Haptic-Study presented in Chapter 9. It provides their descriptive statistics, showing measures of the central tendency and the variability of the dependent variables. It additionally provides the participants' inter-individual characteristics and resulted arm weightings of the Haptic-Study.

Weight Study: Descriptive statistics per sizing technique (SU, SC, SE) and Gz load (1.5*g*, 2*g*, 2.3*g*)

Measures of *frequency of correct target hits and percentage error rate*:

Sizing Technique	Correct Hits	Error [%]
	Mean ± SD	Mean ± SD
	Gz = 1.5*g*	
SU	20.70 ± 2.63	0.53 ± 2.44
SC	21.20 ± 2.76	0.66 ± 1.81
SE	21.07 ± 2.30	0.66 ± 1.71
	Gz = 2.0*g*	
SU	17.56 ± 4.22	2.69 ± 5.51
SC	18.31 ± 4.37	1.85 ± 5.41
SE	19.85 ± 2.48	1.00 ± 2.58
	Gz = 2.3*g*	
SU	18.45 ± 3.57	2.40 ± 4.66
SC	16.88 ± 3.84	2.89 ± 5.16
SE	18.95 ± 2.39	2.23 ± 4.69

Measures of *response time and pointing speed*:

Sizing Technique	Response time [ms]	Speed [mm/ms]
	Mean ± SD	Mean ± SD
	Gz = 1.5*g*	
SU	1007.35 ± 173.11	0.085 ± 0.017
SC	1016.00 ± 173.11	0.065 ± 0.012
SE	1029.43 ± 164.93	0.101 ± 0.018
	Gz = 2.0*g*	
SU	1012.15 ± 195.92	0.085 ± 0.019
SC	1039.77 ± 202.71	0.058 ± 0.012
SE	1031.43 ± 152.39	0.106 ± 0.017
	Gz = 2.3*g*	
SU	1073.56 ± 184.70	0.079 ± 0.015
SC	1068.76 ± 210.42	0.053 ± 0.012
SE	1078.08 ± 159.12	0.105 ± 0.016

Measures of *Euclidean distance and percentage accuracy*:

Sizing Technique	Euclidean Distance [mm]	Accuracy [%]
	Mean ± SD	Mean ± SD
	Gz = 1.5*g*	
SU	4.23 ± 2.49	60.10 ± 23.46
SC	3.58 ± 1.96	56.59 ± 23.71
SE	5.38 ± 2.99	58.51 ± 23.13
	Gz = 2.0*g*	
SU	4.78 ± 2.74	54.93 ± 25.79
SC	3.80 ± 1.87	49.13 ± 24.99
SE	5.85 ± 3.34	57.47 ± 24.30
	Gz = 2.3*g*	
SU	5.02 ± 2.50	52.72 ± 23.58
SC	3.36 ± 1.71	51.94 ± 24.45
SE	5.65 ± 3.14	60.24 ± 22.09

Haptic Study: Resulted arm weights per haptic-feedback method (nHF, HF) and Gz load (1*g*, 1.3*g*, 1.5*g*, 1.8*g*)

Participants' inter-individual characteristics and the resulted weight attached to the arm per Gz load:

Participant	m_{body} [kg]	m_{arm} [kg]	m_{add} [kg]		
			1.3*g*	**1.5*g***	**1.8*g***
S1	88	4.73	1.42	2.37	3.78
S2	66	3.55	1.07	1.78	2.84
S3	76	4.08	1.22	2.04	3.26
S4	78	4.20	1.26	2.09	3.36
S5	74	3.98	1.19	1.99	3.18
S6	80	4.3	1.29	2.15	3.44
S7	88	4.73	1.42	2.37	3.78
S8	60	3.22	0.97	1.61	2.58
S9	80	4.30	1.29	2.15	3.44
S10	69	3.71	1.11	1.86	2.97
S11	77	4.14	1.24	2.07	3.31
S12	67	3.60	1.08	1.8	2.88
S13	69	3.71	1.11	1.86	2.97
S14	81	4.37	1.31	2.19	3.49
S15	61	3.28	0.98	1.64	2.62
S16	95	5.11	1.53	2.56	4.09
S17	85	4.57	1.37	2.29	3.66
S18	62	3.34	1.00	1.67	2.67
S19	86	4.63	1.39	2.32	3.70
S20	76	4.09	1.23	2.05	3.27
S21	57	3.07	0.92	1.54	2.46
S22	98	5.27	1.58	2.64	4.22
S23	70	3.07	0.92	1.54	2.46
S24	75	4.09	1.23	2.05	3.27
S25	64	3.07	0.93	1.54	2.46
S26	80	4.09	1.23	2.05	3.27
S27	75	3.07	0.92	1.54	2.46
S28	78	4.09	1.23	2.05	3.27
S29	81	5.27	1.58	2.64	4.22

Haptic Study: Descriptive statistics per interface condition (nHC, HC) and Gz load

Measures of *response time* [ms]:

Interface	Gz	N	Mean	Median	Stddev	Stderr	Min	Max
nHC	1*g*	736	1530.22	1431.00	456.47	16.83	927.00	2964.00
	1.3*g*	714	1454.27	1313.00	432.09	16.17	873.00	2979.00
	1.5*g*	718	1516.25	1404.50	451.38	16.85	892.00	2981.00
	1.8*g*	703	1517.42	1409.00	445.26	16.79	677.00	2906.00
	total	2871	1504.70	1391.00	447.22	8.35	677.00	2981.00
HC	1*g*	781	1230.60	1118.00	338.45	12.11	713.00	2894.00
	1.3*g*	776	1217.40	1119.00	316.04	11.35	710.00	2970.00
	1.5*g*	777	1232.86	1146.00	327.00	11.73	717.00	2878.00
	1.8*g*	776	1259.09	1156.00	349.45	12.54	762.00	2997.00
	total	3110	1234.98	1135.00	333.16	5.97	710.00	2997.00
total		5981	1364.45	1251.00	414.56	5.36	677.00	2997.00

Measures of *response time* [ms]:

Interface	Angle	Gz	N	Mean	Median	Stddev	Stderr	Min	Max
nHC	vertical (0°)	1*g*	135	1566.24	1491.00	463.24	39.87	961.00	2891.00
		1.3*g*	131	1398.14	1269.00	385.30	33.66	886.00	2979.00
		1.5*g*	137	1497.84	1401.00	438.17	37.44	923.00	2855.00
		1.8*g*	140	1526.51	1431.50	439.01	37.10	845.00	2778.00
		total	543	1498.18	1395.00	435.94	18.71	845.00	2979.00
	oblique (45°)	1*g*	411	1534.04	1440.00	453.75	22.38	927.00	2918.00
		1.3*g*	394	1464.37	1302.50	455.09	22.93	873.00	2975.00
		1.5*g*	401	1539.11	1411.00	477.51	23.85	892.00	2981.00
		1.8*g*	379	1515.80	1415.00	452.13	23.22	677.00	2902.00
		total	1585	1513.64	1403.00	460.34	11.56	677.00	2981.00
	horizontal (90°)	1*g*	190	1496.36	1386.00	457.60	33.20	939.00	2964.00
		1.3*g*	189	1472.13	1369.00	412.01	29.97	939.00	2571.00
		1.5*g*	180	1479.32	1389.50	397.36	29.62	929.00	2893.00
		1.8*g*	184	1513.83	1377.50	437.94	32.29	922.00	2906.00
		total	743	1490.39	1377.00	426.62	15.65	922.00	2964.00
	total		2871	1504.70	1391.00	447.22	8.35	677.00	2981.00
HC	vertical (0°)	1*g*	144	1214.56	1111.00	286.93	23.91	854.00	2268.00
		1.3*g*	145	1183.37	1108.00	244.14	20.27	828.00	2106.00
		1.5*g*	144	1201.49	1131.50	289.48	24.12	764.00	2698.00
		1.8*g*	143	1251.43	1154.00	344.32	28.79	764.00	2997.00
		total	576	1212.59	1122.50	293.50	12.23	764.00	2997.00
	oblique (45°)	1*g*	434	1240.68	1123.50	348.74	16.74	784.00	2894.00
		1.3*g*	431	1218.45	1105.00	337.78	16.27	710.00	2970.00
		1.5*g*	431	1226.15	1142.00	320.17	15.42	717.00	2878.00
		1.8*g*	433	1259.26	1151.00	351.03	16.87	778.00	2857.00
		total	1729	1236.17	1135.00	339.74	8.17	710.00	2970.00
	horizontal (90°)	1*g*	203	1220.45	1099.00	350.46	24.60	713.00	2818.00
		1.3*g*	200	1239.79	1143.00	313.02	22.13	900.00	2808.00
		1.5*g*	202	1269.53	1178.50	362.98	25.54	780.00	2761.00
		1.8*g*	200	1264.21	1159.00	351.30	24.84	762.00	2851.00
		total	805	1248.44	1143.00	344.94	12.16	713.00	2851.00
	total		3110	1234.98	1135.00	333.16	5.97	710.00	2997.00
total			5981	1364.45	1251.00	414.56	5.36	677.00	2997.00

Bibliography

Agan, M., Voisinet, L.A. & Devereaux, A. (1998). NASA's Wireless Augmented Reality Prototype (WARP). *Proc. of AIP Space Technology and Applications International Forum*, 420(1), pp. 236-242.

Aguzzi, M., Bosca, R. & Müllerschkowski, U. (2008). Astronaut training in view of the future: A Columbus payload instructor perspective. *Acta Astronautica*, vol. 66, issues 3-4, pp. 401-407.

Aguzzi, M. & Lamborelle, O. (2012). A Proposal of Visual Guidelines for On Board Procedures. *63*rd *International Astronautical Federation Congress (IAC)*, paper-code: IAC-12-B6.1.6.

André-Deshays, C., Israel, I., Charade, O., Berthoz, A., Popov, K. and Lipshits, M. (1993). Gaze control in microgravity: 1. Saccades, pursuit, eye-head coordination. *Journal of Vestibular Research*, vol. 3, pp. 331–343.

Arguello, L. (2009). WEAR WEarable Augmented Reality. URL: https://www.esa.int/gsp/ACT/doc /EVENTS/bmiworkshop/ACT-PRE-BNG-WEAR(BMI_Workshop).pdf, last visit: 28.09.2016.

Atwell, W. (1990). Astronaut Exposure to Space Radiation: Space Shuttle Experience. *SAE Technical Paper 901342*, doi:10.4271/901342.

Ayres, P. & Sweller, J. (2005). The split-attention-principle in multimedia learning. In: R.E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning. New York, NY: Cambridge University Press, pp. 135-146.

Azuma, R.T. (1997). A survey of augmented reality. *Presence, Teleoperators and Virtual Environments*, 6(4), pp. 355-385.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6), pp. 34-47.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In: G. H. Bower (Ed.): The psychology of learning and motivation: Advances in research and theory. New York: Academic Press. vol. 8, pp. 47–89.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), pp. 417-423.

Bagshaw, M. & Cucinotta, F.A. (2008). Cosmic Radiation. In: J.R. Davis, R. Johnson, J. Stepanek (Eds.), Fundamentals of Aerospace Medcine. 4th Edition, Lippincott Williams and Wilkins, pp. 221-235.

Bai, Z. & Blackwell, A.F. (2012). Analytic review of usability evaluation in ISMAR. *Interacting with Computers*, 24(6), pp. 450-460.

Baird, K.M. & Barfield, W. (1999). Evaluating the Effectiveness of Augmented Reality Displays for a Manual Assembly Task. *Proc. of IEEE Virtual Reality (VR'99)*, vol. 4, pp. 250-259.

Bajura, M., Fuchs, H. & Ohbuchi, R. (1992). Merging Virtual Reality with the Real World: Seeing Ultrasound Imagery Within the Patient. *ACM SIGGRAPH Computer Graphics*, Volume 26, Issue 2, pp. 203-210.

