Forschungsbericht 2026-20

Using Mixed Reality in Simulator Based Flight Training for Helicopters

Tanja Dikarew-Martini

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Flugsystemtechnik
Braunschweig

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Forschungsbericht 2026-20

Using Mixed Reality in Simulator Based Flight Training for Helicopters

Tanja Dikarew-Martini

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Flugsystemtechnik
Braunschweig

162 Seiten
58 Bilder
17 Tabellen
96 Literaturstellen

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2026-20
Erscheinungsjahr 2026
DOI: 10.57676/k56p-1k27

Erklärung des Herausgebers
Dieses Werk – ausgenommen anderweitig gekennzeichneter Teile – ist lizensiert unter den Bedingungen der Creative Commons Lizenz vom Typ Namensnennung 4.0 International (CC BY 4.0), abrufbar über https://creativecommons.org/licenses/by/4.0/legalcode

Lizenz
Creative Commons Attribution 4.0 International

Hubschraubersimulation, Mixed Reality Headsets, Pilotenausbildung

Tanja DIKAREW-MARTINI
DLR, Institut für Flugsystemtechnik, Braunschweig

Einsatz von Mixed Reality in der simulationsbasierten Flugausbildung für Hubschrauber
Technische Universität Clausthal

Diese Forschungsarbeit untersucht das Potenzial von Mixed Reality (MR) Headsets, die Ausbildung von Hubschrauberpiloten in einem Simulator zu ergänzen oder zu verbessern. Es wurden drei Studien unter Verwendung von zwei Simulatoren durchgeführt: einem einfachen Hubschraubersimulator mit Fernsehbildschirmen als visuelles Anzeigesystem und einem voll ausgestatteten EC135-Simulator mit einem Domprojektionssystem. In der ersten Studie wurden das monitorbasierte Varjo XR-3-Headset und das See-Through-Headset JVC HMD-VS1D mit herkömmlichen Fernsehbildschirmen verglichen. Obwohl die Flugleistung der Piloten bei allen Systemen vergleichbar war, erzielte die Varjo-MR-Konfiguration höhere Bewertungen in der Präsenz. Ebenfalls sind bei den Piloten kaum Simulationskrankheitssymptome aufgetreten, wodurch das Varjo-Headset zur bevorzugten Wahl für nachfolgende Studien wurde. Für die zweite Studie wurde das Varjo-Headset in den voll ausgestatteten EC135-Simulator integriert. Es zeigte sich, dass die Interaktion im Cockpit unter Verwendung des MR-Headsets ebenfalls machbar war, wenngleich es Einschränkungen gab, wie beispielsweise das eingeschränkte Sichtfeld. Eine vergleichende Analyse der ersten beiden Studien ergab eine positive Korrelation zwischen Präsenz und Flugleistung sowie ein spezifisches Blickverhalten mit dem MR-Headset, bei dem sich die Piloten stärker auf externe visuelle Hinweise konzentrierten. In der dritten Studie wurden standardisierte Flugaufgaben, die aus Checkflügen abgeleitet wurden, zur Bewertung neuer visueller Systeme herangezogen. Insgesamt haben sich MR-Headsets als geeignet für die Ausbildung von Hubschrauberpiloten erwiesen: Obwohl sie derzeit weniger immersiv und komfortabel sind als voll ausgestattete Simulatoren mit Domprojektionssystemen, übertreffen sie einfache Simulatoren mit Fernsehbildschirmen hinsichtlich der Präsenz.

Helicopter Simulation, Mixed Reality Headsets, Pilot Training

(Published in English)

Tanja DIKAREW-MARTINI
German Aerospace Center (DLR), Institute of Flight Systems, Braunschweig

Using Mixed Reality in Simulator Based Flight Training for Helicopters
Clausthal University of Technology

This research investigates the potential of mixed reality (MR) headsets to supplement or enhance helicopter pilot training in a simulated environment. Three studies were conducted using two simulators: a basic helicopter simulator with TV screens as the visual display system and a fully equipped EC135 simulator with a dome projection system. In the first study, the monitor-based Varjo XR-3 headset and the see-through JVC HMD-VS1D headset were compared with conventional TV screens. Although the pilots' flight performance was comparable across all systems, the Varjo MR configuration achieved superior presence ratings and negligible simulation sickness among the pilots, making the Varjo headset the preferred choice for subsequent studies. For the second study, the Varjo headset was integrated into the fully equipped EC135 simulator. It was demonstrated that interaction within the cockpit using the MR headset was feasible, although there were limitations, such as the restricted field of view. A comparative analysis of the first two studies revealed a positive correlation between presence and flight performance, as well as distinct gaze behavior in MR, with pilots focusing more strongly on external visual cues. In the third study, standardized flight tasks, derived from proficiency flights, were used to evaluate new visual systems. Overall, MR headsets have proven to be suitable for the training of helicopter pilots: although they are currently less immersive and comfortable than fully equipped simulators with dome projection systems, they outperform simple simulators with television screens in terms of presence.

Using Mixed Reality in Simulator Based Flight Training for Helicopters

D o c t o r a l T h e s i s
(D i s s e r t a t i o n)

to be awarded the degree

Doctor of Engineering (Dr.-Ing.)

submitted by
Tanja Dikarew-Martini

from Frankfurt am Main

approved by the Faculty of Science
Clausthal University of Technology,

Date of oral examination
17 June 2026

Dean
Prof. Dr. René Wilhelm

Chairperson of the Board of Examiners
Prof. Dr. Kai Kunze

Supervising Tutor
Apl.-Prof. Dr.-Ing. Umut Durak

Reviewer
Prof. Dr. Sven Hartmann
Assoc. Prof. Umberto Saetti

Keywords: Helicopter Simulation, Mixed Reality Headsets, Pilot Training

Abstract

The integration of Mixed Reality (MR) technology into flight training is a promising solution to meet the growing demand for pilots, as the aviation industry seeks more flexible and cost-effective training solutions. This dissertation investigates the suitability of MR headsets in helicopter pilot training, specifically focusing on whether MR can enhance or complement traditional simulators. To address these objectives, three studies were conducted using two distinct helicopter simulators: two studies in a basic dual pilot station simulator with TV screens and one study in the fully equipped EC135 simulator with a projection dome.

The first study compared a monitor-based headset (Varjo XR-3 Focal Edition) with a Mixed Reality (MR) and Virtual Reality (VR) configuration against a see-through headset (JVC HMD-VS1D) with Augmented Reality (AR) configuration and conventional TV displays as the visual system. The results showed that while flight performance was similar across all systems, the Varjo MR configuration yielded higher presence ratings with negligible simulation sickness, leading to the selection of the Varjo headset with the MR configuration for further testing.

In the second study, the Varjo headset was integrated into the fully equipped EC135 cockpit. This study demonstrated the feasibility of a hybrid MR approach using chroma key and cutouts to minimize visual artifacts, enabling successful cockpit interaction and malfunction procedures, however, challenges such as restricted Field of View (FOV) limitations persisted.

A comparative analysis of the first two studies' results found a positive correlation between presence and flight performance. Additionally, the analysis identified distinct gaze behaviors in MR setups, where pilots tended to focus more on external visual cues than with other visual configurations.

The third study focused on the evaluation of proposed standardized flight tasks based on annual proficiency checks to assess new visual systems, such as MR headsets, for helicopter pilot training in industrial settings and the applicability of the used assessment methods.

The findings show that MR headsets can serve as suitable visual systems for helicopter pilot training. While MR currently falls short of the presence and comfort provided by projection domes, it offers higher presence and comparable flight performance to basic mock-up simulators. Furthermore, MR technology offers significant advantages in terms of flexibility and compactness, allowing for easier deployment and increased accessibility of training facilities. As MR technology advances, it holds the potential to enhance or complement training conducted in full-flight simulators.

Acknowledgements

First, sincere thanks go to Professor Umut Durak for his valuable advice and for always finding time for thoughtful discussions.
I am also grateful to the head of the rotorcraft department, Marc Höfinger, for his support and the offer to regularly debate the progress of this work. A special thank-you to my team leader, Malte-Jörn Maibach, for creating a supportive environment for my work and for his help within the ARiTra project. Additional thanks go to my colleague Daniel Greiwe for his assistance throughout the studies conducted for this thesis.
Deep gratitude to my family, whose unwavering support and continual encouragement have been a source of strength throughout the academic journey. A special thank-you goes to my husband, Alexej Dikarew-Martini, for his comprehensive support, whether brainstorming, lending a sympathetic ear or offering programming tips whenever I needed them.
Finally, appreciation is expressed to the ARiTra project partners for the exchange of expertise that enriched this study.

Thank you all for making this work possible.

Braunschweig, June 2026 *Tanja Dikarew-Martini*

Contents

Acknowledgements V

List of Figures X

List of Tables XII

Acronyms XIII

1 Introduction 1
- 1.1 State of the Art . . . 5
  - 1.1.1 Current Helicopter Simulators . . . 5
  - 1.1.2 Current Helicopter Pilot Training . . . 7
  - 1.1.3 Virtual (VR), Mixed (MR) and Augmented Realities (AR) in Research 9
- 1.2 Contribution of This Dissertation . . . 17
  - 1.2.1 Research Objectives . . . 17
  - 1.2.2 Research Questions . . . 18
  - 1.2.3 Outline of the Thesis . . . 19

2 Reference System Architectures for Mixed Reality Helicopter Simulation 21
- 2.1 Simulation Environment . . . 21
  - 2.1.1 Air Vehicle Simulator (Air Vehicle Simulator (AVES)) . . . 21
  - 2.1.2 Dual Pilot Active Sidestick Demonstrator (2PASD) . . . 24
  - 2.1.3 See-through and Monitor-based Headsets Used in This Work . . . 25
- 2.2 Enhancing Helicopter Simulation with See-through and Monitor-based Headsets . . . 27
  - 2.2.1 Selection of Simulator Platform for Mixed Reality (MR) Development 27
  - 2.2.2 Simulation Framework . . . 28
  - 2.2.3 Rendering Software . . . 31
  - 2.2.4 Techniques for Creating Mixed Reality (MR) Configurations With Monitor-based Headsets . . . 33
  - 2.2.5 Hardware Setup in Dual Pilot Active Sidestick Demonstrator (2PASD) 35
  - 2.2.6 Latency Measurement . . . 38

3 Potentials and Limitations of Mixed Reality (MR) in Helicopter Simulation 45
- 3.1 Study 1: Comparison of See-Through and Monitor-Based Headsets . . . 45
  - 3.1.1 Experimental Setup . . . 45
  - 3.1.2 Task Definition . . . 48
  - 3.1.3 Assessment Methods . . . 51
  - 3.1.4 Pilot Experience . . . 59

3.2 Results of Study 1 60
3.2.1 Results of Questionnaires 60
3.2.2 Results of Flight Performance 65
3.2.3 Further Findings 69
3.2.4 Conclusion of Study 1 and Selection of Mixed Reality (MR) Headset for Further Studies 70
3.3 Mixed Reality (MR) Integration Into Fully Equipped EC135 Helicopter Cockpit 72
3.4 Study 2: Investigation of Mixed Reality (MR) in a Fully Equipped Cockpit . 75
3.4.1 Experimental Setup 75
3.4.2 Task Definition 76
3.4.3 Assessment Methods 78
3.4.4 Pilot Experience 79
3.5 Results of Study 2 80
3.5.1 Results of the Questionnaires 80
3.5.2 Flight Performance Results 88
3.6 Comparative Analysis of Flight Performance for Both Studies 92
3.6.1 Overview of all Evaluated Configurations 92
3.6.2 Evaluation of all Configurations Regarding PQ in Relation to Flight Performance 94
3.6.3 Eye Tracking Evaluation 102
3.7 Specific Implementations of MR in Helicopter Simulation 104
3.7.1 Setup Characteristics of the Evaluated VR, MR and AR Configurations104
3.7.2 Mixed Reality (MR) Implementations in Two Different Simulation Environments (2PASD and AVES) 105
3.7.3 Differences in Application of Monitor-based and See-Through Headset106
3.7.4 Answer to Subordinate Research Question (1.1) 109
3.8 Potentials and Limitations of Using MR in Helicopter Simulation 111
3.8.1 Summarized Pilot Feedback and Further Observations of Configurations in Dual Pilot Active Sidestick Demonstrator (2PASD) 111
3.8.2 Summarized Pilot Feedback and Further Observations of Configuration in Air Vehicle Simulator (AVES) 113
3.8.3 Answer to Subordinate Research Question (1.2) 115

4 Evaluation Assessment Methods for MR Helicopter Simulation 117
4.1 Development of Flight Maneuvers for Evaluation of Mixed Reality (MR)-based Visual Systems 117
4.1.1 Aerial Work Flight Task 118
4.1.2 Ground Work Flight Task 118
4.2 Study 3: Proposing a Standard Maneuver for Evaluation of New Visual Systems for Helicopter Simulation 121
4.2.1 Experimental Setup 121
4.2.2 Task Definition 122

4.2.3 Assessment Methods . . . 123
4.2.4 Pilot Experience . . . 123
4.3 Results of Study 3 . . . 124
4.3.1 SSQ Results of Flight Tasks . . . 124
4.3.2 Eye Tracking Evaluation of Flight Tasks with the Conventional Setup 126
4.3.3 Pilot Feedback on Aerial Work Task . . . 127
4.3.4 Pilot Feedback on Ground Work Task . . . 128
4.3.5 Conclusion of Study 3 . . . 129
4.4 Discussion of the Applicability of Assessment Methods . . . 130
4.4.1 Assessment with Standardized Questionnaires . . . 130
4.4.2 Assessment of Flight Performance . . . 133
4.4.3 Correlation of Presence Questionnaire (PQ) With Mission-Task-Elements (MTE) Flight Performance . . . 134
4.4.4 Assessment Using Proposed Flight Tasks . . . 135
4.4.5 Assessment With Eye Tracking . . . 136
4.4.6 Answer to Subordinate Research Question (1.3) . . . 138

5 Overall Assessment and Knowledge Transfer 139
5.1 Answering the Main Research Question . . . 139
5.2 Knowledge Transfer for Industrial Applications . . . 141

6 Conclusion and Outlook 145
6.1 Conclusion . . . 145
6.2 Outlook . . . 148

A Standardized Questionnaires 149
A.1 Presence Questionnaire [65] . . . 149
A.2 Slater, Usoh, and Steed (SUS) Questionnaire (A Presence Questionnaire) [67] . . . 151
A.3 Simulation Sickness Questionnaire . . . 153

Bibliography 155

List of Figures

1.1 LINK Trainer . . . 2
1.2 Flight Simulators at German Aerospace Center (DLR) . . . 3
1.3 Reality-Virtuality (RV) Continuum by Milgram et al. Adapted From [17], ©1994 IEICE. . . . 9
1.4 Headset Types and Functionality Adapted From [20] . . . 11
1.5 Overview of the Chapters 3 to 5 . . . 20