Barfield, W., Rosenberg,C. & Furness, T.A. (1995). Situation awareness as a function of frame of reference, computer-graphics eyepoint elevation, and geometric field of view. *Intern. Journal of Aviation Psychology*, 5(3), pp. 233-56.

Bauer, U., Kiernan, P., Wolff, M. & Martignano, M. (2003). Authoring and Viewing of Electronic ISS Crew Procedures. *ESA STM-268, ESTEC Noordwijk*.

Bay, H., Tuytelaars, T. & Van Gool, L. (2008). SURF: Speeded Up Robust Features. *Journal of Computer Vision and Image Understanding*, 110(3), pp. 346–359.

Behrends, J., Bischofberger, J. & Deutzmann, R. (2009). Duale Reihe Physiologie. *Thieme*, 1st ed.

Berger, M., Gerstenbrand, F., De Col, C., Grill, L., Muigg, A., Kozlovskaja, I., Burlatchkova, N., Sokolov, A., Babaev, B. & Borisov, M. (1993). [Movement disorders in weightlessness]. *Wien Med Wochenschr. 1993*; 143(23-24), pp. 614-619.

Billinghurst, M., Bowskill, J., Dyer, N. & Morphett, J. (1998). Spatial information displays on a wearable computer. *IEEE Computer Graphics and Applications*, 18(6), pp. 24-31.

Bimber, O. & Raskar, R. (2005). Spatial Augmented Reality: Merging Real and Virtual Worlds. A K Peters/CRC Press.

Biocca, F., Tang, A., Owen C. & Xiao, F. (2006). Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. *Proc. of SIGCHI Conf. Human Factors in Comp. Systems (CHI'06)*, pp. 1115-1122, 2006.

Bischoff, R. & Kazi, A. (2004). Perspectives on Augmented Reality based human-robot interaction with industrial robots. *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '04)*, vol. 4, pp. 3226-3231.

Bock, O., Howard, I.P., Money, K.E. & Arnold, K.E. (1992). Accuracy of aimed arm movements in changed gravity. *Aviation Space Environmental Medicine*, 63(11), pp. 994–998.

Bock, O., Fowler, B. & Comfort, D. (2001). Human sensorimotor coordination during spaceflight: an analysis of pointing and tracking responses during the "Neurolab" Space Shuttle mission. *Aviation Space Environmental Medicine*, 72(10), pp. 877-883.

Bock, O., Abeele, S., & Eversheim, U. (2003). Sensorimotor performance and computational demand during short-term exposure to microgravity. *Aviat Space Environ Med. 2003 Dec;* 74 (12), pp. 1256-1262.

Bonin, G. (2005). Physiological Issues in Human Spaceflight: Review and Proposed Countermeasures. *Biomedical Engineering and Biomechanics,* MAAE 4906/MECH 5801.

Boud, A.C., Haniff, D.J., Baber, C. & Steiner, S.J. (1999). Virtual reality and augmented reality as a training tool for assembly tasks. *Proc. of IEEE International Conference on Information Visualization*, pp. 32-36.

Bowman, D., Johnson, D., & Hodges, L. (1999). Testbed evaluation of virtual environment interaction techniques. *Proc. of ACM Symposium on Virtual Reality Software and Technology (VRST '99)*, pp 26-33.

Bowman, D.A. & Wingrave, C.A. (2001). Design and evaluation of menu systems for immersive virtual environments. *Proc. of IEEE Virtual Reality (VR '01)*, pp. 149-156.

Bowman, D., Kruijff, E., LaViola, J., & Poupyrev, I. (2004). 3D User Interfaces: Theory and Practice. *Addison-Wesley*.

Boyd, A., Fortunato, A., Wolff, M. & Oliveira, D.M. (2016). mobiPV: A new, wearable real-time collaboration software for Astronauts using mobile computing solutions. *SpaceOps 2016 Conference*, Daejeon, Korea.

Brinckmann, E. (1999). Spaceflight opportunities on the ISS for plant research--the ESA perspective. *Adv. Space Res.*, 24(6), pp. 779-88.

Bringoux, L., Blouin, J., Coyle, T., Ruget, H. & Mouchnino, L. (2012). Effect of gravity-like torque on goal-directed arm movements in microgravity. *J Neurophysiol*, 107(9), pp. 2541-2548.

Buxton, W., Hill, R. & Rowley. P. (1985). Issues and Techniques in Touch-Sensitive Tablet Input. *Proc. of ACM SIGGRAPH Computer Graphics*, 19(3), pp. 215-224.

Buxton, W. & Myers, B.A. (1986): A Study In Two-Handed Input. *Proc. of SIGCHI Conference on Human Factors in Computing Systems (CHI '86)*, pp 321-326.

Byers, J.C., Bittner, A.C. & Hill, S.G. (1989) Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? *Advances in Industrial Ergonomics and Safety*, pp. 481-485

Cain, B. (2007) A Review of the Mental Workload Literature. Report RTO-TR-HFM-121-Part-II. Defense Research and Development Canada, Toronto.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E. & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1), pp. 341-377.

Carr, C.E., Schwartz, S.J. & Rosenberg, I. (2002). A Wearable Computer for Support of Astronaut Extravehicular Activity. *Proc. of International Symposium on Wearable Computers (ISWC '02)*, pp. 23-30.

Carriot, J., Bringoux, L., Charles, C., Mars, F., Nougier, V. & Cian, C. (2004). Perceived body orientation in microgravity: effects of prior experience and pressure under the feet. *Aviat Space Environ Med.*, 75(9), pp. 795-9.

Caudell, T.P. & Mizell, D.W. (1992). Augmented Reality: An application of heads-up display technology to manual manufacturing processes. *Proc. of the Twenty-Fifth Hawaii International Conference on System Sciences,* vol.2, pp.659–669.

Chandler, P. & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, vol. 62, pp. 233-246.

Cherry, E.C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *Acoustical Society of America*, 25 (5), pp. 975-79.

Chintamani, K., Van Lierde, B., Maloney, S, Kiernan, P., Oliveira, D.M. & Wolff, M. (2013). Wearable Crew Support Technology on the International Space Station: The Mobile Procedure Viewer (mobiPV). *Proc. of the Human Factors and Ergonomics Society Europe Chapter 2013 Annual Conference*, pp. 279-289.

Chintamani, K., Van Lierde, B., Ilzkovitz, M., Kiernan, P., Chadwick, K., Vachovski, M., Wolff, M. & Oliveira, D.M. (2015). Mobile Procedure Viewer Short Duration Mission, mobiPV-SDM. *ESA TEC-ED & TEC-SW Final Presentation Days.*

Chittaro, L. & Burigat, S. (2004). 3d location-pointing as a navigation aid in virtual environments. *Proc. of the working conference on Advanced Visual Interfaces (AVI '04)*, pp. 267–274.

Cho, C., Yang, H., Kim, G.J & Han, S.H. (2002). Body-based Interfaces. *Proc. of the IEEE International Conference on Multimodal Interfaces (ICMI '02)*, pp. 466-472.

Clarke, A.H., Grigull, J., Müller, R., & Scherer, H. (2000). The three-dimensional vestibulo-ocular reflex during prolonged microgravity. *Exp Brain Res*, vol. 134, pp. 322–334.

Clauser, C.E., McConville, J.T., Young, J.W. (1969). Weight, volume, and center of mass of segments of the human body. *AMRL Technical Report 69-70. Wright-Patterson Air Force Base.*

Clément, G, Vieville, T., Lestienne, F. & Berthoz, A. (1986). Modifications of gain asymmetry and beating field of vertical optokinetic nystagmus in microgravity. *Neurosci. Lett*, vol. 63, pp. 271–274.

Clément, G. & Reschke, M. (1996). Neurosensory and Sensory-motor Functions. In: D. Moore, P. Bie, H. Oser (Eds.), Biological and Medical Research in Space. Springer, Berlin, NY, pp 178-258.

Clément, G. (1998). Alteration of eye movements and motion perception in microgravity. *Brain Research Reviews*, vol. 28, pp.161–172.

Clément, G. (2007). Using your Head Cognition and Sensorimotor Functions in Microgravity. *Gravitational and Space Biology*, 20(2), pp. 65-78.

Clément, G. (2011). Fundamentals of Space Medicine. *Space Technology Library*, *Springer*, 2nd ed.

Cohen, M. (1973). Elevator illusion: Influences of otolith organ activity and neck propriovception. *Perception & Psychophysics*, 14(3), pp. 401-406.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ: Erlbaum.

Cooper, G.E. & Harper, R.P. (1969). The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. Technical Report TN D-5153, NASA.

Coquillart, S. and Wesche, G. (1999). The virtual palette and the virtual remote control panel: A device and an interaction paradigm for the Responsive Workbench™. *Proc. of IEEE Virtual Reality (VR '99)*, pp. 213–216.

Cowan, N. (1995). Attention and memory: An integrated framework. *Oxford Psychology Series, No. 26. Oxford University Press*.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsiderationof mental storage capacity. *Behavioral Brain Science*, 24(1), pp. 87-114.

Craig, A.B., Sherman, W.R. & Will, J.D. (2009). Developing Virtual Reality applications: Foundations of Effective Design. *Morgan Kaufmann Publ Inc*.

CSA. (2014). Working in Space. URL: http://www.asc-csa.gc.ca/eng/astronauts/living-working.asp, last visit: 28.09.2016.

Curtis, D., Mizell, D., Gruenbaum, P., & Janin, A. (1999). Several Devils in the Details: Making an AR Application Work in the Airplane Factory. *Proc. of International Workshop on Augmented Reality (IWAR '98)*, pp. 47-60.

Datcu, D., Cidota, M., Lukosch, S, Oliveira, D.M. & Wolff, M. (2014). Virtual Co-location to Support Remote Assistance for Inflight Maintenance in Ground Training for Space Missions. *Proc of. International Conference on Computer Systems and Technologies (CompSysTech '14)*, pp. 134-141.

Datcu, D., Cidota, M., Lukosch, S, Oliveira, D.M. Agostinh, S. & Wolff, M. (2014a). Augmented Reality Enabled Collaboration. URL: https://indico.esa.int/indico/event/71/contribution/3/material/1/0.pdf, last visit: 28.09.2016.

De Crescenzio, F., Fantini, M., Persiani, F., Di Stefano, L., Azzari, P. & Salti, S. (2011). Augmented Reality for Aircraft Maintenance Training and Operations Support. *IEEE Computer Graphics and Applications*, 31(1), pp. 96-101.

Debevec, P. (1998). Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proc. of the 25th annual conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*, pp. 189-198.

Department of Defense. 1999. Design criteria standard, human engineering. *Technical Report MIL-STD-1472F*.

Dettmann, J., Reitz, G. & Gianfiglio, G. (2007). MATROSHKA—The first ESA external payload on the International Space Station. *Acta Astronautica*, 60(1), pp. 17-23.

Dezfuli, N., Khalilbeigi, M., Huber, J., Müller, F. & Mühlhäuser, M. (2012) Leveraging the palm surface as an eyes-free tv remote control. *Proc. of Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*, pp. 2483-2488.

Dix, A., Finlay, J., Abowd, G.D & Beale, R. (2003). Human Computer Interaction (Third Edition). *Prentice Hall, Pearson Education Limited, Harlow*.

Dow, S., MacIntyre, B., Lee, J., Oezbek C., Bolter, J.D. & Gandy, M. (2005). Wizard of Oz support throughout an iterative design process. *Journal of IEEE Pervasive Computing*, 4(4), pp. 18-26.

Drascic, D. & Milgram, P. (1996). Perceptual Issues in Augmented Reality, Proc. of SPIE, vol. 2653: Stereoscopic Displays and Virtual Reality Systems III, pp. 123-134.

Dünser, A., Grasset, R., Seichter, H., & Billinghurst, M. (2007). Applying HCI Principles to AR Systems Design. HIT Lab NZ, University of Canterbury, New Zealand. URL: http://ir.canterbury.ac.nz/bitstream/10092/2340/1/12604890_2007-MRUI-Applying_HCI_principles.pdf.

Durrant-Whyte, H. & Bailey, T. (2006). Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. *Robotics and Automation Magazine*, 13(2), pp. 99-110.

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66(3), pp. 183-201.

Edwards, D.G. & Berry, J.J. (1987). The efficiency of simulation-based multiple comparisons. *Biometrics*. 43(4), pp. 913-928.

Elliott, D., Helsen W.F. & Chua, R. (2001). A Century Later: Woodworth's (1899) Two-Component Model of Goal-Direct Aiming. *Psychol. Bulletin*, 127(3), pp. 342-357.

Elliott, D. & Khan, M. (2010). Vision and Goal-Directed Movement - Neurobehavioral Perspectives. *Human Kinetics*.

Ellis, S.R. (2000). Collision in space. *Ergonomics in Design*, vol. 8, pp.4–9.

ESA. (2007). Personal Digital Assistants in Space. URL: http://www.esa.int/Our_Activities /Human_Spaceflight/Astrolab/Personal_digital_assistants_in_space, last visit: 28.09.2016.

ESA. (2009). European Astronaut Selection. URL: http://www.esa.int/Our_Activities/Human_ Spaceflight/European_Astronaut_Selection, last visit: 28.09.2016.

ESA. (2015a). European User Guide to Low Gravity Platforms. URL: http://www.esa.int/Our_ Activities/Human_Spaceflight/Research/European_user_guide_to_low_gravity_platforms, last visit: 28.09.2016.

ESA. (2015b). Columbus Payload Accommodation. URL: http://www.capcomespace.net/dossiers /ISS/europe/columbus/Nouveau dossier/colaccom.pdf, last visit: 28.09.2016.

ESA. (2015c). Biolab. URL: http://wsn.spaceflight.esa.int/docs/Factsheets/8 Biolab LR.pdf, last visit: 28.09.2016

ESA. (2015d). User Support and Operation Centres (USOCs). URL: http://www.esa.int/Our_ Activities/Human_Spaceflight/Columbus/User_Support_and_Operations_Centres_USOCs, last visit: 28.09.2016.

ESA. (2015e). Hands-Free in Space and under Water. URL: http://www.esa.int/Our_Activities /Human_Spaceflight/Astronauts/Hands-free_in_space_and_under_water, last visit: 28.09.2016.

Espindola, D.; Silva, B.; Weis, A.; Botelho, S.; Pereira, C.E. (2013). Mobile Advanced Visualization Applied to Oil and Gas Industry Systems. *Proc. of Symposium on Computing and Automation for Offshore Shipbuilding (NAVCOMP'13)*, pp. 93-98.

Fails, J.A. & Olsen D.Jr., (2002). Light Widgets: Interacting in Every-Day Spaces. *Proc. of International Conference on Intelligent User Interfaces (IUI '02)*, pp. 63-69.

Feiner, S., MacIntyre, B. & Seligmann D. (1993a). Knowledge-based augmented reality. *Communications of the ACM*, 36(7), pp. 52-62

Feiner, S., MacIntyre, B., Haupt, M. & Solomon E. (1993b). Windows on the world: 2D windows for 3D augmented reality. *Proc. of ACM Symp. on User Interface Software and Technology (UIST '93)*, pp. 145-155.

Feiner, S., MacIntyre, B. Höllerer, T., & Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Proc. of ISWC '97 (First IEEE Int. Symp. on Wearable Computers)*, pp. 74–81.

Fischler, M.A. & Bolles, R.C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), pp. 381-395.

Fisk, J., Lackner, J.R. & DiZio, P. (1993). Gravitoinertial force level influences arm movement control. *Journal of Neurophysiology*., 69(2), pp. 504–511.

Fite-Georgel, P. (2011). Is there a reality in industrial augmented reality? *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11)*, pp. 201-210.

Fitts, P.M. (1954). The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. In *Journal of Experimental Psychology*, 47 pp. 381-391.

Fitts, R.H., Riley, D.R. & Widrick, J.J. (2001). Functional and structural adaptations of skeletal muscle to microgravity. *J Exp Biol.*, 204(PT 18), pp. 3201-3208.

Fitzmaurice, G.W., Ishii, H. & Buxton, W (1995). Bricks: Laying the Foundations for Graspable User Interfaces. *Proc. of the ACM SIGCHI Conference on Human Factors in Computing (CHI'95)*, pp. 442-449.

Fogt, N., Uhlig, R., Thach, D.P. & Liu, A. (2002). The influence of head movement on the accuracy of a rapid pointing task. *Optometry. 2002 Nov*; 73(11), pp. 665-673.

Fowler, B., Meehan, S. & Singhal, A. (2008). Perceptual-motor performance and associated kinematics in space. *Hum Factors. 2008 Dec*; 50(6):879-92.

Friedrich, W. (2002). ARVIKA-Augmented reality for development, production and service. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '02)*, pp. 3–4, 2002.

Fritz, N., Meyer, T., Blum, T., Lemke, H.U., Ilzkovitz, M., Nevatia, Y., Nolden, M., Wegner, I., Weinlich, M., Breitkreutz, R., Wein, W., Lazerges, M., Angerer, O. & Navab, N. (2009). CAMDASS: An Augmented Reality Medical Guidance System for Spaceflights. *Proc. of Computer Assisted Radiology and Surgery (CARS '09)*.

Gabbard, J.L., Hix, D. & Swan, J.E. (1999). User Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*, 19(6), pp. 51-59.

Gabbard, J.L. & Hix, D. (2001). Researching Usability Design and Evaluation Guidelines for Augmented Reality (AR) Systems. Laboratory for Scientific Visual Analysis, Virginia Tech, USA. URL: http://www.sv.vt.edu/classes/ESM4714/Student_Proj/class00/gabbard/index.html.

Gabbard, J.L. & Swan, J.E. (2008). Usability Engineering for Augmented Reality : Employing User-based Studies to Inform Design. *IEEE Transactions on Visualization and Computer Graphics*, 14(3), pp. 513-525.

Gales, M. & Young, S. (2007). The Application of Hidden Markov Models in Speech Recognition. *Foundations and Trends in Signal Processing*, 1(3), pp. 195–304.

Garcia, J.Q. (2011). A study on PDAs for onboard applications and technologies and methodologies. Personal and Ubiquitous Computing. 15(5), pp. 457-478.

Gay-Bellile, V., Bourgeois, S., Tamaazousti, M. & Collette, S.N. (2012). A mobile markerless Augmented Reality system for the automotive field. *Workshop on Tracking Methods and Applications, IEEE International Symposium on Mixed and Augmented Reality (ISMAR '12)*.

Gawron, V.J. (2008). Human Performance, Workload, and Situational Awareness Measures Handbook. CRC Press, Taylor & Francis Group, LLC.

Gervautz, M & Schmalstieg, D. (2012). Anywhere Interfaces Using Handheld Augmented Reality. *IEEE Computer*, 45(7), pp. 26-31.

Gimeno, J., Morillo, P., Orduña, J.M. & Fernández, M. (2012). An Occlusion-Aware AR Authoring Tool for Assembly and Repair Tasks. *Proc. of GRAPP/IVAPP '12,* pp. 377-386.

Glasauer, S. & Mittelstaedt, H. (1997). Perception of spatial orientation in different g-levels. *Journal of Gravitational Physiology*, vol. 4(2), pp. 5-8.

Glasauer, S. & Mittelstaedt, H. (1998). Perception of spatial orientation in microgravity. *Brain Res Rev*, 28(1-2), pp. 185-193.

Gleason, T.K. (2008) The Vestibular system. In: P.M. Cone (Ed.), Neuroscience in Medicine. Human Press, pp. 591-600.

Goeke, M., König, P. & Gramann K. (2013). Different strategies for spatial updating in yaw and pitch path integration. *Frontiers in Behavioral Neuroscience*, 7(5), pp. 1-13.

Goldstein, E.B. (2010). Cognitive Psychology: Connecting Mind, Research and Everyday Experience. *Wadsworth Publishing*.

Gopher, D. and Donchin, E. (1986). Workload – An examination of the concept. In: K.R. Boff, L. Kaufman, J.P. Thomas (Eds.), Handbook of Perception and Human Performance. Volume 2. Cognitive Processes and Performance. John Wiley and Sons Inc., pp. 41-1 - 41-49.

Grabherr, L., Karmali, F., Bach, S., Indermaur, K., Metzler, S. & Mast, F.W. (2007). Mental own-body and body-part transformations in microgravity. *Journal of Vestibular Research*, 17 (5–6), pp. 279–287.

Gramann K. (2013) Embodiment of Spatial Reference Frames and Individual Differences in Reference Frame Proclivity. *Spatial Cognition and Computation*, 13(1), pp. 1-25.

Graybiel, A. & Kellogg R.S. (1967) Inversion illusion in parabolic flight: its probable dependence on otolith function. Aerospace Med, vol. 38, pp. 1099-1103.

Grenon, S.M., Xiao, X., Hurwitz, S., Ramsdell, C.D., Sheynberg, N., Kim, C., Williams, G.H. & Cohen, R.J. (2005). Simulated microgravity induces microvolt T wave alternans. *Ann Noninvasive Electrocardiol.*, 10(3), pp. 363-370.

Grimm P., Haller M., Paelke V., Reinhold S., Reimann C., Zauner J. (2002). AMIRE - Authoring Mixed Reality. *Proc. of the First IEEE International Augmented Reality Toolkit Workshop*.

Grossman, T. & Balakrishnan, R. (2004) Pointing at trivariate targets in 3D environments. *Proc. of ACM Conference on Human Factor in Computing Systems (CHI '04)*, pp. 447-454.

Grubert, J., Hamacher, D., Mecke, R., Böckelmann, I., Schega, L., Huckauf, A., Urbina, M., Schenk, M., Doil, F. & Tumler, J. (2010). Extended investigations of user-related issues in mobile industrial AR. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '10)*, pp. 229-230.

Guiard, Y. (1987). Symmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. *Journal of Motor Behavior*, 19(4), pp. 486-517.

Guardiera, S., Schneider, S., Noppe, A., Strüder, H.K. (2008). Motor performance and motor learning in sustained +3Gz acceleration. In *Aviation Space and Environmental Medicine*, vol. 79(9), pp. 852-9.

Gurfinkel, V.S., Lestienne, F., Levik, YuS, Popov, K.E. & Lefort, L. (1993). Egocentric references and human spatial orientation in microgravity. II. Body-centred coordinates in the task of drawing ellipses with prescribed orientation. *Exp Brain Res*, 95(2), pp. 343-348.

Gustafson, S.G., Rabe, B. & Baudisch, P.M. (2013). Understanding palm-based imaginary interfaces: the role of visual and tactile cues when browsing. *Proc. of SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, pp. 889-898.

Güven, S. & Feiner, S. (2003). Authoring 3D hypermedia for wearable augmented and virtual reality. *Proc. of the 7th International Symposium on Wearable Computers (ISWC '03)*, pp. 118-126.

Hackos, J. & Redish, J. (1998). User and Task Analysis for Interface Design. *Chichester: Wiley*.

Hain, T.C. & Helminski, J. (2014). Anatomy and Physiology of the Normal Vestibular System. In: S.J. Herdmann, R. Clendaniel (Eds.), Vestibular Rehabilitation, 4th Edition. F.A. Davis Company, pp. 2-19.

Hampshire, A., Seichter, H., Grasset, R. & Billinghurst, M. (2006). Augmented reality authoring: Generic context from programmer to designer. *Proc. of the 18th Australia conference on Computer-Human Interaction (OZCHI '06)*, pp. 409-412.

Haringer, M. & Regenbrecht, H. (2002). A pragmatic approach to Augmented Reality Authoring. In *Proc. of ISMAR '02*, p. 237.

Harm, D.L., Parker, D.E., Reschke, M.F. & Skinner, N.C. (1998). Relationship between selected orientation rest frame, circular vection and space motion sickness. *Brain Research Bulletin*, vol. 47, pp. 497 –501.