2.1 Air Vehicle Simulator AVES . . . 22
2.2 AVES Dome Projection Resolution ©project:syntropy GmbH/domeprojection.com . . . 23
2.3 Dual Pilot Active Sidestick Demonstrator (2PASD) Simulator ©DLR. CC BY-NC-ND . . . 24
2.4 Mixed Reality Headsets . . . 25
2.5 Field of View Comparison Adapted from [53], CC BY 4.0 . . . 26
2.6 2PASD System Architecture . . . 29
2.7 MR/VR System Architecture . . . 30
2.8 Different Mixed Reality (MR) Rendering Options . . . 32
2.9 Mixed Reality (MR) Software Implementations in Dual Pilot Active Sidestick Demonstrator (2PASD) . . . 34
2.10 Greenscreen Setup in 2PASD ©DLR. CC BY-NC-ND . . . 36
2.11 Mixed Reality (MR) Setup With Varjo Headset in 2PASD . . . 37
2.12 JVC Setup with "Self-illuminated Black" Colored Regions Used for Masking ©DLR. CC BY-NC-ND . . . 38
2.13 Time Delay Analysis . . . 40
2.14 2PASD Signal Flow in Basic Simulation Setup . . . 41
2.15 2PASD Signal Flow With Experimental System . . . 42
2.16 Time Delays of Both Setups . . . 43

3.1 Visual System Configurations for Study 1 . . . 47
3.2 Mission-Task-Elements Test Course - Visual Cues for Hover Maneuver ©DLR. CC BY-NC-ND . . . 49
3.3 Slalom Test Course ©DLR. CC BY-NC-ND . . . 50
3.4 Different Eye Tracking Systems . . . 59
3.5 PQ Total Scores (TS) . . . 61
3.6 PQ Subscales . . . 62
3.7 MTE Performance of Pilot B . . . 69
3.8 Air Vehicle Simulator (AVES) EC135 Cockpit With Colored Projection Dome for Mixed Reality (MR) Setup ©DLR. All rights reserved. . . . 72

3.9 Mixed Reality (MR) System Architecture in AVES 73
3.10 Combination of Chroma Keying and Masking for Mixed Reality (MR) Environment in Air Vehicle Simulator (AVES) 74
3.11 Combination of Chroma Keying and Masking for Mixed Reality (MR) Environment in Dual Pilot Active Sidestick Demonstrator (2PASD) 74
3.12 Visual System Configurations for Study 2 76
3.13 Cockpit Instruments ©DLR. CC BY-NC-ND 78
3.14 PQ Total Score for all Three Maneuvers in Air Vehicle Simulator (AVES) . . . 82
3.15 PQ Subscales 83
3.16 SSQ for all Three Maneuvers in Air Vehicle Simulator (AVES) 86
3.17 SSQ Subscales 87
3.18 MTE Flight Performance in AVES of Pilot B 89
3.19 All Tested Configurations ©DLR. CC BY-NC-ND 93
3.20 Linear Regression 96
3.21 Piece-wise Linear Regression 97
3.22 PQ and Mapped Mission-Task-Elements (MTE) Scores for Hover Maneuver 99
3.23 PQ and Mapped Mission-Task-Elements (MTE) Scores for Slalom Maneuver 100
3.24 2nd Degree Polynomial 101
3.25 ITC/OTW Ratios for Hover Task 102
3.26 ITC/OTW Ratios for Slalom Task 103
3.27 Mixed Reality (MR) Configurations 106

4.1 Aerial Work Task ©DLR. CC BY-NC-ND 119
4.2 Ground Work Task ©DLR. CC BY-NC-ND 120
4.3 Mission-Task-Elements (MTE) Lateral Repositioning ©DLR. CC BY-NC-ND . 120
4.4 Configurations of Study 3 122
4.5 SSQ Scores for Aerial Work (AW) and Ground Work (GW) 125
4.6 ITC/OTW Ratio for Trainingflights With Tobii Pro Glasses 3 127

5.1 ARiTra Simulator 142
5.2 ARiTra Simulator Cockpit with Green Screen ©DLR. All rights reserved. . . 143
5.3 Updated Ground Work Task ©DLR. CC BY-NC-ND 143

6.1 Multi-Crew-Cooperation (MCC) Application with Mixed Reality (MR) ©DLR. All rights reserved. 148

List of Tables

3.1 Overview of Visual Systems. 46
3.2 Overview of Tested Configurations. 47
3.3 Mission-Task-Elements (MTE) Hover Performance Parameters based on [63]. 49
3.4 Mission-Task-Elements (MTE) Slalom Performance Parameters based on [63]. 49
3.5 Assessment Methods 51
3.6 Pilot Experience. 59
3.7 Flight Performance of Slalom Mission-Task-Elements (MTE) in 2PASD . . . 66
3.8 Flight Performance of Hover Mission-Task-Elements (MTE) in 2PASD 68
3.9 Overview of Visual Systems 75
3.10 Overview of Tested Configurations 76
3.11 Flight Performance of Hover in AVES 88
3.12 Flight Performance of Slalom in AVES 90
3.13 Mission-Task-Elements (MTE) Score and PQ Score Mapping 95
3.14 Comparison of Correlation Coefficients Between Mapped MTE Scores and Presence Questionnaire (PQ) Scores for Different Regression Approaches . 97

4.1 Test Matrix of Study 3 122

A.2 Simulation Sickness Questionnaire and the calculation of nausea (N), oculomotor disturbance (O) and disorientation (D) based on [71] and for calculation clarification [92]. 153

Acronyms

Acronym	Definition
2D	Two-Dimensional
2PASD	Dual Pilot Active Sidestick Demonstrator
3D	Three-Dimensional
AATD	Advanced Aviation Training Device
ACAH	Attitude Command Attitude Hold
ACT/FHS	Active Control Technology/Flying Helicopter Simulator
ADS-33	Aeronautical Design Standard
AOI	Area of Interest
API	Application Programming Interface
AR	Augmented Reality
ATD	Aviation Training Device
AV	Augmented Virtuality
AVES	Air Vehicle Simulator
BATD	Basic Aviation Training Device
CAD	Caution and Advisory Display
CF	Control Factors
CFR	Code of Federal Regulations
CLS	Control Loading System
CON	Conventional Setup
CPL(H)	Commercial Pilot License (Helicopter)
CS-FSTD(H)	Certification Specifications for Helicopter Flight Simulation Traning Devices
DF	Distraction Factors
DLR	German Aerospace Center
EASA	European Aviation Safety Agency
EC	Experiment Computer
FAA	Federal Aviation Administration
FCS	Flight Control System
FDD	Focus Distance Distribution
FFS	Full Flight Simulator
FLI	First Limit Indicator
FNPT	Flight and Navigation Procedures Trainer
FOV	Field of View
FTD	Flight Training Devices

continued on next page

GC	Graphics Computer
HEMS	Helicopter Emergency Medical Services
HMD	Head Mounted Display
IC	Interface Computer
ICAO	International Civil Aviation Organization
IFR	Instrument Flight Rules
IG	Image Generator
ITC	In-The-Cockpit
ITC/OTW ratio	In-The-Cockpit/Out-The-Window Ratio
LBA	German Federal Aviation Office
LCD	Liquid Crystal Display
LiDAR	Light Detection and Ranging
MCC	Multi-Crew-Cooperation
MR	Mixed Reality
MRC	Mixed Reality Computer
MTE	Mission-Task-Elements
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NATO	North Atlantic Treaty Organization
NLR	Netherlands Aerospace Centre
NRC	National Research Council of Canada
OEI	One Engine Inoperative
OLP	Optical Line Pair
OTW	Out-The-Window
PCCR	Pupil-Center Corneal Reflection
PFD	Primary Flight Display
PIC	Pilot in Command
PPL(H)	Private Pilot License (Helicopter)
PQ	Presence Questionnaire
r	Correlation Coefficient
RC	Rate of Climb
RF	Realism Factors
RGB	red, green, blue
RMIT	Royal Melbourne Institute of Technology
ROS	Robot Operating System
SAR	Search and Resecue
SAS	Stability Augmentation System
SCM	System Control Module
SDK	Software Development Kit
SF	Sensory Factors

continued on next page

SSQ	Simulation Sickness Questionnaire
SUS	Slater Usaoh and Steed
TCP	Transport Control Protocol
TS	Total Score
TU Delft	Delft University of Technology
UDP	User Datagram Protocol
UMD	University of Maryland
VE	Virtual Environment
VEMD	Vehicle and Engine Multifunction Display
VFR	Visual Flight Rules
VMC	Visual Meteorological Conditions
VR	Virtual Reality
VTOL	Vertical Take-off and Landing

1 Introduction

Flying has always been a dream of mankind. The feeling of freedom, experiencing the world from above or even very close up when flying through canyons, a flight simulator allows you to experience all of this on the ground in a safe environment.

History of Aircraft Simulation

It all began with the first pioneering simulator device that was developed and built by Edward Link in 1929, called the LINK trainer. The simulator, that was also referred to as the "Blue Box" (Figure 1.1) allowed pilots to learn how to fly on the ground, which was cost-effective and risk-free. The LINK Trainer could move around in three dimensions and had all the important instruments for navigation and control. An opaque canopy made it possible to simulate night and instrument flying. Instructors and students communicated with each other through headsets, enabling instructors and students to interact just like they would in an actual aircraft. Initially, this simulator was also used as a coin-operated fun ride in amusement parks, but Edward Link was convinced that pilots should acquire flying skills with this simulator and opened a flying school in the basement of his father's piano and organ company. [1]
The turning point and thus the start of the simulation industry came in 1934 when the US Post Service asked the US Army to carry their air mail. As the pilots were not experienced with night and instrument flight, five planes crashed within a couple of days. To prevent further accidents, the US Army purchased the first LINK trainers in an attempt of preventing further accidents by on-ground pilot training. In total, Edward Link sold over 10.000 LINK trainers around the world. The success of the Link instrument flight trainer led to the development of further training devices.[1]

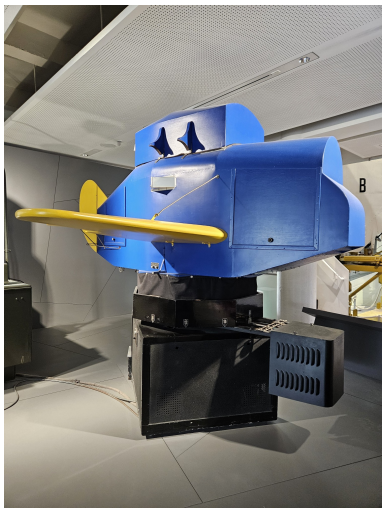
As aircraft complexity increased, pilots required more comprehensive training to operate them safely. To achieve this, replica mock-ups of specific aircraft types were necessary, allowing to simulate realistic scenarios. Moreover, the simulator should accurately replicate aircraft dynamics to effectively model the interactions between the vehicle and its environment. Recent advancements in computing capabilities have enabled these simulations to undergo substantial enhancements across various aspects, including improved aircraft modeling, motion platform simulation and enhanced graphical rendering. [1]

The development of visual out-the-window (OTW) scenes for flight trainers has a long history dating back to the beginnings of such training devices. An accurate and realistic OTW view is important for training simulation, as it usually is the main cue used by pilots

[1]This paragraph is based on [1]



(a) LINK Trainer [2]



(b) LINK Trainer at Deutsches Museum ©Tanja Dikarew-Martini. All rights reserved.

Figure 1.1: LINK Trainer

when operating an aircraft. Several approaches to generating visual scenes have been explored over the years. In the 1950s, the point-light source projection or shadowgraph method was particularly popular for helicopter simulators, where low ground velocity required a more precise simulation of surroundings. It is an optical technique used to create a visual representation of a scene. The light source is filtered through a series of screens or transparencies, which create a silhouette of the surrounding environment (e.g., buildings, trees, terrain). The filtered light is then projected onto a screen inside the simulator, creating a visual representation of the outside scene. [2]
However, this technique had limitations, such as simplified representation of the environment and lack of depth perception. Although the shadowgraph-based visual system was promising, its restrictive characteristics hindered its application and overall success [3].

As technology progressed, advanced visual systems like dome visual projection superseded point-light source projection in flight simulators. While this upgrade provided more realistic and engaging visuals, it also introduced the challenge of larger and more expensive simulator hardware (Figure 1.2a). However, the emergence of Virtual Reality (VR) and Mixed Reality (MR) headsets opened up new avenues for pilot training devices, offering flexible and mobile solutions for immersive simulation experiences (Figure 1.2b).

[2] This paragraph is based on [1]



(a) Air Vehicle Simulator (AVES) at DLR
©DLR. CC BY-NC-ND 4.0



(b) Dual Pilot Active Sidestick Demonstrator (2PASD) With MR Setup at DLR
©DLR. All rights reserved.

Figure 1.2: Flight Simulators at German Aerospace Center (DLR)

The Relevance of Helicopter Training

The phrase "Übung macht den Meister" (practice makes perfect) is a well-known adage in the context of acquiring skills and learning. In the field of aviation, this principle is particularly relevant when it comes to training helicopter pilots. To ensure safe and effective flight operations, helicopter pilots require extensive training due to the demanding nature of their missions. Because of their Vertical Take-off and Landing (VTOL) capability, helicopters can operate in confined or remote areas without the need for conventional landing infrastructure. As a result, they are often used in critical missions such as Helicopter Emergency Medical Services (HEMS), Search and Resecue (SAR), firefighting, aerial construction, offshore operations and many more.
Effective training is needed for mitigating the risk of accidents and ensuring the safe execution of flight operations. Helicopter pilot training is composed of both theoretical and practical components. Theoretical aspects cover topics such as aviation law, navigation and meteorology. These learning outcomes can be acquired through either joint instruction or self-study. In the practical training phase, standard maneuvers are trained alongside emergency procedures. Currently, most practical training takes place on real helicopters, but according to the European Aviation Safety Agency (EASA) Rotorcraft Safety Roadmap [4], simulators have proven to be an effective, safe and affordable means of training pilots. Within the roadmap, EASA encourages the use of new training devices, such as VR and MR, which can reduce the amount of hardware required for effective training, making simulators smaller, cost-effective and more portable.
The utilization of qualified flight simulators for training purposes enables flight training in a safe training environment and even the recognition of simulated hours as equivalent to actual flight time.

Traditional flight simulators, e.g. Full Flight Simulator (FFS), are often expensive and require specialized simulator personnel and flight instructors to operate. In contrast, MR technologies are gaining interest from the aviation industry as training tools. Implementing MR in simulation has the potential to reduce both time and cost by substituting some real-aircraft flight hours with relatively affordable simulated training [5].

The projection system with dome technology is a primary factor contributing to the overall dimensions of FFS. Replacing this visual setup with MR headsets would significantly reduce the simulator's size, thereby enhancing its portability and practicality. MR headsets enable the blending of virtual and real-world elements depending on the training requirements for specific training capabilities. This enables the use of real cockpits with their control units, switches and displays while simultaneously eliminating the need for a projection dome to simulate external views.
In addition, MR offers the ability to look anywhere, especially downwards, an option that is often restricted by the projection dome used in FFS. In helicopter simulations, especially for flight tasks requiring visual references on the ground, such as take-off, landing or hovering close to the ground, this limitation in FFS can affect altitude and speed estimation in the simulator, as noted by pilots in interviews.