Harm, D.L., Reschke, M.F. & Wood S.J. (2014). Spatial orientation and motion perception in microgravity. In: R. Hoffman, P. Hancock, M. Scerbo, R. Parasuraman, J. Szalma (Eds.), The Cambridge Handbook of Applied Perception Research. Cambridge: Cambridge University Press, vol. 2, chap. 45, pp. 912–929.

Harrison, C., Tan, D. & Morris, D. (2010). Skinput: appropriating the body as an input surface. *Proc. of Proc. of Conf. Human Factors in Computing Systems (CHI '10)*, pp. 453-462.

Harrison, C., Benko, H. & Wilson, A.D. (2011). OmniTouch: Wearable Multitouch Interaction Everywhere. *Proc. of 24th Annual ACM Symposium on User interface Software and Technology (UIST'11)*, pp. 441-450.

Harrison, C. & Faste, H. (2014) Implications of location and touch for on-body projected interfaces. *Proc. of Conference on Designing Interactive Systems (DIS '14)*, pp. 543-552.

Hart, S.G. & Staveland, L.E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: P.A. Hancock, N. Meshkati (Eds.), Human Mental Workload. Amsterdam: North Holland Press.

Hart, S.G. (2006). NASA-Task Load Index (NASA-TLX); 20 Years Later. *Proc. of the Human Factors and Ergonomics Society 50th Annual Meeting*, pp. 904-908.

Hartmann, K., Ali, K., Strothotte, T. (2004). Floating Labels. Applying Dynamic Potential Fields for Label Layout. In *Smart Graphics*, vol. 3031, Springer, pp. 101–113.

Henderson, S. & Feiner, S. (2007). Augmented Reality for Maintenance and Repair (ARMAR). *Technical Report AFRL-RH-WP-TR-2007-0112,* United States Air Force Research Lab.

Henderson, S. & Feiner, S. (2008). Opportunistic Controls: Leveraging Natural Affordances as Tangible User Interfaces for Augmented Reality. *Proc. of ACM Virtual Reality Software and Technology (VRST '08)*, pp. 211-218.

Henderson, S. & Feiner, S. (2009). Evaluating the Benefits of Augmented Reality for Task Localization in Maintenance of an Armored Personnel Carrier Turret. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '09)*, pp. 135-144.

Henderson, S. & Feiner, S. (2010). Opportunistic Tangible User Interfaces for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 16(1), pp. 4-16.

Henderson, S. & Feiner, S. (2011a). Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 17(10), pp. 1355-1368.

Henderson, S. & Feiner, S. (2011b). Augmented Reality in the Psychomotor Phase of a Procedural Task. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11)*, pp. 191-200.

Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G. & Verplank, W. (1992). ACM SIGCHI Curricula for Human-Computer Interaction. *Report of the ACM SIGCHI Curriculum Development Group, ACM*. URL: http://old.sigchi.org/cdg/, last visit: 28.09.2016.

Hix, D. & Hartson, H. (1993). Developing User Interfaces: Ensuring Usability through Product & Process. *John Wiley and Sons*.

Hockey, G.R.J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45(1-3), pp. 73-93.

Holick, M.F. (2000). Microgravity-induced bone loss—will it limit human space exploration? *The Lancet*, 355(9215), pp. 1569–1570.

Höllerer, T., Hallaway, D., Tinna, N. & Feiner, S. (2001). Steps Toward Accommodating Variable Position Tracking Accuracy in a Mobile Augmented Reality System. *Proc. of International Workshop on Artificial Intelligence in Mobile System (AIMS '01)*, pp. 31-37.

Hopf, J.M., Boehler, C.N., Schoenfeld, M.A., Mangun, G.R. & Heinze, H.J. (2012) Attentional Selection for Locations, Features, ad Objects in Vision. In: G.R.. Mangun (Ed.), Neuroscience of Attention: Attentional Control and Selection. 1st edition, Oxford University Press, pp. 3-29.

Howard, I.P, Jenkin, H.L. & Hu, G. (2000). Visually-induced reorientation illusions as a function of age. *Aviat Space Environ Med*, 71(9 Suppl), pp. A87-A91.

Howard, I.P & Hu, G. (2001). Visually induced reorientation illusions. *Perception*, 30(5), pp. 583-600.

Huang, W., Alem, L. & Livingston M.A. (2013). Human Factors in Augmented Reality Environments. *Springer Science+Business Media*, New York.

Huber, P.J. (1981). Robust Statistics. John Wiley & Sons, New York.

Ishii, H. & Ullmer, B. (1997). Tangible Bits: Towards seamless interfaces between people, bits and atoms. *Proc. of the ACM SIGCHI Human Factors in Computer Systems Conference*, pp. 234-241.

Ishii, H., Bian, Z., Fujino, H., Sekiyama, T., Nakai, T., Okamoto, A., Shimoda, H., Izumi, M., Kanehira, Y. & Morishita, Y. (2007). Augmented reality applications for nuclear power plant maintenance work. *Proc. of International Symposium on Symbiotic Nuclear Power Systems for 21st Century (ISSNP '07)*, pp. 262-268.

ISO/DIS 9241-9. (2000). Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices. International Standard, International Organization for Standardization.

Jennings, R.T., Sawin, C.F. & Barratt, M.R. (2008). Space Operations. In: J.R. Davis, R. Johnson, J. Stepanek (Eds.), Fundamentals of Aerospace Medcine. 4th Edition, Lippincott Williams and Wilkins, pp. 516-551.

Jersild, A.T. (1927). Mental set and shift. Archives of Psychology, no. 89.

Jones, P.M. (2010). Human Performance in Space. *Reviews of Human Factors and Ergonomics*, 6(1), pp. 172-197.

Kahneman, D. (1973). Attention and Effort. *Prentice-Hall Inc., Englewood Cliffs, New Jersey*.

Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (1999). Managing Split-attention and Redundancy in Multimedia Instruction. *Applied Cognitive Psychology*, 13(4), pp. 351-371.

Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist*, 38(1), pp. 23-31.

Kanas, N. & Manzey, D. (2008). Space psychology and psychiatry (2nd revised edition). *Dordrecht: Springer*.

Kato, H. & Billinghurst, M. (1999). Marker Tracking and HMD calibration for a video-based augmented reality conferencing system. *Proc. of the 2nd International Workshop on Augmented Reality (IWAR '99)*, pp. 85-94.

Katuntsev, V.P. (1998). [Decompression sickness-one of the vital problems of aerospace medicine]. *Aviakosm Ekolog Med.*, 32(6), pp. 11-20.

Kazennikov, O.V. & Wiesendanger M. (2005) Goal synchronization of bimanual skills depends on proprioception. *Neurosci Lett.*, 388(3), pp. 153-6.

Kenefick, R.W., Cheuvront, S.N., Castellani, J.W. & O'Brien, C. (2008). Thermal Stress. In: J.R. Davis, R. Johnson, J. Stepanek (Eds.), Fundamentals of Aerospace Medcine. 4th Edition, Lippincott Williams and Wilkins, pp. 206-220.

Khuong, B.M., Kiyokawa, K., Miller, A., La Viola, J.J., Mashita, T. & Takemura, H. (2014). The effectiveness of an AR-based context-aware assembly support system in object assembly. *Proc. of IEEE Virtual Reality (VR'14)*, pp. 57-62.

Kieffaber, P.D. & Hetrick, W.P. (2005). Event-related potential correlates of task switching and switch costs. Psychophysiology, 42(1), pp. 56-71.

Kim, D., Hilliges, O., Izadi, S., Butler, A.D., Chen, J., Oikonomidis, I. & Olivier, P. (2012) Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor. *Proc. of ACM symposium on User Interface Software and Technology (UIST '12)*, pp. 167-176.

Kishishita, N., Kiyokawa, K., Kruijff, E., Orlosky, J., Mashita, T., Takemura, H. (2014). Analysing the Effects of a Wide Field of View Augmented Reality Display on Search Performance in Divided Attention Tasks. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'14)*.

Knöpfle, C., Weidenhausen, J., Chauvigné, L. & Stock, I. (2005). Template Based Authoring for AR Based Service Scenarios. *Proc. of IEEE Virtual Reality (VR '05)*, pp. 237-240.

Kohli, L. & Whitton, M. (2005). The Haptic Hand: Providing User Interface Feedback with the Non-Dominant Hand in Virtual Environments. *Proc. of Graphics Interface (GI '05)*, pp. 1-8.

Kohli, L., Whitton, M.C., Brooks, F.P. (2012). Redirected touching: The effect of warping space on task performance. In *Proc of 3DUI'12*, pp. 105-112.

Kölsch, M. & Turk, M. (2002). Keyboards without keyboards: A survey of virtual keyboards. *Technical Report 2002-21*, UCSB, Santa Barbara, CA.

Kornilova, L.N. (1997). Orientation Illusions in Spaceflight. J Vestib Res, 7(6), pp. 429-439.

Krevelen D.W.F. van & Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. International Journal of Virtual Reality, 9(2), pp. 1-20.

Kruijff, E., Swan, J.E. & Feiner, S. (2010). Perceptual issues in augmented reality revisited. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'10)*, pp. 3–12.

Kurillo, G., Mihelj, M., Munih, M. & Bajd, T. (2007). Multi-Fingered Grasping and Manipulation in Virtual Environments Using an Isometric Finger Device. *Presence Teleoperators and Virtual Environments archive*, 16(3), pp. 293-306.

Lackner, J.R. & Graybiel, A. (1981). Illusions of postural, visual, and aircraft motion elicited by deep knee in the increased gravitoinertial force phase of parabolic flight. Evidence for dynamic sensory-motor calibration to earth gravity force levels. *Exp Brain Res*, 44(3), pp. 312-316.

Lackner, J.R. & Graybiel, A. (1983). Perceived orientation in free-fall depends on visual, postural, and architectural factors. *Aviat Space Environ Med*, 54(1), pp. 47-51.

Lackner, J.R., DiZio P. & Fisk, J.D. (1992). Tonic vibration reflexes and background force level. *Acta Astronautica;* vol. 26, pp. 133–136.

Lackner, J.R. & DiZio P. (1992). Gravitoinertial force level affects the appreciation of limb position during muscle vibration. *Brain Res;* 592(1-2), pp. 175–180.

Lackner, J.R. (1993). Orientation and movement in unusual force environments. *Psychol Sci*, 4(3), pp. 134-142.

Lackner, J.R. & DiZio, P. (1996). Motor function in microgravity: movement in weightlessness. *Current Opinion in Neurobiology*, 6(6), pp. 744-750.

Lackner, J.R., Rabin, E. & DiZio, P. (2000). Fingertip contact suppresses the destabilizing influence of leg muscle vibration. *J Neurophysiol*, 84(5), pp. 2217-24.

Lackner, J.R. & DiZio, P. (2000) Human orientation and movement control in weightless and artificial gravity environments. *Exp Brain Res*, vol. 130, pp. 2–26.

Lackner, J.R. & DiZio, P. (2006). Space motion sickness. *Experimental Brain Research*, vol. 175, pp. 377–399.

Latash, M. (2007). Neurophysiological Basis of Movement. *Human Kinetics*, 2nd ed.

Lecuyer, A., Coquillart, S., Kheddar, A., Richard, P. & Coiffet, P. (2000). Pseudo-Haptic Feedback: Can Isometric Input Devices Simulate Force Feedback? *Proc. of the IEEE Virtual Reality Conference (VR '00)*, pp. 83-90.

Ledermann, F. & Schmalstieg, D. (2005). APRIL: A High-level Framework for Creating Augmented Reality Presentations. *Proc. of Virtual Reality (VR '05)*, pp. 187-194.

Lee, G.A., Nelles, C., Billinghurst, M., & Kim, G.J. (2004). Immersive Authoring of Tangible Augmented Reality Applications. *Proc. of the Third IEEE International Symposium on Mixed and Augmented Reality (ISMAR '04)*, pp. 172-181.

Lee, T. & Höllerer, T. (2007) Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking. *Proc. of 11th IEEE International Symposium on Wearable Computers (ISWC '07)*, pp. 1-8.