Furthermore, with the increase of single-pilot aircraft, e.g. VTOLs, future pilot students will be trained with aircraft that provide only one set of controls, inhibiting the instructor to intervene in emergency situations. In this scenario, MR technologies could be beneficial not only because they provide an immersive simulation environment, which is necessary for effective training of pilots prior to flying the aircraft, but also because of their adaptability for further development of new aircraft systems.
Although it offers several advantages, MR also presents challenges and limitations. Currently, both the EASA and the Federal Aviation Administration (FAA) are working on the establishment of certification standards for VR training simulators. However, the effects of using MR in helicopter simulation training have not been thoroughly investigated (Section 1.2.1).

1.1 State of the Art

This chapter provides an overview of the state of the art. The chapter begins with an explanation of the existing standards for the qualification of helicopter simulators, followed by an overview of the different simulator types. Subsequently, the current path to obtaining a helicopter pilot license is described. Furthermore, the terms VR, MR and Augmented Reality (AR) are defined, the different display technologies used and their impact on user experience are detailed. Finally, the current state of research in the field is presented, which provides a comprehensive basis for the subsequent discussion on the integration of VR, MR and AR in helicopter simulation.

1.1.1 Current Helicopter Simulators

In addition to regular helicopter flight training, as already mentioned, simulators have already been proven to be effective and safe for various aspects of pilot training and therefore the intention to use simulations for training is also part of the EASA Rotorcraft Safety Roadmap [4], which calls for the development of new types of training devices. The current requirements for helicopter training devices in Europe are summarized in the Certification Specifications for Helicopter Flight Simulation Traning Devices (CS-FSTD(H)) [6] by EASA.

For helicopter flight simulation training, the CS-FSTD(H) defines three main categories of devices: Full Flight Simulator, Flight Training Devices and Flight and Navigation Procedures Trainer:

1. A **Full Flight Simulator (FFS)** is a highly realistic replica of a specific helicopter type's cockpit, including all the necessary equipment and computer programs to simulate ground and flight operations. It features a visual system that provides an outside view from the cockpit, as well as a force-cueing motion system that simulates the feel of the aircraft in flight. FFSs are further categorized depending on their capability as level A, B, C or D, where D is the highest qualification.
2. A **Flight Training Devices (FTD)**, on the other hand, is also a replica of a specific helicopter type's instruments, equipment and controls, but it does not require a visual or force-cueing motion system. It simulates ground and flight conditions to the extent that the systems installed in the device appear to function as they would in an actual aircraft. Depending on their capability, the FTDs are further categorized as level 1, 2 or 3, with level 3 devices providing the highest capabilities.
3. Lastly, there is the **Flight and Navigation Procedures Trainer (FNPT)**, which represents the helicopter cockpit environment, including all the necessary equipment and computer programs to simulate flight operations. While it does not provide an equal level of realism as a FFS or FTD, it is still an effective tool for training pilots in

various procedures and operations. FNPTs are further categorized as level I, II and III, depending on their capabilities.

The requirements specified by the CS-FSTD(H) differ across the three device categories according to the Easy Acess Rules [7]. For example, regarding the minimum Field of View (FOV) of the visual system of a FFS level D, a continuous cross-cockpit with a minimum visual FOV of 180° horizontal and 60° vertical is required for each pilot. Alternatively, for FFS level A, a FOV of 45° horizontal and 30° vertical simultaneously for each pilot is sufficient.
For FTD level 1 and FNPT level I, the minimum visual FOV is not explicitly listed in specifications. However, for FTD level 3 and FNPT level III, the requirement is a continuous, cross-cockpit, minimum visual FOV with 150° horizontal and 60° vertical for each pilot.

The qualification standards for simulation devices for the United States are published in 14 Code of Federal Regulations (CFR) Part 60 [8] which are comparable to the EASA CS-FSTD(H). The specification defines the categories FFS and FTD similiar to the EASA. But instead of the FNPT, the category Aviation Training Device (ATD) is specified. The ATD includes two primary types of entry-level training tools: the Basic Aviation Training Device (BATD) and the Advanced Aviation Training Device (AATD). These devices are predominantly utilized for initial training, private and commercial pilot training programs. They are characterized as less sophisticated, fixed-base systems and are qualified in accordance with FAA Advisory Circular (AC) 61-136B [9].

Currently, the established requirements are changing due to the growing technical possibilities and the increasing demands on helicopter operations and qualified pilots. In 2021 and 2022, two helicopter training devices (Robinson R22 Beta II and Airbus Helicopters H125) using VR headsets have been qualified by EASA (as FNPT II and FTD level 3, respectively). The additional requirements for VR technology-based helicopter cockpits are published in the "CS-FSTD(H) Special Conditions for the use of Head Mounted Display (HMD) combined with a motion platform with reduced envelope" [10]. Besides EASA activities, the FAA qualified the VR training simulator by Loft Dynamics in 2024. In 2025, a "Virtual Extended Reality (VxR)" training system from Leonardo Helicopters with a Varjo Headset has been qualified by EASA as FTD level 3 and by FAA as FTD level 7. The different qualification levels can be attributed to the similar but different specification standards of EASA and FAA. In 2025, the first flight simulator for the Diamond DA42 aircraft type from Brunner Elektronik was also qualified by EASA. This is the first MR simulator for an aircraft in which pilots can see themselves and the original cockpit in the monitor-based MR headset.

Furthermore, the FAA has recently published rules [11] to integrate future powered-lift aircraft into the National Airspace System (NAS), including guidelines for certifying pilots and clarifying operating rules. As part of this effort, the FAA is adopting new frameworks for training and certifying single-pilot crews with simplified flight controls. This update also opens up opportunities for dedicated simulation training for single-pilot aircraft.

Finally, with the "FSTD Certification Guidance Document" [12] the International Civil Aviation Organization (ICAO) has developed a guide to harmonize qualification standards for flight simulators based on existing standards such as for example EASA CS-FSTD(H) and FAA CFR Part 60. The goal is to reduce costs and resources for pilots and aviation authorities while increasing safety through the use of simulation and modern training technology. The document provides guidelines for qualifying flight simulators, including setting requirements for performance and documentation. It aims to help aviation authorities carry out a qualified evaluation and issue qualification for a flight simulator. Additionally, the guide should assist other countries' aviation authorities in recognizing the qualifications of flight simulators without needing further evaluations.

These efforts show that the authorities recognize the necessity for updating and adapting their current regulations, thereby creating an opportunity to explore innovative aviation training technologies such as MR.

1.1.2 Current Helicopter Pilot Training

This section describes the path to become a helicopter pilot. The requirements of EASA and FAA are described to get an overview of the current helicopter training.
There are several ways to become a helicopter pilot. One option is to join the military, where each country typically offers its own training programs. On the other hand, if an individual chooses to pursue their pilot's license privately, they must adhere to national aviation standards, such as those set by the FAA in the United States or the EASA in Europe.

In the European Union, the first milestone is to receive the Private Pilot License (Helicopter) (PPL(H)). This, according to EASA standards in Easy Access Rules for Aircrew [13], requires flight training with at least 45 hours where 25 hours are of dual flight instruction and 10 hours of supervised solo flight, including five hours of solo cross country flight. Five hours may be completed in a FNPT or FFS.
After receiving the license, pilots need to collect 155 flight hours before they are eligible for the Commercial Pilot License (Helicopter) (CPL(H)). Of those 155 flying hours, the pilot needs to be in the role of Pilot in Command (PIC) for 50 hours, which means that for these hours the pilot bears responsibility for the aircraft during operation. If the pilot did not obtain an instrument rating which allows to fly under Instrument Flight Rules (IFR) beforehand, 30 hours should be completed in dual flight instruction for training IFR, from which up to ten hours may be conducted in a simulator. The instrument rating enables pilots to operate safely in bad weather conditions. Without this rating, pilots are restricted to Visual Flight Rules (VFR), which limit them to fly only in Visual Meteorological Conditions (VMC). In addition to the licenses PPL(H) and CPL(H), different kinds of ratings may be granted. These include, for example, type ratings for specific helicopter types and a rating for Multi-Crew-Cooperation (MCC). MCC training aims to prepare pilots for multi-crew

operations by developing technical and non-technical skills to ensure safe and efficient flight operations. For each rating different rules apply for using simulators or taking actual flight hours.
In addition, each type rating needs to be renewed by proficiency checks with an authorized examiner. As the EASA has currently made significant changes to its Easy Access Rules for Aircrew, these proficiency checks can already be performed, for example, on the H125 VR Training Device of Loft Dynamics [14]. This notable revision removes the requirement for helicopter pilots to undergo training, skill tests and proficiency checks using FFS, now allowing them to use other types of CS-FSTD(H).

In the United States, according to FAA 14 CFR PART 61 [15], a person applying for a private pilot certificate in airplanes, helicopters and gyro-planes must log at least 40 hours of flight time, which includes at least 20 hours of flight training with an authorized instructor and 10 hours of solo flight training. They must also have a minimum of three hours of cross country flight experience, three hours of nighttime flight experience and three hours of instrument time to be eligible for certification. With a PPL(H), the pilot is eligible for an instrument rating course which allows to fly under IFR. To obtain the CPL(H), a person must log at least 150 hours of flight time as a pilot. This includes at least 100 hours in powered aircraft, with 50 hours specifically in helicopters. Additionally, they must complete 100 hours of PIC flight time, which includes at least 35 hours in helicopters and 10 hours in cross-country flights in helicopters. The person must also undergo 20 hours of training on specific areas of operation, such as flying solely by reference to instruments using a view-limiting device, cross-country flights in daytime and nighttime conditions and preparation for the practical test with an authorized instructor. This training must include at least five hours on instrument flying, two hour cross-country flights in helicopters (one in daytime and one in nighttime) and three hours with an instructor. Finally, they must complete 10 hours of solo flight time in a helicopter or perform PIC duties for 10 hours with an authorized instructor on board. This training should include a cross-country flight with landings at three points, flying in night VFR conditions and takeoffs and landings in the traffic pattern.

Concerning usage of simulators for training, the CFR PART 61 states that, for example, pilots can use flight simulators or training devices to gain instrument experience for their certification, as long as they have an authorized instructor observing and verifying the training. This simulated time can also be used to keep their instrument skills up to date. To log this training, pilots will need to document it in their logbook or training record with details about the device, how long they spent training and what was covered during the session.

The current state of helicopter training is characterized by a focus on actual helicopter flights and less on simulator training, but this status quo is likely to shift due to new technologies for simulators like VR and MR.

1.1.3 Virtual (VR), Mixed (MR) and Augmented Realities (AR) in Research

In this chapter, firstly, the definitions of VR, MR, AR and Augmented Virtuality (AV) used in this thesis are explained. This is followed by a brief overview of the most common display technologies. Additionally, implications of MR technology identified by related research works are briefly summarized. The chapter is concluded by a research overview of the application of VR and MR in aircraft simulation and operation.

Definition of VR, MR, AR and AV

Different terminology is employed to distinguish various realities in both scientific and business domains. Milgram et al.'s [16] reality-virtuality continuum is widely accepted to define and differentiate between various realities. It ranges from the real environment on one end to the Virtual Environment (VE) / VR on the other end (Figure 1.3). According to Milgram et al., **VR** is characterized by a state where an individual is fully immersed within a synthetic environment, while the physical/real world is governed by the laws of physics. **MR** represents a continuum of characteristics where the boundaries between the real and virtual worlds may be blended in various degrees. Therefore, MR can be further divided into AR and AV. In literature, **AR** seems to be related more to see-through displays (Section 1.1.3). Nevertheless, without the relation to a display technology, it is often described as the extension of the real world with virtual elements. In contrast, **AV** is defined as a computer-generated world that additionally displays real elements.

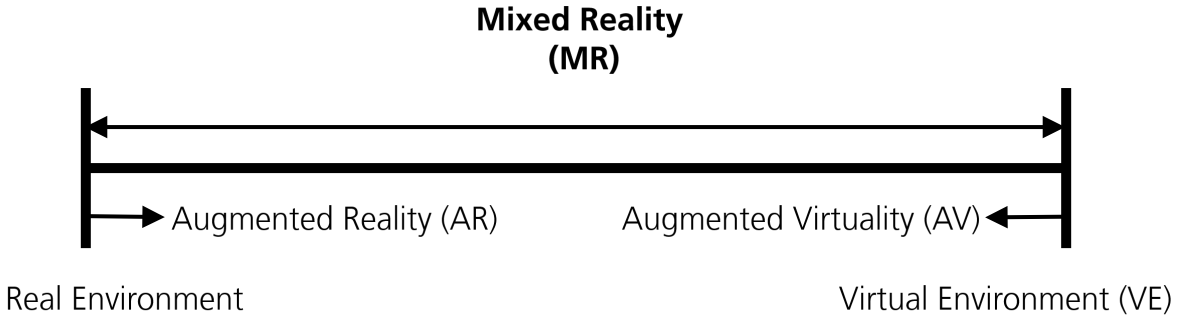


Figure 1.3: Reality-Virtuality (RV) Continuum by Milgram et al. Adapted From [17], ©1994 IEICE.

Headset Types

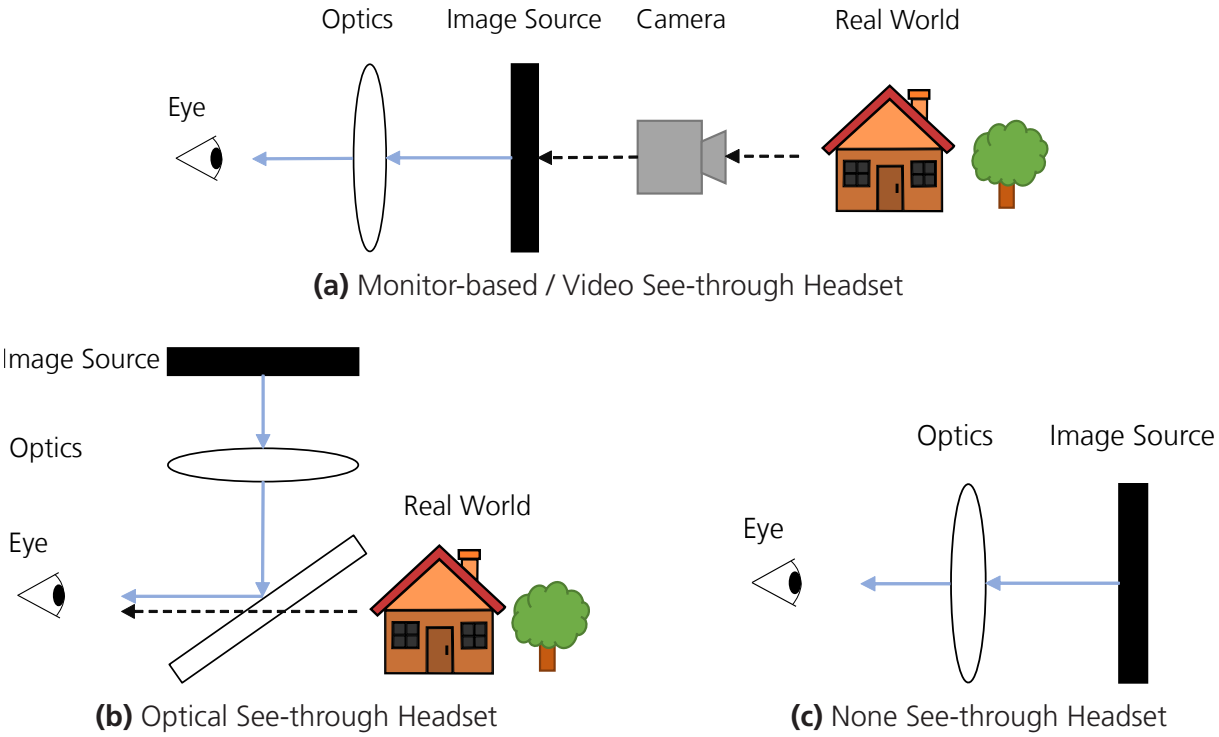
MR and VR can be generated by three distinct headset technologies [18, 19]:
see-through displays, monitor-based displays and none see-through displays.