Lepetit, V., Vacchetti, L., Thalmann, D. & Fua, P. (2003). Fully Automated and Stable Registration for Augmented Reality Applications. *Proc. of the International Symposium on Mixed and Augmented Reality (ISMAR '03)*, pp. 93-102.

Lepetit, V. & Fua, P. (2005). Monocular model-based 3D tracking of rigid objects. *Journal of Foundations and Trends in Computer Graphics and Vision*, 1(1), pp. 1-89.

Lepetit, V., Moreno-Noguer, F. & Fua, P. (2009). EP*n*P: An Accurate $O(n)$ Solution to the P*n*P Problem. *International Journal of Computer Vision*, 81(2), pp. 155-166.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathmatics*, II(2), pp. 164-168.

Levine, B.D., Pawelczyk, J.A., Zuckerman, J.H., Zhang, R., Fu, Q., Iwasaki, K., Ray, C. & Blomqvist, C.G. (2003). Neural Control of the Cardiovascular System in Space. In: J.C. Buckey, J.L. Homick (Eds.), The Neurolab Spacelab Mission: Neuroscience Research in Space. NASA, pp. 175-185.

Li, N. & Duh, H.B.L. (2013). Cognitive Issues in Mobile Augmented Reality: An Embodied Perspective? In *Huang, T., Alem, L. and Livingston, M., Human Factors in Augmented Reality Environments. Springer Science+Business Media*, New York 2013, pp. 109-135.

Lindeman, R.W., Sibert, J.L. & Hahn, J.K. (1999). Hand-Held Windows: Towards Effective 2d Interaction in Immersive Virtual Environments. *Proc. of IEEE Virtual Reality Conference (VR' 1999)*, pp. 205-212.

Lipshits, M., Bengoetxea, A., Cheron, G. & McIntyre, J. (2005). Two reference frames for visual perception in two gravity conditions. *Perception*, vol. 34, pp. 545-555.

Liu, L., van Liere, R., Nieuwenhuizen, C., and Martens, J.-B. (2009). Comparing aimed movements in the real world and in virtual reality. *Proc. of Virtual Reality Conference (VR '09)*, pp. 219-222.

Liu, L. & Liere, R. (2009). Designing 3D selection techniques using ballistic and corrective movements. *Proc. of the 15th Joint virtual reality Eurographics conference on Virtual Environments (JVRC'09)*, pp. 1-8.

Loomis, J., Golledge, R. & Klatzky, R. (1993). Personal guidance system for the visually impaired using GPS, GIS, and VR technologies. *Proc. of Conf. on Virtual Reality and Persons with Disabilities*.

Looser, J., Grasset, R., Seichter, H. & Billinghurst, M. (2006). OSGART - A Pragmatic Approach to MR. *IEEE International Symposium on Mixed and Augmented Reality (ISMAR '06), Industrial Augmented Reality Workshop*.

Looser, J. (2007). AR Magic Lenses: Addressing the Challenge of Focus and Context in Augmented Reality. Thesis Doctoral, University of Canterbury, Christchurch, New Zealand.

Looser, J., Billinghurst, M., Grasset, R. & Cockburn, A. (2007). An evaluation of virtual lenses for object selection in augmented reality. *Proc of. GRAPHITE'07*, pp. 203-210.

Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*. 60(2), pp. 91–110.

Lucas, B.D. & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *Proc. of the 7th International Joint Conference on Artificial Intelligence*, pp. 674–679.

Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Linton, P.M., Wierwille, W.W., Zaklad, A.L., Bittner, A.C. & Wherry, R.J. (1989). Operator Workload: Comprehensive Review and Evaluation of Operator Workload Methodologies. *Technical Report 851, Army Research Institute for the Behavioral and Social Sciences, Alexandria*.

MacIntyre, B., Gandy, M., Bolter, J.D., Dow, S. & Hannigan, B. (2003). DART: The Designer's Augmented Reality Toolkit. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'03)*, pp. 329-330.

MacKenzie, I.S. (1992). Fitts' law as a research and design tool in human-computer interaction. In *Journal of Human-Computer Interaction*, vol. 7, pp. 91–139.

Manzoni, D. (2009). Vestibulo-Spinal Reflexes. In: M.D. Binder, N. Hirokawa, U. Windhorst (Eds.), Encyclopedia of Neuroscience, Springer Berlin Heidelberg, pp. 4245-4250.

Markov-Vetter, D., Muehl, J., D. Schmalstieg & Sorantin, E. (2009). 3D augmented reality simulator for neonatal cranial sonography. *International Journal of Computer Assisted Radiology and Surgery (CARS'09)*, Springer , 4(1), pp. 19-20.

Markov-Vetter, D., Mittag, U. & Staadt, O. (2009). An Augmented Reality supported Rack Guidance System in Space Flight. *Proc. of Joint Virtual Reality Conference of EGVE – EuroVR (JVRC'09)*, pp. 33-34.

Markov-Vetter, D., Moll, E. & Staadt, O. (2012). Evaluation of 3D Selection Tasks in Parabolic Flight Conditions: Pointing Task in Augmented Reality User Interfaces. *Proc. of* the 11th International Conference on Virtual Reality Continuum and Its Applications in Industry (*VRCAI'12)*, Singapore, pp.287-293.

Markov-Vetter, D., Moll, E. & Staadt, O. (2013a). Verifying Sensorimotoric Coordination of Augmented Reality Selection under Hyper- and Microgravity. *International Journal for Advanced Computer Science (IJACS)*, 3(5), pp. 217-226.

Markov-Vetter, D., Moll, E. & Staadt, O. (2013b). The Effect of Hyper- and Microgravity on Visuomotor Coordination of Augmented Reality Selection in Correlation with Spatial Orientation and Haptical Feedback. *Proc of 64th International Astronautical Congress (IAC 2013)*, 1(14), pp. 35-47.

Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2013c). The Impact of Altered Gravitation on Performance and Workload of Augmented Reality Hand-Eye-Coordination: Inside vs. Outside of Human Body Frame of Reference. *Proc. of Joint Virtual Reality Conference of EGVE – EuroVR* (*JVRC'13)*, Paris, France, pp. 65-72.

Markov-Vetter, D. & Staadt, O. (2013d). A Pilot Study for Augmented Reality Supported Procedure Guidance to Operate Payload Racks On-Board the International Space Station. *Proc. of the International Symposium of Mixed and Augmented Reality (ISMAR'13)*, Adelaide, Australia.

Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2014a). A Proof-Of-Concept Study on the Impact of Artificial Hypergravity on Force-Adapted Target Sizing for Direct Augmented Reality Pointing. *Proc. of IEEE Virtual Reality Conference (VR'14)*, pp. 95-96.

Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2014b). A Proof-Of-Concept Study on the Impact of Artificial Hypergravity on Force-Adapted Target Sizing for Direct Augmented Reality Pointing. *Proc. of IEEE Virtual Reality Conference (VR '14)*, pp. 95-96.

Markov-Vetter, D., Zander, V., Latsch, J. & Staadt, O. (2015). The Influence of Gravity-Adapted Target Resizing on Direct Augmented Reality Pointing Under Simulated Hypergravity. In *Proc. of 10th International Joint Conference on Computer Graphics Theory and Applications (GRAPP'15)*, pp. 401-411.

Markov-Vetter, D., Zander, V., Latsch, J., & Staadt, O. (2016). Enhancement of Direct Augmented Reality Object Selection by Gravity-Adapted Resizing. In Book: J. Braz, J. Pettré, P. Richard, A. Kerren, L. Linsen, S. Battiato, F. Imai (Eds.), Computer Vision, Imaging and Computer Graphics - Theory and Applications. International Joint Conference, VISIGRAPP 2015, Berlin, Germany, March 11-14, 2015. *Springer International Publishing Switzertland*, pp. 75-96.

Marmolejo, J., Gernux, C.G. & Blaser, R.W. (1989). A Helmet Mounted Display Demonstration Unit for a Space Station Application. *Proc. of 19th Intersociety Conference on Environmental Systems*.

Marmolejo, J. (1994). Helmet-Mounted Display and Associated Research Activities Recently Conducted by the NASA/Johnson Space Center. *Proc. of SPIE - The International Society for Optical Engineering*, vol. 2218, pp.281-291.

Marmolejo, J. (1996). An Electronic Cuff Checklist Information Display for Extravehicular Activity. *Proc of. 26th International Conference on Environmental Systems*.

Marner, M., Smith, R., Walsh, J. & Thomas, B. (2014). Spatial User Interfaces for Large Scale Projector-Based Augmented Reality. *IEEE Computer Graphics and Applications Magazine,* Issue 6, pp. 74-82.

Marquardt, D.W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics*, 11(2), pp. 431-441.

Martignano, M. (2006). PDAs in Space - Development of onboard PDA based applications. *ESA*, URL: ftp://ftp.estec.esa.nl/pub/wm/anonymous/wme/Web/PDAs_in_SpaceMauriziov1.ppt, last visit: 28.09.2016.

Maschke, P., Oubaid, V. & Pecena, Y. (2011). How Do Astronaut Candidate Profiles Differ From Airline Pilot Profiles? Results From the 2008/2009 ESA Astronaut Selection. *Aviation Psychology and Applied Human Factors*, 1(1), pp. 38-44.

Mayer, R.E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13(2), pp. 125-139.

Mechtcheriakov, S., Berger, M., Molokanova, E., Holzmueller, G., Wirtenberger, W., Lechner-Steinleitner, S., De Col, C., Kozlovskaya, I. & Gerstenbrand, F. (2002). Slowing of human arm movements during weightlessness: the role of vision. *Eur J Appl Physiol*, 87(6), pp. 576-583.

Medenica, Z., Kun, A.L., Paek, T. & Palinko, O. (2011). Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. *Proc. of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'11)*, pp. 265-274.

Memi, E. (2006). Now see this - Boeing's working on augmented reality, which could change space trainings, ops. *Boeings Frontiers*, October 2006. p. 21.

Meshkati, N. (1988). Heart rate variability and mental workload assessment. Human Mental Workload. *Advanced in Psychology*, vol. 52, pp. 101-115.

Metz, R. (2015). Why NASA Wants Microsoft's HoloLens in Space. MIT Technology Review. URL: https://www.technologyreview.com/s/541126/why-nasa-wants-microsofts-hololens-in-space, last visit: 28.09.2016.

Meyer, D.E., Abrams, R.A., Kornblum, S., Wright, C.E. & Smith, J.E.K. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review*, 95(3), 340-370.

Millberg, J. (2011). Generierung von Keyframes als Basis für markerloses Tracking in 3D Augmented Reality Anwendungen. *Projektbericht*, Hochschule Bonn-Rhein-Sieg.

Millberg, J. (2012). Markerloses, modellbasiertes Echtzeit-Tracking für AR-Applikationen. *Master Thesis*, Hochschule Bonn-Rhein Sieg.

Millberg, J. (2013). Markerloses, modellbasiertes Echtzeit-Tracking für AR-Applikationen. GI Informatiktage, poster session.

Milgram, P. & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, Vol. E77-D, No. 12, December 1994.

Miller, G.A. (1994). The magical number seven, plus or minus two: some limits on our capacity for processing information. 1956. *Psychological Review*, 101(2), pp. 343–52.

Mine, M.R., Brooks Jr., F. P. & Sequin, C. H. (1997). Moving Objects in Space: Exploiting Proprioception Virtual-Environment Interaction. *Proc. of SIGGRAPH '97*, pp. 19–26.

Mittelstaedt, H. (1989). The role of the pitched-up orientation of the otoliths in two recent models of the subjective vertical . *Biological Cybernetics*, 61(6) , pp. 405–416.

Money, K.E. & Cheung, B.S. (1991). Alterations of proprioceptive function in the weightless environment. *J Clin Pharmacol*, 31(10), pp. 1007-1009.

Moray, N. (1967). Where is attention limited? A survey and a model. *Acta Psychologica*, vol. 27, pp. 84-92.

Morphew, E., Balmer, D.V. & Khoury, G.J. (2001). Human performance in space. *Ergonomics in Design*, 9(4), pp. 6–11.

Mühlhäuser, M. & Gurevych, I. (2008). Handbook of Research on Ubiquitous Computing Technology for Real Time Enterprises. Information Science Reference; 1 edition, Hershey, PA.

Mulloni, A., Seichter, H. & Schmalstieg, D. (2011). Handheld augmented reality indoor navigation with activity-based instructions. *Proc. of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'11)*, pp. 211-220.

Mura, K., Gorecky, D. & Meixner, G. (2012). Involving Users in the Design of Augmented Reality-Based Assistance in Industrial Assembly Tasks. In: S. Trzcielinski, W. Karwowski (Eds.), Advances in Ergonomics in Manufacturing. CRC Press Inc, pp. 232-240.

NASA. (2000). International Space Station User's Guide. URL: http://www.spaceref.com/iss/ops /ISS.User.Guide.R2.pdf, last visit: 28.09.2016.

NASA. (2015). Reference Guide to the International Space Station. URL: http://www.nasa.gov/sites /default/files/atoms/files/np-2015-05-022-jsc-iss-guide-2015-update-111015-508c.pdf, last visit: 28.09.2016.

NASA JPL. (2002). Wiring the Fashion Trend of the Future. URL: http://www.jpl.nasa.gov/news /news.php?feature=491, last visit: 28.09.2016.

NASA JSC. (2014). Augmented Reality for Operations and Trainings. URL: http://www.spaceref.com/news/viewsr.html?pid=45164, last visit: 28.09.2016.

NASA JSC. (2015). Vestibular Experiments in Spacelab (178072). URL: https://lsda.jsc.nasa.gov /scripts/experiment/exper.aspx?exp_index=1, last visit: 28.09.2016.

Neisser, U. & Lazar, R. (1964). Searching for Novel Targets. *Perceptual and Motor Skills*, vol. 19, pp. 427-432.

Neumann, U. & Majoros, A. (1998). Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. *Proc. of IEEE Virtual Reality (VR '98)*, pp. 4–11.

Nevatia, Y., Chintamani, K., Meyer, T., Blum, T., Runge, A. & Fritz, N. (2011). Computer aided medical diagnosis and surgery system: towards automated medical diagnosis for long term space missions. *Proc. of 11th Symposium on Advanced Space Technologies in Robotics and Automation (ASTRA '11).*

Nielsen, J. (1993). Usability Engineering. *San Francisco: Morgan Kaufmann*.

Norman, D.A. (2002). The Design of Everyday Things. *Reprint edition, Basic Books*.

Ockerman, J.J. & Pritchett, A.R. (1998). Preliminary Investigation of Wearable Computers for Task Guidance in Aircraft Inspection. *Proc. of Int. Symp. on Wearable Computers (ISWC '98)*, pp. 33-40.

Ohlenburg, J., Herbst, I., Lindt, I., Fröhlich, T. & Broll, W. (2004). The MORGAN framework: enabling dynamic multi-user AR and VR projects. *Proc. of VRST'04*, pp. 166-169.

Oman, C.M. (1988). The role of static visual-orientation cues in the etiology of space motion sickness. *Proc. of Symposium on Vestibular Organs and Altered Environments*, Houston, TX: NASA Johnson Space Center, pp. 25-38.

Oman, C.M., Lichtenberg, B.K. & Money, K.E. (1990). Space motion sickness monitoring experiment: Spacelab 1. In: G.H. Crampton (Ed.), Motion and Space Sickness. Boca Raton: CRC Press, pp. 217-246.

Oman, C.M., Howard, I.P., Smith, T., Beall, A.C., Natapoff, A., Zacher, J.E. & Jenkin, H.L. (2003). The Role of Visual Cues in Microgravity Spatial Orientation. In: J.C. Buckey, J.L. Homick (Eds.), The Neurolab Spacelab Mission: Neuroscience Research in Space. Results from the STS-90, Neurolab Spacelab Mission. NASA/SP-2003-535, pp. 69-81.

Oman, C. (2007). Spatial Orientation and Navigation in Microgravity. In: F. Mast, L. Jancke (Eds.), Spatial processing in navigation, imagery and perception. Springer, pp. 209-247.

Ong, S.K., Yuan, M.L. & Nee, A.Y.C. (2008). Augmented Reality Applications in Manufacturing: A Survey. *International Journal of Production Research*, 46(10), pp. 2707-2742.

Osterlund, J. & Lawrence, B. (2012). Virtual reality: Avatars in human space flight training. *Acta Astronautica*, vol. 71, pp. 139-150.

Palmieri, R.M., Ingersoll, C.D. & Hoffmann, M.A. (2004). The Hoffmann Reflex: Methodologic Considerations and Applications for Use in Sports Medicine and Athletic Training Research. *Journal of Athletic Training*, 39(3), pp. 268–277.

Paloski, W.H., Black, F.O., Reschke, M.F., Calkins, D.S. & Shupert, C. (1993). Vestibular ataxia following shuttle flights: effects of microgravity on otolith-mediated sensorimotor control of posture. *American Journal of Otology*, 14(1), pp. 9-17.

Pang, Y., Nee, A.Y.C., Ong, S.K., Yuan, M. & Youcef-Toumi, K. (2006) Assembly feature design in an augmented reality environment. *Assembly Automation*, 26(1), pp.34-43.

Park, S., Toole, T. & Lee, S. (1999). Functional roles of the proprioceptive system in the control of goal-directed movement. *Percept Mot Skills*, 88(2), pp. 631-647.

Park, Y., Lepetit V. & Woo, W. (2011). Extended keyframe detection with stable tracking for multiple 3d object tracking. *IEEE Trans. Vis. Comput. Graph.*, 17(11), pp. 1728-1735.

Parker, D.E., Reschke, M.F., Arrott, A.P., Homick, J.L. & Lichtenberg, B.K. (1985). Otolith tilt-translation reinterpretation following prolonged weightlessness: implications for pre-flight training. *Aviation, Space, and Environmental Medicine*, vol. 56, pp. 601-606.

Parmet, A.J. & Ercoline, W.R. (2008). Spatial Orientation in Flight. In: J.R. Davis, R. Johnson, J. Stepanek (Eds.), Fundamentals of Aerospace Medcine. 4th ed., Lippincott Williams and Wilkins, pp. 142-205.

Paul, D.B. (1990). Speech Recognition Using Hidden Markov Models. *The Lincoln Laboratory Journal*, 3(1), pp. 41-62.

Pereira Do Carmo, J., Gordo, P.R., Martins, M., Rodrigues, F. & Teodoro, P. (2006). Study of a Direct Visualization Display Tool for Space Applications. *Proc. of 6th International Conference on Space Optics*.

Peters, R.A. (1969). Dynamics of the vestibular system and their relation to motion perception, spatial disorientation, and illusions. *NASA CR-1309*, pp. 1-223.

Pick, S., Hentschel, B., Tedjo-Palczynski, I., Wolter, M., Kuhlen, T. (2010). Automated positioning of annotations in immersive virtual environments. In *Proc. of JVRC'10*, pp. 1-8.

Piekarski, W. & Thomas, B.H. (2003). ThumbsUp: Integrated Command and Pointer Interactions for Mobile Outdoor Augmented Reality Systems. *Proc. of International Conference on HCI*, pp. 185-186.

Platonov, J., Heibel, H., Meier, P. & Grollmann, B. (2006). A Mobile Markerless AR System for Maintenance and Repair. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '06)*, pp. 105-108.

Posner, M.I. (1980). Orienting of attention. *The Quarterly journal of experimental psychology*, 32 (1), pp. 3-25.

Prätorius, M., Valkov, D., Burgbacher, U. & Hinrichs, K. (2014). DigiTap: an eyes-free VR/AR symbolic input device. *Proc. of ACM Symposium on Virtual Reality Software and Technology (VRST '14)*, pp. 9-18.

Proctor, R.W. & Van Zandt, T. (2008). Human Factors in Simple and Complex Systems. *2nd edition, CRC Press*.

Proske, U. & Gandevia, S.C. (2009). The kinaesthetic senses. *J Physiol*, 587(17), pp 4139–4146.

Pruett, C., Kordi, M. & Damann, V. (2013). Process and Medical Analysis of the 2008-09 European Space Agency Astronaut Selection Campaign (ASC). *Aerosol Medical Association 84th Annual Scientif Meeting*.

Pruzin, S.J., Hurt, M.T., Compton, J.M. & Snyder, R.D. (2003). SSP 50253. *Operations Data File Standards. Rev J, NASA, International Space Station Program*, Johnson Space Center, Houston, 2003.

Porter, M.E. (1998). Competitive Strategy: Techniques for Analyzing Industries and Competitors. *Free Press*.

Porter, S.R., Marner, M.R., Smith, R.T., Zucco, J.E. & Thomas, B.H. (2010). Validating Spatial Augmented Reality for Interactive Rapid Prototyping. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '10)*, pp. 265-266.

Posner, M.I. (1980). Orienting of attention. *Journal of Experimental Psychology*, 32, pp. 3-25.

Poupyrev, I., Tomokazu, N., and Weghorst, S. (1998). Virtual notepad: Handwriting in immersive VR. *Proc. of the Virtual Reality Annual International Symposium (VRAIS ' 98)*, pp. 126–132.

Poupyrev, I., Tan, D., Billinghurst, M., Kato, H., Regenbrecht, H. & Tetsutani, N. (2001). Tiles: A mixed reality authoring interface. *Proc. of INTERACT'01*, pp. 334-341.

Raczynski, A. & Gussmann, P. (2004). Services and training through augmented reality. *Proc. of European Conf. on Visual Media Production (CVMP '04)*, pp. 263–271.

Regenbrecht, H., Baratoff, G. & Wilke, W. (2005). Augmented Reality Projects in the Automotive and Aerospace Industry, *IEEE Computer Graphics and Applications,* 25(6), pp. 48–56.

Reid, G.B., & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In: P.A. Hancock, N. Meshkati (Eds.), Human mental workload. Amsterdam: Elsevier, pp. 185–218.

Reiners, D., Stricker, D., Klinker, G. & Müller, S. (1999). Augmented Reality for Construction Tasks: Doorlock Assembly. *Proc. of Int. Workshop on Augmented Reality (IWAR '98)*, pp. 31-46.

Rekimoto, J. (1998). Matrix: A Realtime Object Identification and Registration Method for Augmented Reality. *Proc. of Asia Pacific Computer-Human Interaction (APCHI '98)*, pp. 63-68.

Reschke, M.F., Anderson, D.J. & Homick, J.L. (1984). Vestibulospinal reflexes as a function of microgravity. *Science*, 225(4658), pp. 212-214.

Reschke, M.F., Harm, D.L., Parker, D.E., Sandoz, G.R., Homick, J.L., & Vanderploeg, J.M. (1994). Neurophysiologic aspects: space motion sickness. In: A.E. Nicogossian, C. Leach Huntoon, S.L. Pool (Eds.), Space Physiology and Medicine. 3rd ed. Philadelphia: Lea & Febiger, pp. 228-260.

Pels, G.F., Compton, J.M., Banfield, R. & Boriack, S. (2004). Operations Data File Standards. *International Space Station Program. SSP 50253, Revision L,* International Space Station Program, NASA, Johnson Space Center, Houston, Texas, U.S.

Riemann, B.L. & Scott M. Lephart, S.M. (2002). The Sensorimotor System, Part I: The Physiologic Basis of Functional Joint Stability. *Journal of Athletic Training*, 37(1), pp. 71–79.

Robertson, C.M., MacIntyre, B. & Walker, B.N. (2008). An Evaluation of Graphical Context When the Graphics are Outside of the Task Area. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 73-76.

Roeber, H., Bacus, J. & Tomasi, C. (2003).Typing in Thin Air: The Canesta Projection Keyboard—A New Method of Interaction with Electronic Devices. *Proc. of Conf. on Human Factors in Computing Systems (CHI '03)*, pp. 712-713.