Monitor-based headsets or video see-through headsets feature cameras attached to the front, capturing a stereo video stream of the real environment. This stream is displayed to the user alongside virtual content, creating a blended image (Figure 1.4a). The MR headset provides a realistic visual experience by allowing for controlled ambient lighting and the ability to determine which objects should be virtual or real, offering a realistic blended image. However, the concept presents specific challenges, primarily requiring high camera and display resolution with minimal latency to ensure a seamless blended image. Additionally, adapting to varying lighting conditions is essential, as these conditions can significantly impact the quality and accuracy of the blended image. [3]

Optical see-through headsets, which are often referred to as AR, contain a optical combiner which reflects the virtual image from the display into the user's eyes and allows them to see through the optics to the real environment (Figure 1.4b). As a result, the user sees the VR elements as an overlay on the real world.
One of the advantages of see-through headsets is the direct view of the real environment, which enhances the feeling of presence in the real world. Nevertheless, there are also major limitations with these devices. Firstly, the virtual elements can obscure the real objects and the colors may be altered due to blending. Secondly, these devices display the virtual elements as overlays using additional light, which hinders the ability to selectively dim the environment, resulting in limited control of contrast. [3]

None see-through headsets, also known as VR headsets, operate by displaying a virtual scene through a simple optical system consisting of a collimating lens placed in front of the image source (Figure 1.4c). This setup creates an immersive environment for the user, where the entire real world is blocked from view. As it is isolated from the real world, a VR headset can provide saturated colors and high contrast independently of surrounding conditions. However, the isolation can also lead to the feeling of being disconnected from the real world. [3]

[3] This paragraph is based on [18, 19]



(a) Monitor-based / Video See-through Headset

(b) Optical See-through Headset

(c) None See-through Headset

Figure 1.4: Headset Types and Functionality Adapted From [20]

Impact of Using See-through and Monitor-based Headset

One of the primary concerns associated with the use of VR headsets is the phenomenon of simulation sickness.

The most widely supported explanation for simulation sickness is the **sensory conflict theory** [21]. This theory states that simulation sickness occurs due to a discrepancy between visual and vestibular signals, which disrupts the brain's ability to accurately interpret spatial orientation.

To address this, a North Atlantic Treaty Organization (NATO) specialist team under HFM-MSG-323 (Human Factors and Medicine Modeling and Simulation Group) developed Guidelines for Mitigating Cybersickness in VR Systems [22]. Their goal was to identify best practices, design solutions and countermeasures based on scientific literature, to reduce cybersickness/simulation sickness in VR applications.

Based on the Guidelines for Mitigating Cybersickness in VR Systems, the most important factors that can cause simulation sickness are briefly outlined below, as they are important for the upcoming studies.

Display Related Factors [4]

- **Latency (Time Lag):** The delay between user input (e.g., head movement) and visual output, consisting of sensor, processing and display delays. The aim is to minimize this delay to prevent vestibular-ocular mismatches, which can cause discomfort or nausea.

- **Display Resolution:** Determined by the number of pixels, it directly affects visual sharpness, detail perception and immersion by reducing pixelation and enhancing spatial accuracy in a VE.

- **Screen Flicker:** A visual artifact is caused by inconsistent frame rates or display refresh rates. It can induce eyestrain and simulation sickness by disrupting stable visual feedback required for natural motion perception.

- **Refresh Rate:** Each rendered visual output in a simulation is called a frame, processed and displayed in discrete time intervals. For VR applications, a sustained frame rate of at least 90 frames per second is recommended.

- **Motion Blur:** Artificial motion blur can reduce perceived latency but may compromise depth perception, indirectly affecting sickness. Excessive blur (from slow refresh rates) worsens visual degradation, increasing cognitive load.

Visual Scene Factors [4]

- **Visual Scene Density and Altitude Above Terrain:** Low-altitude navigation (e.g., flight, driving) produces high-density visual cues, increasing simulation sickness risk compared to high-altitude scenes.

- **Embodiment:** The illusion of controlling a virtual body can reduce discomfort if real and virtual bodies align spatially/temporally. However, its relationship with simulation sickness remains unclear.

- **Predictability:** Unexpected movements (e.g., roller coaster twists) worsen sickness. Visual cues (e.g., arrows, trails) help users anticipate changes, mitigating discomfort.

- **FOV**: VR's narrower FOV compared to reality creates unnatural delays in motion perception, increasing simulation sickness. Amplifying head/hand movements to compensate for FOV discrepancies (e.g., exaggerating motions to match wider perceived fields) can reduce symptoms by aligning vestibular and visual inputs.

- **Velocity/Direction Changes:** Frequent shifts in visual flow increase sickness but may reduce vection (the illusion of self-motion).

[4]These factors are based on the NATO Guidelines [22]

- **Visual Realism:** While high visual realism in VR improves immersion and training effectiveness, it paradoxically increases the risk of simulation sickness due to increased perceptual demands and vestibular-visual conflict, where the brain expects natural movement cues (e.g. balance cues) that VR cannot provide, exacerbating sensory mismatch.

Further Factors [5]

- **Head Movements:** There is a strong link between head movements and simulator sickness. Minimizing abrupt changes in velocity/direction and ensuring low-latency tracking is therefore important.
- **Eye Fixation:** Fixating the eyes on a stable target (like a head-tracked object) often reduces simulation sickness by cutting visual motion cues, but its effectiveness varies depending on how it interacts with head movements and visual-vestibular conflict.
- **Duration:** Numerous studies with simulators and virtual reality head-mounted displays have shown that prolonged exposure to VE exacerbates simulation sickness.
- **Head Tracking:** Precise head tracking is essential for aligning virtual and physical movements. However, errors in motion capture, such as those caused by obstructions, interference or calibration issues, can disrupt spatial coherence and cause simulation sickness.

Further research on implications of VR headsets besides the guidelines are listed in the following:

- **Distortion in Visual Feedback:** Distortion warps geometry, causing straight edges to appear curved and compressing or expanding distances, which leads to inaccurate depth perception and size estimation. The study in [23] highlights that distortion in visual feedback, particularly due to factors like depth displacement, restricted FOV, weight of the MR headset and lack of depth cues, has a substantial negative impact on hand-eye coordination and performance in tasks performed with a video see-through headset.
- **Distance Estimation Errors:** This error in a see-through headset is the systematic mismatch between the actual physical distance of real-world objects and the distance that users perceive through the device's camera-based view. The factors contributing to distance estimation errors in virtual reality are e.g. the headset's FOV and weight [24]. Limited FOV can lead to overestimation or underestimation of distance based on the inclusion or exclusion of cues like texture information. A heavy headset tilts the wearer's head slightly, causing the brain to interpret virtual objects as being positioned lower in the view and therefore to judge them as closer than they actually are.

[5] These factors are based on the NATO Guidelines [22]

- **Visual Adaption:** Humans can adapt to visual displacement to some extent. However, if the amount of visual displacement is increased, the task performance improvement due to training decelerates significantly. This means that, while humans can adapt to some level of visual disturbance, further improvement through training gets increasingly difficult as the distortion increases [25].
- **Weight:** The weight of a MR headset on the head correlates positively with the pressure load on the head and the discomfort reported by users [26]. By minimizing the overall weight and optimizing the support architecture, the resulting forces are distributed more evenly on the head, reducing the perceived strain. Consequently, weight reduction and ergonomic support design are the most effective strategies for improving user comfort in VR systems.
- **Presence:** The feeling of presence in a VE is also affected by factors such as display resolution, the latency and display frame rate [27]. These factors show that they not only contribute to simulation sickness, but also significantly influence the perceived sense of presence in VE.

Research of VR and MR in Aircraft Simulation and Operation

VR and MR have gained significant attention for their potential applications in aircraft simulation and operation. In aviation, research has already been conducted on the use of VR and MR headsets in flight training. A scoping review conducted by [5] suggests that MR has significant potential for integration into flight training programs, with several contributing factors identified, including time and monetary savings, as well as enhanced training capabilities. However, limitations have also been uncovered in the research, which warrant further investigation and refinement.

Research indicates that individuals who undergo VR training demonstrate comparable or even superior performance compared to those trained through traditional methods [28–30]. VR flight training has also shown potential for decreasing training duration and reducing costs [31, 32]. However, VR flight training also has limitations, particularly concerning simulation sickness (Section 1.1.3) and breaks in presence, which can result in the participant feeling disconnected from the physical world. This can lead to restrictions in operation, especially when pilots need to interact with physical cockpit controls, such as buttons or switches [33]. A study compared a VR flight simulator with a physical flight simulator for cockpit familiarization [34]. The results showed that participants were able to successfully complete check procedures in both simulators, however, some individuals reported symptoms of simulation sickness after using the VR flight simulator.

National Aeronautics and Space Administration (NASA) and Systems Technology, Inc. conducted a flight test using an MR headset to simulate an air-to-air refueling maneuver [35]. The pilot commanded a Learjet 3 aircraft while wearing the MR headset, which displayed

a simulated tanker aircraft with a probe within the visual field of the pilot. The primary objective was to reduce risk due to very low distances between the two aircraft during the maneuver, by replacing the actual tanker aircraft with a simulated aircraft. Based on the findings from previous in-flight experiments, further flight tests were conducted using the Fused Reality Flight System in 2015 [36]. Following a familiarization phase, the pilots were tasked with performing specific maneuvers, such as precision landings. Observations showed that visual cueing issues arose due to the limited FOV presented by the MR system, which had a significant impact on the pilots' ability to accurately perform tasks.
Additionally, the effects of FOV restrictions on the head movement of rotorcraft pilots have been examined [37]. The findings show that as the FOV decreases, the pilot's head movement adapts to compensate for the reduced visual information. When faced with limited FOV, the pilot must move his head to see visual cues that are outside the visible area, whereas a larger FOV allows the pilot to view visual cues in their peripheral vision.
The National Research Council of Canada (NRC) conducted flight trials where a pilot wore a VR headset to simulate helicopter operations, including ship landing tasks [38]. Several rating scales were used in the study to assess the pilots' subjective experiences of workload during the simulation-based tasks. The findings show that the prediction of a specific rating scale can be improved by combining physical workload indicators such as pilot control movements and heart rate with ratings scales, rather than relying on a single strategy.

Research continues with various aircraft simulators. Besides the German Aerospace Center (DLR) simulators (Section 2.1), the University of Liverpool is also using the HELIFLIGHT simulator, which offers re-configurable flight simulation with dynamic control loading and a wide-view collimated visual display together with a virtual instrument panel [39]. The SIMONA research simulator of Delft University of Technology (TU Delft) is another example, capable of re-configuring for fixed-wing, rotary-wing or ground-vehicle platforms, with interchangeable force-feedback controls, a fully programmable touchscreen instrument panel, a wide-view collimated display generated by three projectors and six high-performance hydraulic actuators that provide realistic cabin-motion cues [40].

Interest in aircraft simulation remains high, driving the development of new VR simulators. For instance, the MoVR-X at the Royal Netherlands Aerospace Centre (NLR) features a modular design with a 6-DOF short-stroke motion platform that can be quickly linked to a variety of MR headsets, helicopter controls and flight models [41]. The University of Maryland (UMD) is using a VR motion-simulator based on Lockheed Martin's Prepar3D equipped with a VR plug-in and a compact motion platform that reproduces accurate flight dynamics for its research [42]. The Eight360 NOVA Simulator of Royal Melbourne Institute of Technology (RMIT) is another recent innovation, offering a two-metre-wide single-operator VR sphere that delivers full 360° motion, immersive audio-visual-haptic cues and integrated biometric data, thereby expanding the possibilities for aviation research [43].

In conclusion, VR and MR have shown potential for application in aircraft simulation and operation. However, limitations, such as simulation sickness, breaks in presence and limited FOV, need to be addressed. It is important to note that the current research has primarily focused on VR for fixed-wing aircraft and more studies are needed to investigate the use of VR and MR in rotary-wing aircraft, which have different requirements for training because of their VTOL capabilities and differing mission spectrum.

1.2 Contribution of This Dissertation

This dissertation seeks to make significant contributions to both current research and emerging trends in the industry by investigating the potential benefits of MR technology for helicopter training and examining the implications of this innovation.

1.2.1 Research Objectives

The existing literature lacks a comprehensive understanding of the potential benefits and limitations of MR technology in helicopter training. Specifically, there is a significant knowledge gap in several areas:

- **Evaluation of Different MR Displays Technologies for Helicopter Pilot Training:** This thesis assesses the effectiveness of different MR display technologies, such as monitor-based and see-through displays. The evaluation will provide valuable insights into the suitability of the different display technologies for MR-based helicopter simulation training.
- **Comparison of Pilot Performance Between MR and Conventional Simulators:** This thesis investigates the impact of MR technology on pilot performance and compares it to conventional simulators. This investigation is essential as it helps to understand the suitability of MR technology in helicopter simulation training and its potential to replace or supplement traditional training methods.
- **Investigation of the Concept of Presence for Helicopter Simulations:** This thesis explores the role of presence in assessing the suitability of MR-based helicopter simulations. Presence is a relevant aspect of MR technology, as it enables pilots to immerse themselves in a VE.
- **Creation of a Scientific Knowledge Base for MR Simulators**: The aim of this thesis is to provide a comprehensive understanding of the key considerations for MR simulators. This knowledge should ensure that MR simulators meet the needs of pilots and training organizations and should also serve as a basis for the qualification process of simulators.

The thesis contributes to the development of novel helicopter flight simulators and provides guidelines for MR simulators based on scientific research.

1.2.2 Research Questions

As outlined in the previous sections, the motivation for this work is to learn about the impact of using MR headsets during flight training.
Therefore, the overall aim is to investigate whether a MR headset is suitable for the simulation training of helicopter pilots. This leads to this dissertation's main research question:

Main Research Question (MRQ) 1: Can a MR headset serve as a suitable visual system for training helicopter pilots?

To address the MRQ 1 the following subordinate research questions were formulated.

Subordinate Research Question (1.1): What are the key considerations for the application of MR headsets into helicopter simulation?

The subordinate question RQ 1.1 aims to identify the key considerations for implementation as well as for the use case of MR in helicopter simulation. The implementation considerations should address technical characteristics of MR setups and should provide guidelines in designing and optimizing MR simulators. The use case considerations should address relevant factors regarding the application of MR in a simulator. Therefore the MR headset was integrated in two distinct simulation environments: a basic pilot station equipped with TV screens and a fully equipped EC135 cockpit with a dome-projection visual system. The lessons learned form the basis to answer RQ 1.1.

Subordinate Research Question (1.2): What are the potentials and limitations of using MR headsets for helicopter pilot simulation training compared to other visual systems (TV screens and dome projection system)?

To evaluate potentials and limitations, two studies were conducted.

Study 1 (basic pilot station): Three visual systems, two MR headset technologies (a monitor-based and a see-through headset) and a TV screen were compared to each other. Pilots performed hands-on maneuvers (hover and slalom) with each visual system. The evaluation focused on presence, simulation sickness and flight performance. Results from Study 1 were used to select the most promising MR headset for subsequent testing.

Study 2 (fully equipped EC135 cockpit): The MR headset selected in Study 1 was evaluated against the dome projection visual system. In addition to the hands-on maneuvers used in Study 1, the pilots executed a malfunction procedure to investigate cockpit interaction with an MR headset. Study 2 complements Study 1 and answers RQ1.2 by identifying MR's potentials and limitations relative to TV screens and dome projection systems.

Subordinate Research Question (1.3): How can new systems, such as MR headsets, be evaluated to assess the quality of a visual simulation?

Finally, the assessment methods and flight tasks used in the studies were re-examined to determine their applicability in evaluating visual systems. This analysis involved considering the applicability of the assessment methods to new visual systems and identifying a wider range of flight tasks that could be used to evaluate MR headsets in industrial settings. Two individual flight tasks combining multiple elements were defined for this purpose.

1.2.3 Outline of the Thesis

The thesis is structured into five chapters. In the current **Chapter 1**, the topic of MR in helicopter training simulation was introduced, including an overview of the state of the art of current helicopter training and simulators, the motivation for investigation of MR for helicopter pilot training in this thesis and the derived research questions.

In order to investigate the adaptability and of MR technology, a MR headset was integrated into two different simulators at German Aerospace Center (DLR), which are described in detail in **Chapter 2**. The chapter also outlines the simulation framework and explains the software and hardware adaptations implemented for MR integration. Additionally, a latency measurement was carried out to assess the delay between the pilot's control inputs and the display output of the VE in the MR headset.

For investigating the key considerations, potentials and limitations of MR in helicopter training simulation, the approach and the resulting chapter structure are shown in Figure 1.5 and are shortly described in the following:

Two studies have been conducted and are presented in **Chapter 3**. **Study 1** was conducted in the Dual Pilot Active Sidestick Demonstrator (2PASD), comparing two different headset technologies for MR and their possible configurations in the context of helicopter training. The results of this study provide insights into the impact of MR technologies by presenting subjective ratings from presence and simulation sickness questionnaires, as well as flight performance ratings assessed from recorded aircraft state data.
Subsequently, the steps required to integrate the MR headset into the fully equipped Air Vehicle Simulator (AVES) are described, followed by a discussion of the optimization of the MR setup based on the findings from the previous study.
Next, the chapter presents **Study 2** conducted in the AVES. In addition to the first study, the interaction with the cockpit is investigated. For comparability, the same assessment methods are used as in Study 1. Subsequently, a comparative analysis of all configurations employed in the two studies is presented, with a focus on questionnaire ratings, flight performance and eye tracking data.

In **Chapter 4** the evaluation of the applicability of the utilized assessment methods is discussed. In addition, in **Study 3**, pilots evaluate novel flight tasks designed to assess the suitability of MR technologies. The tasks were developed based on the findings of this thesis and shall be used in an upcoming evaluation of a MR helicopter simulator in an industrial setting.

In **Chapter 6**, a comprehensive summary of this thesis is provided, along with an answer to the main research question. The chapter also describes how the findings of the studies have been successfully integrated into a helicopter simulator in an industrial setting. Furthermore, the chapter presents an outlook to future areas of research based on the results of this work.

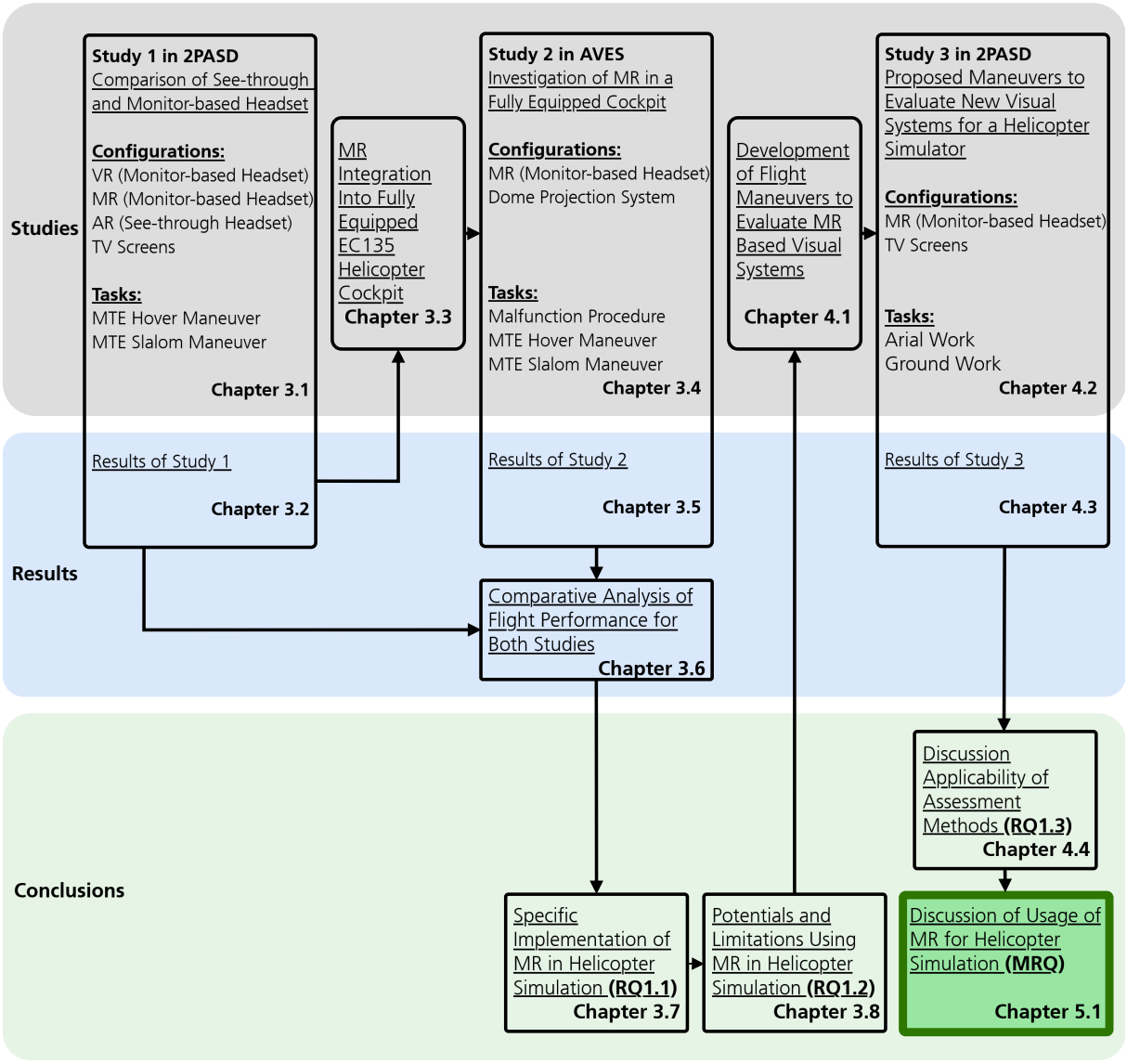


Figure 1.5: Overview of the Chapters 3 to 5

2 Reference System Architectures for Mixed Reality Helicopter Simulation

To assess the potential of MR technology, a MR headset was integrated into two simulators at the German Aerospace Center (DLR). This setup enabled the investigation of MR's adaptability in different simulation environments. Furthermore, this chapter provides a detailed description of the simulation framework and the software and hardware adaptations implemented for MR integration. To validate the quality of the MR setup, latency measurements were performed.

2.1 Simulation Environment

This section describes the helicopter simulation environment at the DLR that is used for the studies in this thesis. The two different simulators (Air Vehicle Simulator - AVES and Dual Pilot Active Sidestick Demonstrator - 2PASD) have been the fundamental platforms where a MR headset was integrated to investigate the opportunities, challenges and limitation of the MR technology.

2.1.1 Air Vehicle Simulator (AVES)

The AVES facility [44] at DLR provides multiple cockpits (Airbus 320, Dassault Falcon 2000, Dornier 328 and Airbus Helicopter EC135) and an advanced future aircraft cabin which are exchangeable between a fixed and a motion platform (Figure 2.1a). The focus of this work is placed on the EC135 Cockpit, which is a replica of the Active Control Technology/Flying Helicopter Simulator (ACT/FHS) research helicopter [45]. The cockpit cell features flight instruments such as a Primary Flight Display (PFD), a Vehicle and Engine Multifunction Display (VEMD) and a Caution and Advisory Display (CAD) to provide the pilots with all necessary information.
The pilot can maneuver the helicopter using the control inceptors of the Control Loading System (CLS). The control inceptors consist of a long pole stick in the center, the collective lever and the pedals (Figure 2.1b). To simulate the flight dynamics behavior of the ACT/FHS, DLR's non-linear helicopter model HeliWorX is used [46]. The helicopter model includes a set of modular components, including the fuselage, horizontal stabilizer, vertical stabilizer, main rotor and tail rotor. For modelling of dynamic inflow, the widely used model by Pitt and Peters [47] is implemented. In addition, the main rotor is modeled by fully articulated rotor blades, with an equivalent hinge offset and spring restraint to represent the basic



(a) AVES ©DLR. CC BY-NC-ND

(b) EC135 Cockpit ©DLR. CC BY-NC-ND

Figure 2.1: Air Vehicle Simulator AVES

flapping and lagging natural frequencies. Each main rotor blade is assumed to be a rigid blade. Aerodynamic forces and moments are calculated using the blade element theory. In addition to basic helicopter simulation, the EC135 simulator also includes a replication of ACT/FHS's flight experimental system, featuring identical hardware and software to replicate the behavior of the research aircraft during simulation. The experimental system provides simulated sensor data, data acquisition, data recording, telemetry and interface functions (e.g. user interfaces and data displays). Furthermore, the software includes an experimental Flight Control System (FCS) for research purposes, which allows to choose between different flight control laws and upper modes during simulation.

The conventional visual system of AVES features nine high resolution projectors providing a FOV of 240° × 90°, see Figure 2.2. The resolution on the projection dome, expressed in inch per Optical Line Pair (OLP), varies between 3.31'/OLP to 4.95'/OLP. Three of the projectors are mounted centrally at the front of the dome and are set in portrait (vertical) orientation. These projectors cover the upper region of the visual field. Two additional projectors are placed on the left- and right-hand sides of the dome to fill the peripheral areas on each side. The remaining four projectors are positioned in the middle of the dome and are tilted downward, projecting the lower portion of the image.
In order to maintain an accurate color representation and a uniformly bright image, a color and brightness calibration process is conducted regularly at certain intervals. First, the color spectra are measured from each individual projector in order to adjust the red, green, blue (RGB) channel gains and thus ensure accurate color reproduction. Secondly, as the projectors are positioned at different distances from the screen, the luminance of each projector is set individually so that the combined image appears uniformly bright across the entire screen.

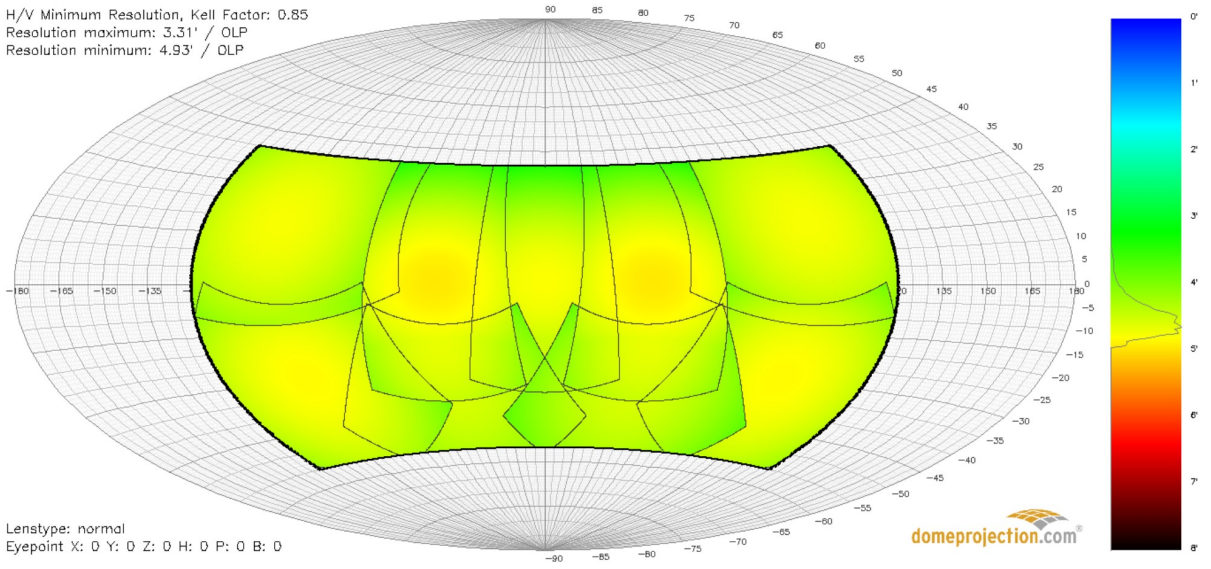


Figure 2.2: AVES Dome Projection Resolution
©project:syntropy GmbH/domeprojection.com

The blending is calibrated with a system that uses six calibration cameras and LED markers. The measurement data is used to create a mesh that determines the geometry of the individual projector's image, allowing the blending masks required for seamless image overlap to be created accordingly.
To create an image with a projection system, it is important to define a specific eye point from which the image is viewed from the cockpit. In a typical cockpit with two pilots, for example, this eye point would be at the level of the eyes of the PIC (right side of the cockpit). However, if the viewpoint is defined for the PIC, the co-pilot's view is affected by an angle error and is therefore distorted/not displayed correctly, which is a limitation of dome projection systems.
One solution to this problem could be to employ collimated optics in the display. Collimated optics emit light rays that are parallel, causing the image to appear at a distance. This arrangement reduces directional (parallax) error for both pilots simultaneously. However, collimated displays providing the required FOV, 240° horizontally and 90° vertically, were not available when AVES was built.
The DLR Institute of Flight Systems developed its own PC-based image generation system named AVES Image Generator (IG), which utilizes an OpenGL-based rendering engine [48] based on the OpenSceneGraph (OSG) [49] Application Programming Interface (API). The system operates at a frame rate of 60 Hz without synchronization of the images. For displaying the flight instruments, DLR's software 2indicate is used.
The projection dome of AVES measures approx. 3 m in diameter and is mounted on a electro-pneumatic hexapod motion cueing system, which allows for up to 1.5 m of vertical movement (Figure 2.1a). The overall assemblies mass exceeds 10 t.