Rogers, R.D. & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231.

Rohs, M., Oulasvirta, A., Suomalainen, T. (2011). Interaction with magic lenses: real-world validation of a Fitts' Law model. In *Proc. of CHI'11*, pp. 2725-2728.

Roll, R., Gilhodes, J.C., Roll, J.P. & Popov, K. (1998). Proprioceptive information processing in weightlessness. *Exp Brain Res*, vol. 122, pp. 393-402.

Rowe, D., Silbert, J. & Irwin, D. (1998). Heart rate variability: indicator of user state as an aid to human-computer interaction. *Proc. of Conf. on Human Factors in Computing Systems (CHI '98)*, pages 480–487.

Salas-Moreno, R.F., Newcombe, R.A., Strasdat, H., Kelly, P.H.J. & Davison, A.J. (2013). SLAM++: Simultaneous Localisation and Mapping at the Level of Objects. *Proc. of CVPR 2013*, pp. 1352-1359.

Salmen, F., Thomson, C. & Müllerschkowski, U. (2011). European Payload Training for ISS Astronauts. *62nd International Astronautical Congress*, IAC-11-B6.3.2.

Salonen, T. & Sääski, J. (2008). Dynamic and Visual Assembly Instruction for Configurable Products Using Augmented Reality Techniques. In: X.-T. Yan, C. Jiang, B. Eynard (Eds.), Advanced Design and Manufacture to Gain a Competitive Edge. Springer, pp. 23-32.

Sanders, M.S. & McCormick, E.J. (1982). Human Factors in Engineering and Design. McGraw-Hill Inc.,US.

Scheid, F., Nitsch, A., König, H., Arguello, L., De Weerdt, D., Arndt, D. & Rakers, S. (2010). Operation of European SDTO at Col-CC. *Proc. of SpaceOps 2010 Conference*.

Schmalstieg, D., Encarnação, L.M. & Szalavári, Z. (1999). Using transparent props for interaction with the virtual table. *Proc. of the ACM Symposium on Interactive 3D Graphics (I3D '99)*, pp. 147-153.

Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavari, Z., Encarnação, L.M., Gervautz, M. & Purgathofer, W. (2002). The Studierstube Augmented Reality Project. *Presence: Teleoperators and Virtual Environments*, 11 (1), pp. 33-54.

Schmalstieg, D., Langlotz, T. & Billinghurst, M. (2010). Augmented Reality 2.0. In: G. Brunnett, S. Coquillart, G. Welch (Eds.), Virtual Realities, Dagstuhl Seminar 2008. Springer, 2011, pp. 13-37.

Schmalstieg, D. & Höllerer, T. (2016). Augmented Reality: Principles and Practice. *Addison Wesley*.

Schuber, M., Seibt, D. & Anken, R. (2013). BIOLAB on the International Space Station (ISS): Facility and Experiments. *Current Biotechnology*. ISSN 2211-5501.

Schuler, H. (2007). Assessment Center zur Potenzialanalyse. *Hogrefe Verlag*.

Schumacher, E. & Weimer, M. (2006). Multiple Vergleiche mit der SAS-Prozedur MIXED. In: K. Kaiser, R.-H., Bödeker (Eds.), Statistik und Datenanalyse mit SAS. Proceedings der 10. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE), pp. 171-187.

Schwald, B., Figue, J., Chauvineau, E., Vu-Hong, F., Robert, A., Arbolino, M., Schnaider, M., Laval, B. de, Dumas de Rauly, F., Anez, F.G., Baldo, O. & Santos, J. (2001). STARMATE: Using augmented reality technology for computer guided maintenance of complex mechanical elements. *Proc. of eBusiness and eWork Conference (e2001)*, pp.196-202.

Schwald, B. & Laval, B.D. (2003). An Augmented Reality System for Training and Assistance to Maintenance in the Industrial Context. *Proc. of Winter School of Computer Graphics*, pp. 425-432.

Schwerdtfeger, B., Frimor, T., Pustka, D. and Klinker, G. (2006). Mobile Information Presentation Schemes for Supra-adaptive Logistics Applications. In. Z. Pan, A. Cheok, M. Haller, R.W.H Lau, H. Saito and R. Liang (Eds.), Advances in Artificial Reality and Tele-Existence. 16th International Conference on Artificial Reality and Telexistence (ICAT '06), *Springer Berlin Heidelberg*. pp. 998-1007.

Schwerdtfeger B. & Klinker, G. (2008). Supporting Order Picking with Augmented Reality. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 91-94.

Schwerdtfeger, B., Reif, R., Gunthner, W.A., Klinker, G., Hamacher, D., Schega, L., Bockelmann, I., Doil, F., & Tumler, J. (2009). Pick-by-Vision: A First Stress Test. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '09)*, pp. 115-124.

Seichter, H., Looser, J. & Billinghurst, M. (2008). ComposAR: An intuitive tool for authoring AR applications. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'08)*, pp. 177-178.

Selvatici, A.H.P., Costa, A.H.R. & Dellaert, F. (2008). Object-based Visual SLAM: How Object Identity Informs Geometry. *Proc. of IV Workshop de Visao Computacional (WVC)*, pp. 17-19.

Sharp, H., Rogers, Y. & Preece, J. (2007). Interaction Design - beyond human-computer interaction. *2nd edition, John Wiley & Sons Ltd*.

Shazad, M. (2011). Detection and tracking of pointing hand gestures for AR applications. *Master Thesis*, Hochschule Bonn-Rhein Sieg.

Shi, J. & Tomasi, C. (1994). Good features to track. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593 - 600.

Shiga, D. (2009). Proposed Mars Trip would exceed NASA's Radiation limits. *New Scientist*, 203(2726), pages 11.

Shneiderman, B. (1987). Designing the User Interface: Strategies for Effective Human-Computer Interaction. *Addison-Wesley Publ. Co., Reading, MA, USA*.

Shneiderman, B., Plaisant, C., Maxine, C. & Jacobs, S. (2009). Designing the User Interface: Strategies for Effective Human-Computer Interaction: Fifth Edition. *Addison-Wesley Publ. Co., Reading, MA, USA*.

Simons, D.J. & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*. 28(9), pp. 1059-74.

Skytec. (2015). Successful experiment execution on board ISS. URL: http://www.skytek.com /successful-experiment-execution-board-iss/, last visit: 28.09.2016.

Smetanin, B.N. & Popov, K.E. (1997). Effect of body orientation with respect to gravity on directional accuracy of human pointing movements. *Eur J Neurosci*, 9(1), pp. 7-11.

Smets, N. & Neerincx, M. (2013). CRUISE Evaluation Report. *TNO 2013 R11379*, URL: http://publications.tno.nl/publication/34606658/KqUvfk/TNO-2013-R11379.pdf, last visit: 28.09.2016.

Steen, van der J. (2009). Vestibulo-Ocular Reflex (VOR). In: M.D. Binder, N. Hirokawa, U. Windhorst (Eds.), Encyclopedia of Neuroscience, Springer Berlin Heidelberg, pp. 4224-4228.

Stelges, K. & Lawrence, B.A. (2012). Human Engineering Modeling and Performance: Capturing Humans and Opportunities. In: V.G. Duffy (Ed.), Advances in Applied Human Modeling and Simulation. CRC Press, pp. 555-564.

Stepanek, J. & Webb, T. (2008). Physiology of Decompressive Stress. In: J.R. Davis, R. Johnson, J. Stepanek (Eds.), Fundamentals of Aerospace Medcine. 4th Edition, Lippincott Williams and Wilkins, pp. 46-82.

Stoakley, R., Conway, M. & Pausch, R. (1995). Virtual Reality on a WIM. *Proc. of the ACM SIGCHI Human Factors in Computer Systems Conference*, pp. 265-272.

Stone, D., Jarett, C., Woodroffe, M. & Minocha, S. (2005). User Interface Design and Evaluation. *Morgan Kaufmann*.

Sutherland, I. E. (1968). A head-mounted three dimensional display. *Proc. of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)*, pp. 757-764.

Swan, J.E. & Gabbard, J.L. (2005). Survey of User-Based Experimentation in Augmented Reality. *Proc. First International Confernce on Virtual Reality (VR '05)*.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), pp. 257-285.

Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty and Instructional Design. *Learning and Instruction*, vol. 4, pp. 293-312.

Sweller, J. & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. Cognition and Instruction, 2(1), pp. 59-89.

Sweller, J., van Merrienboer J.J.G & Paas F.G.W.C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), pp. 251-296.

Sweller, J. (2008) Human Cognitive Architecture. In: J. Spector, M. Merrill, J. Van Merrienboer, M. Driscoll (Eds.), Handbook of Research on Educational Communications and Technology, edn. 3rd, Erlbaum, NY, pp. 369 - 381

Syberfeldt, A., Danielsson, O., Holm, M. & Wang, L. (2015). Visual Assembling Guidance Using Augmented Reality. *Procedia Manufacturing*, vol. 1, pp. 98-109.

Tan, W., Liu, H., Dong, Z., Zhang, G. & Bao, H. (2013). Robust Monocular SLAM in Dynamic Environments. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '13)*, pp. 209-218.

Tang, A., Owen, C., Biocca, F. & Mou, W.M. (2003). Comparative Effectiveness of Augmented Reality in Object Assembly. *Proc. of the ACM SIGCHI Conference on Human Factors in Computing (CHI'03)*, pp. 73-80.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart Rate Variability: standards of measurement, physiological interpretation and clinical use. (1996). *Circulation* 1996 (93), pp. 1043-1065 / *Eur Heart J*, 17(3), pp. 354-81.

Teather, R.J. & Stuerzlinger, W. (2011). Pointing at 3D targets in a stereo head-tracked virtual environment. *Proc. of IEEE Symposium on 3D User Interfaces (3DUI),* pp. 87-94.

Thompson, A.A. & Henriques, D.Y.P. (2011). The coding and updating of visuospatial memory for goal-directed reaching and pointing. *Vision Research*, 51(8), pp. 819–826.

Tingdahl, D., De Weerdt, D., Vergauwen, M. & Van Gool, L. (2011). WEAR++: 3D model driven camera tracking on board the International Space Station. *Proc. of International Conference on 3D Imaging (IC3D '11)*, pp. 1-8.

Tomasi, C., Rafii, A. & Torunoglu, I. (2003). Full-Size Projection Keyboard for Handheld Devices. *Communications of the ACM - A game experience in every application*, vol. 46, no. 7, pp. 70-75.

Tönnis, M., Sandor, C., Lange, C., Klinker, G., Lange, C. & Bubb, H. (2005). Experimental Evaluation of an Augmented Reality Visualization for Directing a Car Driver's Attention. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '05)*, pp. 56-59.

Tönnis, M. & Klinker, G. (2006). Effective Control of a Car Driver's Attention for Visual and Acoustic Guidance Towards the Direction of Imminent Dangers. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '06)*, pp. 13-22.

Treisman, A. & Gelade, G. (1980). A Feature Integration Theory of Attention. *Cognitive Psychology*, vol. 12, pp. 97-136.

Treisman, A. & Schmidt, H. (1982). Illusory Conjunctions in the Perception of Objects. *Cognitive Psychology*, vol. 14, pp. 107-141.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, vol. 255, pp. 114-125.

Tsang, P.S. & Velasquez, V.L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, 39(3), pp. 358-381.

Tuceryan, M., Genc, Y. & Navab, N. (2002). Single-Point Active Alignment Method (SPAAM) for Optical See Through HMD Calibration for Augmented Reality. *Prescence : Teleoperators & Virtual Environments*, 11(3), pp. 259-276.

Tuday, E.C., Meck, J.V., Nyhan, D., Shoukas, A.A. & Berkowitz D.E. (2007). Microgravity-induced changes in aortic stiffness and their role in orthostatic intolerance. *J Appl Physiol*, 102(3), pp. 853-858.