2.1.2 Dual Pilot Active Sidestick Demonstrator (2PASD)

The 2PASD [50] is a fixed base simulator with two pilot seats, each equipped with cyclic and collective sidesticks and pedal inceptors which are connected to the CLS (see Figure 2.3). The primary visualization system consists of five Liquid Crystal Display (LCD) TV screens,

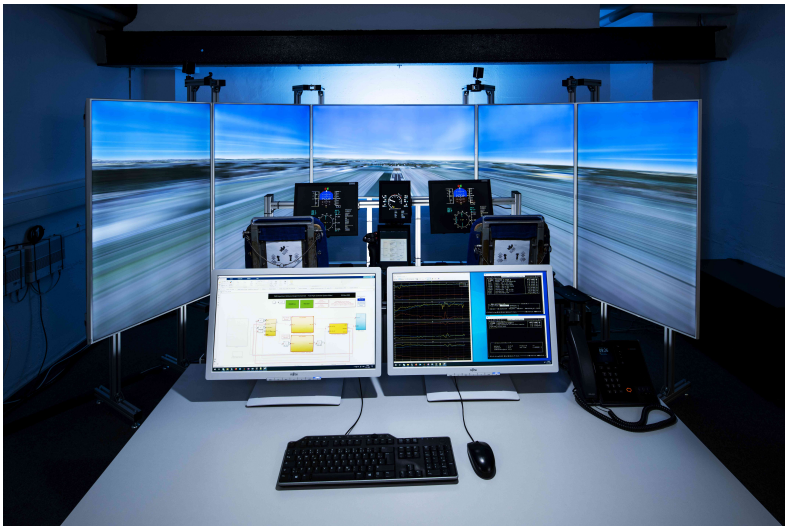


Figure 2.3: 2PASD Simulator ©DLR. CC BY-NC-ND

each measuring 55 inch (139 cm) diagonally and operating at a resolution of 1920×1080 pixels with an image refresh rate of 60 Hz without image synchronization. The screens are arranged in an arc around the pilot stations, with the center screen positioned in landscape orientation and the four outer screens placed in portrait orientation. The FOV covered by the TVs is approximately $150° \times 50°$ (horizontal $\times$ vertical). Additionally, a set of screens is mounted on the pilot stations in front of the seats to display flight instruments. The operator station for configuration and simulation control is located behind the simulator, providing easy access for operators to manage the simulation environment. The 2PASD's design allows for various cockpit configuration arrangements, such as side-by-side, single-pilot or tandem configurations.
The 2PASD features the same experimental system, flight model (HeliWorX), AVES IG and FCS as AVES. The specific software architecture is described in Section 2.2.2. Nevertheless, both simulators differ not only in their visual systems (projection dome vs. TV screens) and the cockpit setup but also in their CLS. In particular, AVES provides long-pole sticks for cyclic and collective inputs, whereas 2PASD relies on sidesticks for these control axes.

2.1.3 See-through and Monitor-based Headsets Used in This Work

For an initial insight into the potentials and limitations of different display technologies, a monitor-based and see-through headsets were selected to investigate these technologies for use in helicopter pilot training. These two headsets were selected because at the time of this work they offered the best characteristics in terms of display resolution, FOV and other critical factors such as weight: The **Varjo XR-3 Focal Edition** and the **JVC HMD-VS1D.**

Varjo XR-3 Focal Edition

The Varjo headset (Figure 2.4a) is a monitor-based headset with a FOV of 115° × 90° (horizontal × vertical). For the study, the Varjo XR-3 Focal Edition is used [51], which has been optimized for an object distance range of 30 cm to 80 cm approximately corresponding to the distance between pilot and cockpit instruments. The Varjo offers a focus area within the rendered image for which the image is displayed in higher resolution. The focus is located in the lower center and has a FOV of 27° × 27°. It provides a much clearer image that makes it easier to read the cockpit instruments (see display resolution in Section 1.1.3). The monitor-based headset weighs 1320 g and operates at a frame rate of 90 Hz, which is the recommended minimum according to NATO guidelines [22] (see refresh rate in Section 1.1.3).
As an additional functionality, the Varjo Base Software enables eye movement recording, allowing for eye tracking evaluations within the studies in this work. In addition to the Varjo Base Software, the Varjo Native Software Development Kit (SDK) provides a set of libraries and a C/C++ API that facilitate the integration and utilization of Varjo headsets' features and capabilities for developers.



(a) Varjo XR-3
©DLR. CC BY-NC-ND



(b) JVC HMD-VS1D
©REISER Simulation and Training GmbH. CC BY-NC-ND

Figure 2.4: Mixed Reality Headsets

JVC HMD-VS1W

The JVC headset (Figure 2.4b) features a see-through display using half-mirror technology, which blends virtual images with the real environment. This results in a natural, pixel-free image, without any peripheral blur or chromatic aberration [52]. It features a FOV of 120° × 45° for virtual image. The ability to see underneath the display extends the FOV to 72° vertically. The JVC headset operates at a frame rate of maximum 72 Hz, which is below the threshold recommended by the NATO guidelines. The JVC headset weighs a little less (980 g) than the Varjo headset, which may be advantageous for comfort (see comfort in Section 1.1.3).
The JVC headset can easily be used as a plug-and-play device. However, the lack of an SDK restricts the flexibility to adapt and fully utilize the possibilities of the see-through glasses. To achieve this, adjustments must be made to the image generator engine itself.

For comparison of the FOVs provided by the headsets and the typical FOV of humans, the different values are illustrated schematically in Figure 2.5. The human FOV is about 200° × 130° [53]. The headsets' FOVs are decreasing the natural visual field of the user. Horizontally, the headsets' FOV differ by only 5° and cover approximately 60 % of humans' horizontal FOV. Vertically, however, the headsets' FOV differ significantly, with the JVC headset covering approximately 35 % of the vertical human FOV and the Varjo headset covering 70 %.

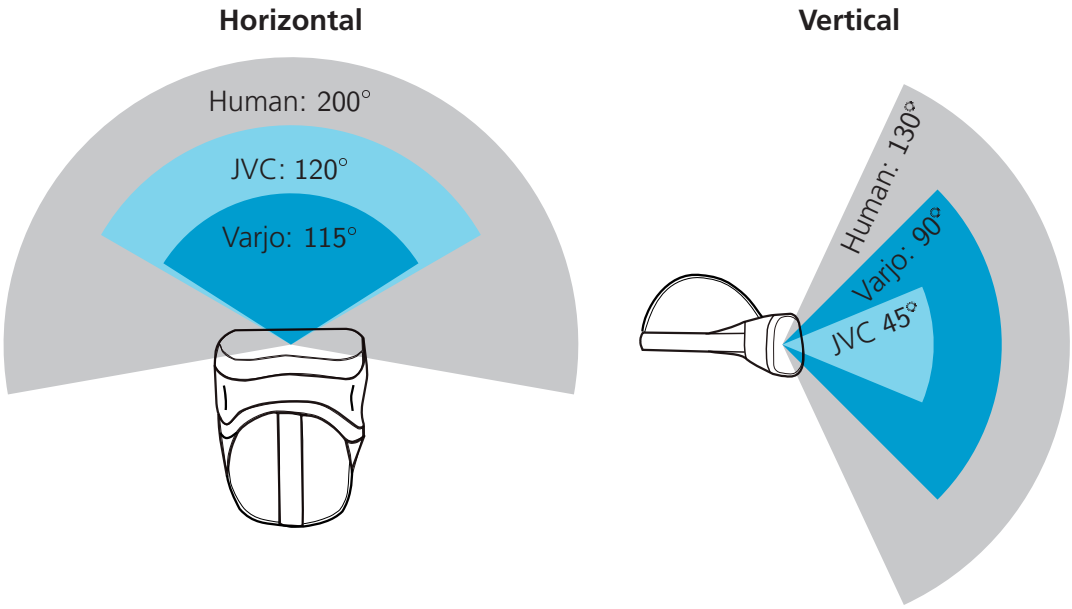


Figure 2.5: Field of View Comparison Adapted from [53], CC BY 4.0

2.2 Enhancing Helicopter Simulation with See-through and Monitor-based Headsets

For the implementation of MR technology into helicopter simulation, the selection of a simulator platform is crucial, as the platform's characteristic may have significant influence on the insights gathered for the usability of MR headsets for helicopter pilot training. Parts of this chapter have already been presented in earlier work by the author's publication in [50]. It presents the steps involved for implementing a MR headset into a simulation environment, including the selection of a simulator platform, presentation of the simulation framework, implementation options for MR configurations with monitor-based headsets, hardware setup for MR and latency measurement for evaluation of the simulation setup.

2.2.1 Selection of Simulator Platform for MR Development

To evaluate the simulator platform options for development of MR and VR applications, the benefits and limitations of the 2PASD were carefully weighed against those of a simulator with a fully equipped cockpit with dome projection like AVES. The 2PASD emerges as an attractive development platform due to several key advantages:

- **Flexible Cockpit Configuration:** The modular cockpit design of the simulator allows experimentation with different pilot setups (e.g., single-pilot or dual-pilot configuration), as well as variation of display configurations. The MR headset offers a highly customizable VE, enabling the virtual world to be merged with reality in multiple ways and therefore can adapt to different simulator configurations.
- **Optimized Green Screen Environments:** Ample space around the cockpit enables large-scale green screen arrangements (see Section 2.2.4 for green screen technique), ensuring robust head-tracking accuracy and minimization of interference from shadows caused by lighting systems. This setup creates a reproducible test environment with uniform illumination and minimal reflective surfaces.
- **Spatial Experimentation Freedom:** Researchers can explore alternative configurations for head tracking layouts and lighting settings without hardware constraints.

Despite these strengths, the 2PASD also has notable limitations:

- **Limited Physical Haptic Interaction:** The absence of a fully equipped cockpit with physical switches and buttons reduces realism and immersion during scenarios requiring tactile feedback.
- **No Motion Platform Integration:** The lack of a motion platform prevents studying interactions between visual cues from MR headsets and motion feedback, which are critical for enabling high-fidelity simulations.

Based on its specific strengths and limitations, the 2PASD was selected as the development platform for integrating MR headsets due to its dual role:

- It functions as a **flexible development environment**, while fulfilling requirements for helicopter simulation and thus allowing for piloted studies.
- The MR system components developed and insights gained through **implementation in the 2PASD can be directly applied to the AVES**, thereby reducing overall development effort and allowing for the utilization of the inherent advantages of both simulators.

2.2.2 Simulation Framework

This section gives a basic overview of the 2PASD's software architecture in order to provide the background of how the MR headset was integrated into the simulation environment. Furthermore, it outlines the data flow needed for displaying MR in order to provide an overview of the processing pipeline implemented for this research work.

The 2PASD simulation framework integrates the CLS and visual system hardware with a comprehensive suite of software modules, including the aircraft dynamics model, to enable the execution of flight simulation. The simulator is implemented as a distributed system running on multiple host computers. The underlying system architecture is provided by the simulation framework 2Simulate [54], which is a distributed real time simulation API developed for AVES. Building upon a common simulation framework allows for sharing software modules between the two simulators.

The 2Simulate framework provides a software library that enables real-time scheduling of simulation tasks and inter-process communication. The framework applies a centralized architecture through a central node, called interface software, which connects the various software modules running on different computers (Figure 2.6).
In 2PASD, this interface software is executed on the Interface Computer (IC), which runs on the QNX Neutrino 7 real-time operating system [55]. The IC software defines simulation tasks that can be executed at configurable frame rates and allows each task to establish a network connection using a User Datagram Protocol (UDP) driver provided by the 2Simulate library.

The IC is linked to the System Control Module (SCM), which handles the control software for the control inceptors during simulation. The SCM processes the stick forces and displacements provided by the pilots and transfers them to the IC, where they are used by the aircraft model to calculate the aircraft response. This model is written in C++ and replicates the flight dynamics of ACT/FHS, including models for the main rotor with dynamic inflow, tail rotor, fuselage, engine and flight control actuators, see Section 2.1.1.
The simulated aircraft's response is transmitted by the interface software from the IC to

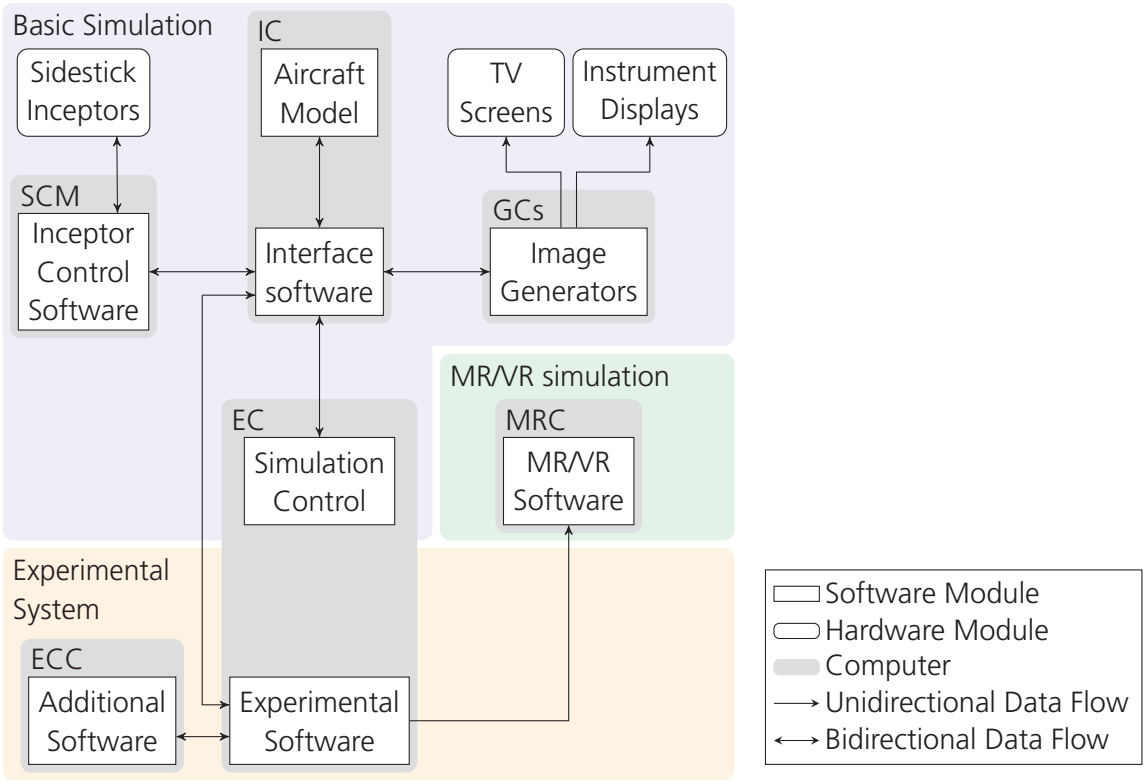


Figure 2.6: 2PASD System Architecture

three Graphics Computers (GCs), where it is used to feed a realistic visual representation of the simulation environment on the visual system. The first GC provides visuals for the central TV screen and flight instruments displays, while the remaining two computers each handle the graphics output for two additional TV screens on the left and right side.
For this work, 2PASD was enabled to additionally support MR/VR headsets by configuring the experimental system software to send the aircraft model's output to the Mixed Reality Computer (MRC), which connects the UDP data stream to the MR/VR environment implemented with Robot Operating System (ROS) (see Section 2.2.2).
From the Experiment Computer (EC), located just behind the simulator's cockpit (Figure 2.3), users can access and configure the simulation settings using the simulation control software, which is executed in a standard desktop environment.
The experimental system of the ACT/FHS (Section 2.1), which provides guidance and control functions for the helicopter, is integrated into the simulation by executing its software on the EC. The interface software forwards pilot control inputs from the SCM to the experimental software, where flight control algorithms generate actuator commands and send them back to the interface software as input to the aircraft model.

Further details regarding the 2PASD simulation framework and integration of the ACT/FHS experimental system may be found in [50].

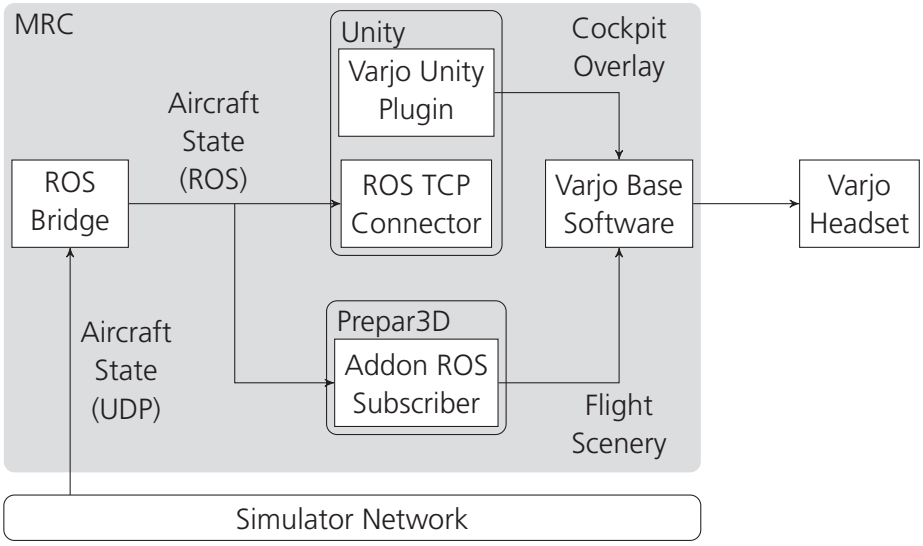


Figure 2.7: MR/VR System Architecture

Data Flow for Displaying Mixed Reality

The data flow of the Varjo headset is described in detail, as it offers more configuration options than the JVC headset due to its SDK. The adaptations made to the simulation environment for integrating the JVC headset are discussed in Section 2.2.5.

The aircraft state data of the helicopter is sent from simulator software to the MRC (Figure 2.7) via UDP. This information is needed to determine the aircraft's and pilots's position and orientation inside the VE and display the image of the Varjo headset accordingly. For the data transfer of the aircraft state to the rendering software (Section 2.2.3), ROS 2 Foxy [56] is used. Originally designed for robotics, ROS provides a robust and flexible middleware for inter-process communication as well as libraries and tools for development of robotic applications. ROS operates through message-based communication between modular software components called nodes, each node performing a specific task. In the configuration established for this work, they enable middleware functionality between the simulator and the MRC to transmit flight data such as the helicopter state vector, calculated by the aircraft dynamics model.

The ROS bridge that receives the helicopter aircraft state from the EC converts the UDP datagrams into standardized ROS messages. These messages are then transmitted to the ROS Transport Control Protocol (TCP) connector which is integrated within the game engine Unity (Section 2.2.3). This connector allows Unity to receive and interpret ROS messages, enabling the dynamic rendering of the helicopter's position and orientation in alignment with the aircraft state data. Additionally, Varjo provides a dedicated Unity plugin tailored for the integration with its hardware. The execution of this Unity application is managed by Varjo Base Software, which ultimately renders the VR content onto the MR headset.

Simultaneously, Prepar3D, which is used for rendering the outside scenery of the VE (Section 2.2.3) also receives the aircraft data from the ROS bridge. The data is transmitted to Prepar3D via an addon ROS Subscriber, which is embedded in the SimConnect framework of Prepar3D. The subscriber acts as a receiver node within SimConnect, enabling the helicopter's position and orientation in the VE to be updated using the incoming aircraft state information. When launching Prepar3D with VR settings enabled, the image is displayed according to the flight position and orientation in the Varjo headset.

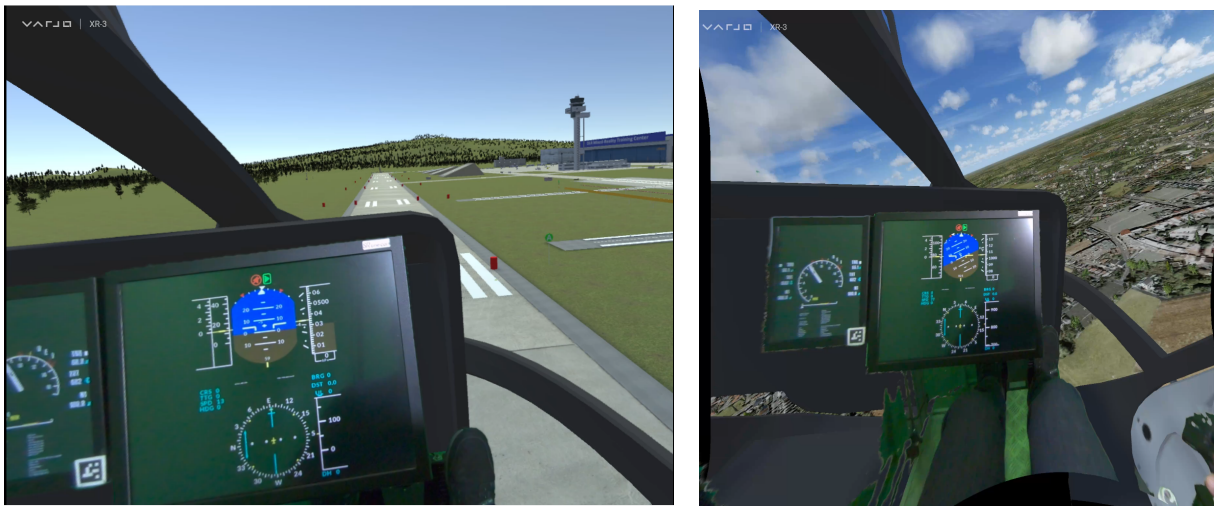
Varjo Base Software coordinates the MR experience by activating both applications: the Unity-based cockpit overlay is displayed over the Prepar3D scenery. This capability results from the multi-application function of the Varjo Base software, which allows several applications to run simultaneously and display their content on the headset. Alternatively, each individual render software may be used standalone for VE visualization.

2.2.3 Rendering Software

To create the VE, two software were employed: **Unity** and **Prepar3D**.
For creating a VE, a user may utilize game engines that support VR/MR technology. The two most capable engines for this use case are Unity and Unreal Engine. For this work, Unity was preferred as it provides a larger number of libraries and support for the interface for the utilized MR/VR headsets.
For displaying the scenery, a flight simulation software shall be used. From the wide range of options, for example Microsoft Flight Simulator 2024, X-Plane or FlightGear, the software Prepar3D was chosen due to its performance and availability at DLR. Prepar3D provides a detailed database for the representation of the earth's topography and features weather simulation.

Unity, a versatile game engine, supports both two-dimensional (2D) and three-dimensional (3D) rendering on multiple platforms. It is widely used by game developers, architects and designers for its flexibility in creating interactive environments. In Unity, the virtual environment can be customised to meet specific needs, such as the construction of detailed 3D helicopter cockpits (Figure 2.8a).

Prepar3D developed by Lockheed Martin, is primarily a flight simulation engine for gaming designed to depict realistic aircraft in various environments such as airports, mountainous regions and coastal landscapes. Its versatility is further enhanced by the SimConnect SDK, which allows developers to customize the software beyond the usual gaming or training scenarios. The software includes dynamic weather and environmental variables such as wind, all of which contribute to realistic simulation. SimConnect allows developers to modify 3D objects and integrate Prepar3D with external frameworks such as the 2PASD simulation framework, enabling data exchange between different software tools.



(a) VE Rendered With Unity
©DLR. CC BY-NC-ND

(b) Usage of Unity and Prepar3D
©DLR. CC BY-NC-ND

Figure 2.8: Different MR Rendering Options

While both Unity (Figure 2.8a) and Prepar3D may be operated independently, each offering its own VE, Varjo Base Software also allows to run both at the same time to combine both VEs in the Varjo Headset. For instance, a modified helicopter cockpit rendered in Unity can be displayed as an overlay within the scenery generated by Prepar3D (Figure 2.8b).

The two sources of the resulting merged image are not explicitly synchronized with each other. However, the positioning and orientation of the user in the VE is based on the aircraft state data received by the MRC, which is identical for both Unity and Prepar3D. Asynchronous displaying of the images might therefore be caused only by different processing delays of the two rendering pipelines and dynamic scheduling effects. These time differences are typically in an order of magnitude hardly perceivable by humans. As a consequence, no synchronization issues between both image sources could be observed.

This combination of both VEs adds the following value: In this configuration of the 2PASD, the helicopter simulator is equipped with active sidesticks instead of the conventional collective and long pole cyclic control sticks. Standard 3D helicopter models supplied with flight simulation software such as Prepar3D do not include sidesticks, as such a control arrangement is atypical for helicopters. To address this gap, a customized cockpit design aligning with the 2PASD setup was developed using Unity. The cockpit was displayed as an overlay, while the rest of the VE was rendered by Prepar3D.

2.2.4 Techniques for Creating MR Configurations With Monitor-based Headsets

To create an immersive and interactive MR scene, several techniques can be employed to merge real-world elements with virtual content. Three key methods that can be used to create a MR setup with the Varjo headset are:

- **Chroma Keying:** Chroma keying is an image processing technique for which image areas showing a specific color are replaced by virtual content. A background in the desired replacement color, usually light green or light blue, is captured by the scene cameras and marks the areas to display the MR. Green or blue are often chosen for the background because of their strong contrast to human skin tones, making it easier to distinguish between real world objects to display in the VE and the green screen background. The resulting composite image combines the foreground subject with virtual content and allows rendered imagery, effects and other elements to be integrated into the scene.

- **Masking Function/Cutouts:** Masking/cutouts provide a means of defining whether virtual content or real-world objects are displayed based on their position in the VE. This can work both ways: The camera captures a real-world object or scene and the mask defines an area where the virtual content will be inserted into the real-world image. Or the other way around, when the VE is created first, the real-world image captured by cameras can be projected onto a specified area within the VE.

- **Depth Camera:** In the context of MR, depth occlusion is a technique that provides the possibility to merge and filter real and virtual elements by utilizing the capabilities of depth recognition. In this way, one may integrate content from both the world and the VE without the use of a green screen solely based on the distance between an object and the headset.

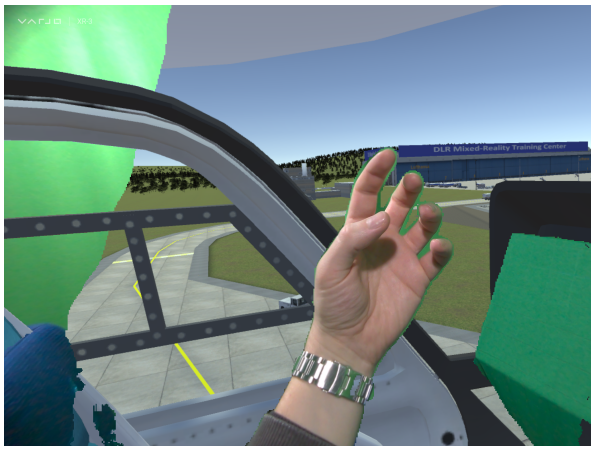
To obtain the best possible MR experience, all three options were tested with the Varjo headset and are described in the following:

- **Masks**, for this application referred to as **cutouts**, were implemented in the VE of the 2PASD. For this purpose, areas were defined in the cockpit where the real world, captured by the headset's cameras is displayed. The implemented cutout regions include the area of the cockpit instruments, where the live image from the headset cameras are embedded into the VE to make the 2PASD instrument displays visible to the pilot. Additionally, a cutout was created within the interior of the cockpit to allow the pilots to see themselves, the helicopter interior and the control inceptors. However, this approach resulted in parts of the body being cut off outside these designated areas, such as when an arm extends out of the cutout region into the area where only the VE is shown (Figure 2.9a). Multiple cutouts were chosen to individually mask cockpit elements and therefore create a more realistic VE.

- The Varjo headset achieves **depth occlusion** (Figure 2.9b) by combining Light Detection and Ranging (LiDAR) technology and stereo imagery to generate high-resolution depth maps and track objects in the scene. In the 2PASD environment however, the technology did not provide satisfying results, as the VE was not displayed accurately beyond the depth occlusion threshold. In addition, there was no accurate delineation around non-occluded objects, for example the pilot's hands.
- **Chroma keying** provides an effective means of merging real-world and virtual elements. An example of chroma keying with the Varjo headset is provided in Figure 2.9c. As can be seen in the image, the user's hand is clearly visible inside the VE and distinguishable from virtual elements. The resulting image has a highly realistic appearance, with only few noticeable distortions.



(a) Usage of Masks/Cutouts
©DLR. CC BY-NC-ND



(b) Usage of Depth Camera
©DLR. CC BY-NC-ND



(c) Usage of Chroma Keying
©DLR. CC BY-NC-ND

Figure 2.9: MR Software Implementations in 2PASD

Selection of MR Software Implementation

For the MR setup that was developed during this work, **chroma keying** was preferred over masking and depth occlusion for creating a MR environment due to the following **advantages**:

- **Seamless Fusion:** Among the three options, chroma keying provided the best merging of VR elements with the real-world environment with the Varjo headset in the 2PASD environment.
- **Realistic View:** Chroma key allows for a unobstructed view of the VE outside the cockpit, while still maintaining visibility of the interior space and the pilot, such as showing the feet of the pilot while allowing them to see the outside scenery through the chin window of the helicopter (small window located at the lower front). In addition, if the pilot points out the window with his hands the whole arm is visible (Figure 2.9c) in comparison to masking/cutouts where the arm is partially covered by the VR image (Figure 2.9a).
- **Clear Delineation:** The hand is clearly defined without any distortions or distractions caused by the VR environment.

Nevertheless, chroma keying also introduces some **disadvantages:**

- **Space Requirements:** A green screen is necessary, which can take up extra space and requires additional setup.
- **Risk of Artifacts:** If the cockpit holds elements colored in green, chroma keying artifacts may appear as "holes" in the real world where the VR is displayed.

2.2.5 Hardware Setup in 2PASD

For both headsets it is necessary to provide head tracking information such that the images of the VE will be displayed according to the pilots head movement. Two techniques are available to provide head tracking: inside-out tracking within the headset or using external head trackers. Whereas with the Varjo headset both tracking technologies can be used, the JVC headset is fully dependent on external tracking.
Varjo's inside-out tracking utilizes the headset's built-in tracking cameras to continuously monitor the environment to pinpoint the user's location and orientation. Nevertheless, the technology does not seem to be mature yet, as it results in strong delays and dropping image frame rates.