Tümler, J., Mecke, R., Schenk, M. , Huckauf, A., Doil, F., Paul, G., Pfister, E., Böckelmann, I. & Roggentin, A. (2008). Mobile Augmented Reality in Industrial Applications: Approaches for Solution of User-Related Issues. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 87-90.

Uhlig, T., Mannel, T., Fortunato, A. & Illmer, N. (2015). Space-to-Ground Communication for Columbus: A Quantitative Analysis. *Scientific World Journal*. 2015: 308031.

Wagner, D. & Schmalstieg, D. (2003). First Steps Towards Handheld Augmented Reality. *Proc. of the 7th International Conference on Wearable Computers (ISWC '03)*, pp. 127-135.

Wagner, D. (2007). Handheld Augmented Reality. Thesis Doctoral, Graz University of Technology, Graz, Austria.

Wang, R.F., Crowell, J.A., Simons, D.J., Irwin, D.E., Kramer, A.F., Ambinder, M.S. & Hsieh, B.B. (2006). Spatial updating relies on an egocentric representation of space: Effects of the number of objects. *Psychonomic Bulletin & Review*, 13, 281–286.

Wang, C.Y., Chu, W.C., Chiu, P.T., Hsiu, M.C., Chiang, Y.H. & Chen, M.Y. (2015). PalmType: Using Palms as Keyboards for Smart Glasses. *Proc. of International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*, pp. 153-160.

Watson, N.V. & Breedlove, S.M. (2012). The Mind's Machine: Foundations of Brain and Behavior. *Sinauer Associates, Inc*., 1st ed.

Watt, D.G.D. (1997). Pointing at memorized targets during prolonged microgravity. *Aviation, Space, and Environmental Medicine*, vol. 68, pp. 99–103.

Webel, S., Bockholt, U., Engelke, T., Peveri, Olbrich, M & Preusche, C. (2011a). Augmented Reality Training for Assembly and Maintenance Skills. *Proc. of BIO Web of Conferences*, pp. 1-4.

Webel, S. Bockholt, U. and Keil, J. (2011b). Design Criteria for AR-Based Training of Maintenance and Assembly Tasks. *Elsevier, Virtual and Mixed Reality*, Part I, HCII 2011, LNCS 6773, pp. 123-132.

Weidenhausen, J., Knoepfle, C. & Stricker, D. (2003). Lessons learned on the way to industrial augmented reality applications, a retrospective on ARVIKA. *Computers and Graphics*, 27(6), pp. 887-889.

Weigel, M., Mehta, V. & Steimle, J. (2014). More than touch: understanding how people use skin as an input surface for mobile computing. *Proc. of SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, pp. 179-188.

Weimer, D. & Ganapathy, S.K. (1989). A Synthetic Visual Environment with Hand Gesturing and Voice Input. *Proc. of ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '89)*, pp. 235-240.

Weiser, M. (1993). Hot topics-ubiquitous computing. IEEE Computer Society, October 1993, 26(10), pp. 71-72.

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. In *Ergonomics*, 3, 189-230.

Wickens, C.D. (1984). Processing resources in attention. In: D.R. Davies (Ed.), Varieties of Attention, Academic Press, New York City, NY., pp. 63-258.

Wickens, C.D. & Hollands, J.G. (2000). Engineering Psychology & Human Performance. *3th edition, Upper Saddle River, NJ: Prentice Hall.*

Wickens, C.D, Lee, J., Liu, Y.D. & Gordon-Becker, S. (2003). An Introduction to Human Factors Engineering. *2n edition, Pearson Education Limited*.

Wickens, C. & Alexander, A. (2009). Attentional tunneling and task management in synthetic vision displays. *International Journal of Aviation Psychology*, 19(2), pp. 182-199.

Wiedenmaier, S., Oehme, O., Schmidt, L. & Luczak, H. (2003). Augmented Reality (AR) for Assembly Processes Design and Experimental Evaluation. *International Journal of Human-Computer Interaction*, 16(3), pp. 497-514.

Wierwille, W.W. (1979). Physiological Measures of Aircrew Mental Workload. *Human Factors*, 21(5), pp. 575-593.

Williges, R.C & Wierwille, W.W. (1979). Behavioral Measures of Aircrew Mental Workload. *Human Factors and Ergonomics Society*, 21(5), pp. 549-574.

Wither, J., DiVerdi, S. & Höllerer, T. (2007). Evaluating Display Types for AR Selection and Annotation. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '07)*, pp. 1-4.

White, S. & Feiner, S. (2009). SiteLens: Situated visualization techniques for urban site visits. *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp. 1117-1120.

Whiteside, T.C.D. (1961). Hand-Eye Coordination in Weightlessness. *Aerosp Med*, vol. 32, pp. 719-725.

Wolff, M. (2011). Crusade - Crew usability demonstrator. URL: http://spaceflight.esa.int/eo/EOI /crew-info-sys/bkgnd/ESA-Crusade-briefing-ISTAR.pdf, last visit: 28.09.2016.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review*, vol. 3, pp. 1-114.

Yavari, F., Towhidkhah, F., Ahmadi-Pajouh, M.A & Darainy, M. (2015). The role of internal forward models and proprioception in hand position estimation. *J Integr Neurosci*, 14(3), pp. 403-18.

Yerkes, R.M. & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Comparative Neurology and Psychology*, 18(5), pp. 459-482.

Young, L.R., Oman, C.M., Watt, D.G., Money, K.E. & Lichtenberg, B.K. (1984). Spatial orientation in weightlessness and readaptation to earth's gravity. Science. 225(4658), pp. 205-208.

Young, L.R., Oman, C.M., Watt, D.G.D., Money, K.E., Lichtenberg, B.K., Kenyon, R.V. & Arrott, A.P. (1986). MIT/Canadian vestibular experiments on the spacelab-1, mission 1. Sensory adaptation to weightlessness and readaptation to one-g: an overview. *Exp Brain Res*, vol. 64, pp. 291–298.

Young, L.R. & Shelhamer, M. (1990). Microgravity enhances the relative contribution of visually-induced motion sensation. *Aviat Space Environ Med*, vol. 61, pp. 525–530.

Young, L.R., Jackson, D.K., Grouleau, N. & Modestino, S. (1992). Multisensory Integration in Microgravity. *Ann N Y Acad Sci*, vol. 656, pp. 340-353.

Young, L.R., Oman, C.M., Merfeld, D., Watt, D., Roy, S., DeLuca, C., Balkwill, D., Christie, J., Groleau, N., Jackson, D.K., Law, G., Modestino, S. & Mayer, W. (1993). Spatial orientation and posture during and following weightlessness: human experiments on Spacelab Life Sciences 1. *J Vestib Res*, 3(3), pp. 231-239.

Young, L.R., Mendoza, J., Groleau, N., Wojcik & P.W. (1996) . Tactile influences on astronaut visual spatial orientation: human neurovestibular studies on SLS-2. *J Appl Physiol*, vol. 81, pp. 44–49.

Yuan, M. L., Ong, S. K. & Nee, A. Y. C. (2008). Augmented reality for assembly guidance using a virtual interactive tool. *International Journal of Production Research*, 46(7), pp. 1745-1767.

Zauner, J., Haller, M., Brandl, A. & Hartman, W. (2003). Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '03)*, pp. 237–246.

Zhou, F., Duh, H.B.-L., & Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 193-202.

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Daniela Markov-Vetter Rostock, 14.10.2016

Theses

Author:	M.Sc. Dipl.-Inf. (FH) Daniela Markov-Vetter
Title:	Human Factors in Instructional Augmented Reality for Intravehicular Spaceflight Activities and How Gravity Influences the Setup of Interfaces Operated by Direct Object Selection
Objective:	To improve the work of astronauts during intravehicular payload operations by means of advanced user interfaces using Augmented Reality.

The success of space missions vitally depends on human performance to complete tasks accurately in a reasonable time. While astronauts are carrying out maintenance and mission specific tasks at International Standard Payload Racks using standardized procedures, they obtain the necessary information and instructions from a laptop computer. This is not only challenged by the human-computer interaction, but also by the adaptation to weightlessness. Because the laptop computer, providing the guidance interface, is laterally installed at the rack construction, astronauts are forced to change their view constantly between the laptop and the region of operation, which can lead to loss of focus and attention, as well as implies increased costs of human information processing, which is conditioned by a high perceptual and cognitive load during the reception of the instruction information and the subsequent task localization. As a consequence, the performance quality of payload operations can suffer from a higher workload level and thus bears the risk of astronaut errors. Using an integrated design, as provided by Augmented Reality (AR), not only ensures that all relevant information is embedded into one field of view, but is also capable to cue visual attention during task localization. Hence, AR has the greatest potential to improve the work of astronauts during intravehicular payload operations. Besides closing a fundamental gap, which requests a field test in a summative way, it is important to clarify where task cues can be displayed to efficiently support localizing and detecting off-screen and in-view targets, which is addressed to solve the trade-off between exocentric and egocentric visualization schemes. Once AR is adopted to assist payload operations, another basic problem is arisen from the lack of appropriate input devices to, for example, enter command-based instructions. Assuming an AR interface is operated by direct touching, it can be placed in a world-, head- or hand-reference. Because the underlying visuomotor coordination will be affected by modified gravity conditions, especially by weightlessness, it is necessary to find out the most adequate reference frame for AR input devices.

The following theses are based on the results of the research conducted in this dissertation:

1. AR is capable to fulfill the requirements of intravehicular payload operations.
2. AR reduces sequence errors and ensures the sequential mode of payload operations, and thus provides a higher level of performance quality than the common guidance method.
3. AR decreases the workload during payload operations compared to the common guidance method, but only for those tasks, which vary in their instructions, requiring constant learning.
4. AR enables a higher level of satisfaction and acceptance from domain application experts compared to the common payload guidance method.

5. Target cues presented by an egocentric display, like a head-mounted display, enhance the visuomotor performance of task localization, compared to an exocentric display, as provided by the common payload guidance interface.
6. Assuming an egocentric display, an egocentric AR visualization of target cues accelerates the visuomotor performance compared to an exocentric screen-stabilized area map.
7. Compared to an egocentric AR navigation, an exocentric screen-stabilized area map leads to the lowest effort spent on dual-task performance during the search of the off-screen targets.
8. Compared to egocentric AR target cueing, an exocentric screen-stabilized area map provides the most relaxed state and thus the lowest level of arousal during task localization.
9. The subjective workload benefits more from an egocentric AR visualization of target cues than from those presented by an exocentric display, as provided by the common guidance interface.
10. Target cues presented by an exocentric area map within an egocentric display provoke egocentric spatial updating during the task localization process.
11. Human individual predisposition for visual spatial processing needs to be considered in AR research, which is addressed to user's spatial navigation skills.
12. Users following an egocentric strategy are performing better with an egocentric display and presentation scheme than those one following an exocentric strategy.
13. Users following an exocentric strategy consume less processing resources during task localization than those one following an egocentric strategy.
14. AR is operable under altered gravity conditions.
15. Handheld AR input devices enhance the quality of visuomotor coordination under short-term hyper- and microgravity conditions compared to world- and head-referenced interfaces.
16. Haptic feedback is superior to non haptic feedback for goal-directed pointing tasks under normogravity.
17. Haptic feedback enhances the quality of visuomotor coordination while operating AR input devices under short-term hyper- and microgravity.
18. Short-term hyper- and microgravity affects visuomotor performance on world-, hand and head-referenced AR interfaces.
19. The visuomotor performance under normogravity does not differ between AR interfaces referenced to the world and those referenced to the hand.
20. Gravity-adapted target resizing affects the performance and workload of direct pointing towards world-referenced AR interfaces under hypergravity conditions.
21. World-referenced AR interfaces, which are resized in an expanded manner, improves the visuomotor performance under hypergravity conditions compared to unmodified and compressed resized one.
22. Increased sensorimotoric load attached to the pointing arm improves the visuomotor performance while handling a nonhaptic head-referenced interface under normogravity.
23. Increased sensorimotoric load attached to the pointing arm does not affect visuomotor performance on a haptic interface, referenced to world objects.