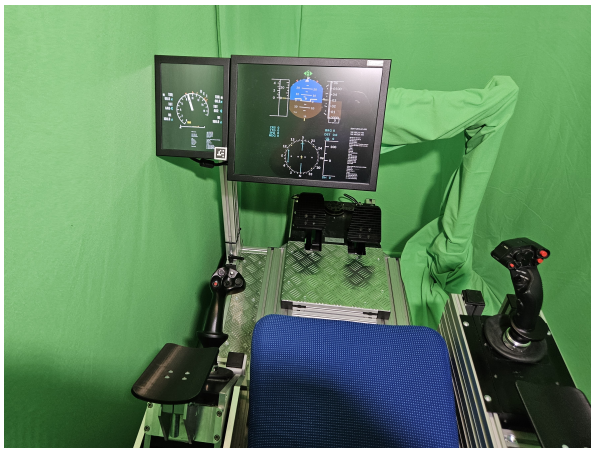
As for both headsets the VIVE Base Stations [57] were compatible, these were installed as external tracking systems for the experimental setup in the 2PASD. Each base station should be positioned in about 1.5 m to 5 m distance from the headset. As the pilot remains in the pilot seat, two tracking station were sufficient to cover the area of interest for tracking.

To use chroma keying for the MR configuration with the Varjo headset, a green screen was installed around the pilot seat (Figure 2.10). Additionally, the HTC Vive Base Stations are mounted on top of the green screen in the left and right corners of the front side.
In addition to the green screen, supporting frames of the 2PASD were covered with green fabric (Figure 2.11a) to avoid that the frames are visible in the VE. The resulting MR configuration is shown in Figure 2.11b, with the right pilot seat, the displays and control inceptors being displayed from the real world captured by camera and the left seat, the helicopter body frame, as well as the outside view being virtual.

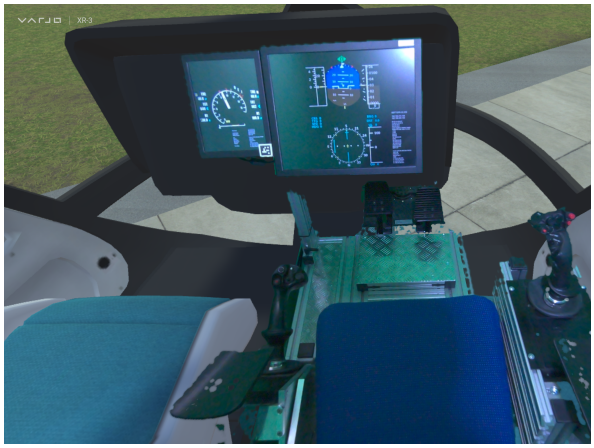
An important aspect which occurred during creation of the test setup is providing sufficient light. Therefore, spotlights (Figure 2.10) have been installed behind the pilot seat to illuminate the 2PASD cockpit. The green screen must be illuminated evenly to ensure accurate color recognition, as different light intensities can alter the color tone perceived by the headset cameras resulting in green screen ares to fall outside the chroma keying threshold. This can lead to image artifacts if the perceived color deviates too much from the predefined chroma keying color.



Figure 2.10: Greenscreen Setup in 2PASD
©DLR. CC BY-NC-ND



(a) Cockpit View With Green Screen Setup
©DLR. CC BY-NC-ND



(b) Cockpit View with MR Setup
©DLR. CC BY-NC-ND

Figure 2.11: MR Setup With Varjo Headset in 2PASD

Customization for the Integration of the JVC Headset

The JVC headset does not offer developers any SDK or alternative interfaces for customizing rendering applications, restricting its use to exclusively Unity or Prepar3D at a time. The data flow setup for the headset's operation aligns with the Varjo system's architecture (Figure 2.7), with the helicopter's aircraft state being transmitted from the ROS bridge to either Unity or Prepar3D. As a plug-and-play solution, the JVC headset enables Unity and Prepar3D to render VE directly onto its display when operated in VR mode.
To achieve transparent cutouts with the see-through headset, the rendered image should display a "self-illuminated" black color in the area of the cut-out. To configure the AR setup, masking elements were integrated into the software to prevent the VE from obscuring the physical 2PASD, while the physical cockpit mockup remained visible through the headset. Therefore, in the Unity cockpit the area where the cockpit instruments are located and the floor were masked in self-illuminated black, so that the pilot sees the 2PASD displays, the cockpit interior (controls, seats, etc.) and themselves (Figure 2.12). For the AR setup with the JVC, a non-reflective environment is beneficial to minimize real world elements shimmering trough the VE overlay. Therefore, fabric was used to cover reflective surfaces, as for example the frames of the 2PASD TV screens.

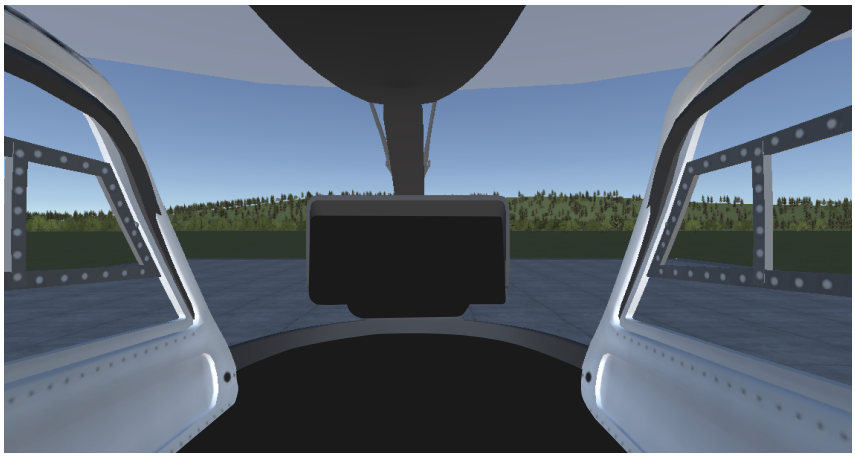


Figure 2.12: JVC Setup with "Self-illuminated Black" Colored Regions Used for Masking ©DLR. CC BY-NC-ND

2.2.6 Latency Measurement

Latency is a critical parameter for the assessment of flight simulators' real-time performance. The CS-FSTD(H) specification requires it to be less than 100 ms for level C and D FFSs and level 3 FTDs [6]. Quantifying latency is crucial as its reduction helps preventing vestibular-ocular mismatches that can lead to discomfort or nausea, ultimately improving the overall user experience in virtual environments (see Latency in Section 1.1.3). While the impact of simulator latency on pilots' presence has not been thoroughly explored, several studies have shown a strong negative effect of high latency on interaction quality [27].
Due to the importance of time delay in simulator environments, in this section the end-to-end latency of the 2PASD is examined, quantifying the time delay from pilot input to visual output on both the TV screens and the Varjo headset.
A simulator's latency can be attributed to several sources, including the computational time consumed by software components, network communication transport delays and asynchronous software execution on distributed machines.
To assess the performance of the 2PASD, the simulator's end-to-end latency for both the conventional and the MR configurations is evaluated. The measured time delay covers the time it takes for a pilot's input to be processed, the aircraft's response to be calculated and the simulated environment to be rendered. The following sections provide an overview of the utilized measurement method and a discussion of the results.

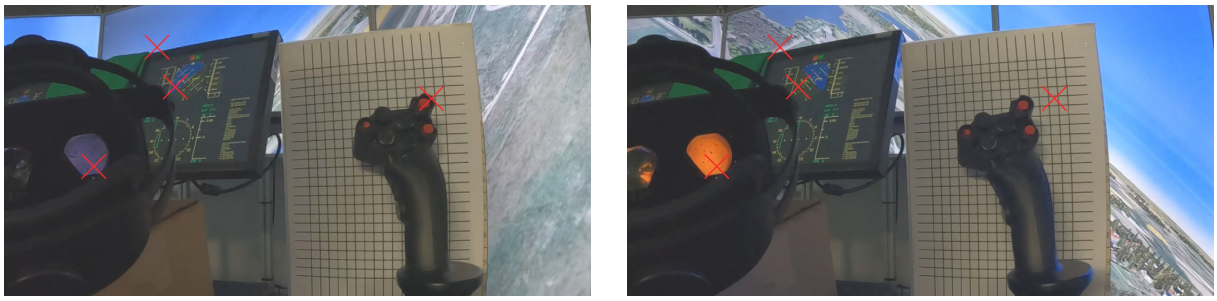
Measurement Method

The end-to-end latency of a system can be measured using a non-invasive approach that utilizes video capturing [58, 59]. The 2PASD's end-to-end latency was measured using the non-invasive setup to evaluate the latency of the standard TV screen visual system, the instrument displays and the MR headset. A high-speed camera was used to record control inceptor movements and visual system responses. The time delay between a movement of the stick and the visual response to this input is determined by counting the number of video frames between both events in the recorded videos.
To generate reproducible control inputs, the inceptor control software was configured to move the cyclic stick with a periodic step signal pattern in the lateral axis. In order for this input to result in a distinct visual response, the dynamics of the aircraft model were replaced by a step signal shaped response in the roll axis. The output roll angle ϕ of the modified aircraft model responds to a change of the stick position δ_y with

$$\phi = \begin{cases} -45^\circ & \text{for } \delta_y < -0.1\,\%\,, \\ 0^\circ & \text{for } -0.1\,\% < \delta_y < 0.1\,\%\,, \\ 45^\circ & \text{for } \delta_y > 0.1\,\%\,. \end{cases} \tag{2.1}$$

By applying the relation from (2.1), the visual systems display a distinct and reproducible change in the image for small movements of the stick. These distinct image changes can easily be observed in the recorded videos and allow for automated processing of the data.

For the extraction of the time delay from recorded video data, a specialized tool was developed using MATLAB [60]. For the analysis, the user is asked to select specific points in the image stream where notable changes in color or brightness occur during the step response. Figure 2.13 shows an example with two images captured during the latency measurements conducted for 2PASD. Figure 2.13a shows an image captured before a stick input, i.e. the idle state, while Figure 2.13b shows an image recorded after the stick was moved and the visual response to the stick input is visible.
The high-speed camera was oriented to capture the movement of the stick, which is visible in the right portion of the images, as well as the instrument displays and the TV screens, visible in the middle and the background of the images, respectively. In addition, the MR headset was placed in front of the camera, such that its displays are facing towards the camera. As the camera is placed far from the headset displays' focal points, the captured image is blurry and shows only a small portion of what a person wearing the headset could see. In order to robustly detect changes in the images displayed by the headset, it was configured to display a very simple VE: the space above the horizon is colored in blue, while the area below the horizon shows orange color. The resulting image is similar to the representation of the artificial horizon on flight displays.
The red crosses in Figure 2.13 are placed by the user and mark characteristic points in the image, with one being placed at the edge of the stick and the other markers being



(a) Initial State
©DLR. CC BY-NC-ND

(b) State After Step Response
©DLR. CC BY-NC-ND

Figure 2.13: Time Delay Analysis

placed at a position where the change in each of the visual systems' images due to the roll response is detectable through a distinct change of color.
The latency analysis tool processes the recorded video frame by frame. When the brightness of the marker on the stick increases significantly, the timestamp of the according frame is captured, as a change of brightness indicates a movement of the stick and hence a control input. The tool then proceeds to step through the video and captures the timestamps of frames when the pixel colors of the other markers change, indicating the response of the respective visual system. The individual time delay of each visual system is calculated as the difference between the timestamp marking the movement of the stick and the timestamp of the respective visual response.

The accuracy of this method for latency measurement depends on the frame rate of the camera being used. For the latency measurement conducted in this work, a camera is utilized that operates at a frame rate of $f = 240\,\text{Hz}$ and therefore provides a temporal resolution of $t_r = 1/f = 4.2\,\text{ms}$ for the measurement. As the initial movement of the stick and the step response of the visual system can occur at any time between successive video frames, the resulting worst-case accuracy is $t_a = 2\,t_r = 8.4\,\text{ms}$.

Setup and Results

To quantify the time delays associated with different simulation setups, measurements were conducted using two simulation modes: the basic simulation mode and the simulation with the experimental system enabled. Additionally, the latency of both fully virtual and mixed reality modes of the MR headset were investigated. Each simulation setup was evaluated by capturing at least 50 step responses for statistical analysis and a robust evaluation.

In the basic simulation setup, the signal flow differs from that of the experimental setup as the experimental system is not utilized. This basic simulation setup has also been investigated to estimate the latency of the experimental system and thus gain additional insight into how large the latency is in the simulator setup with and without the experimental

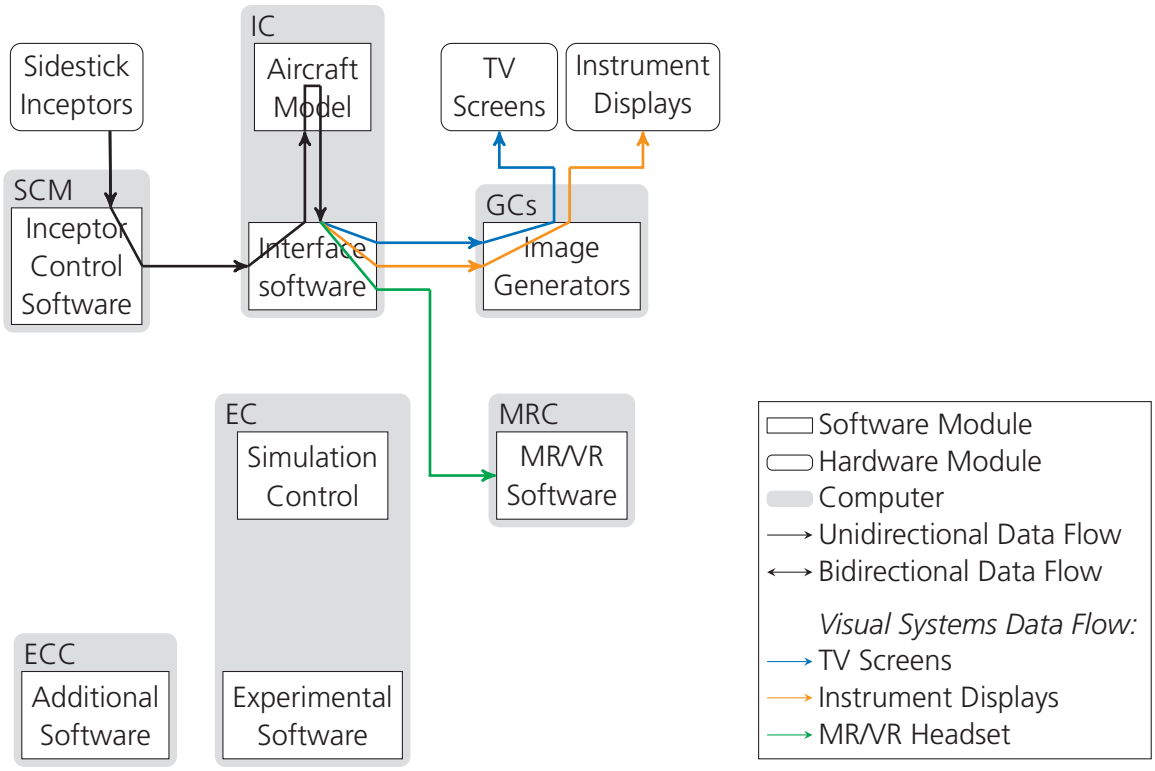


Figure 2.14: 2PASD Signal Flow in Basic Simulation Setup

system. The basic simulation setup deviates from the experimental system's signal flow since the experimental system is not part of this configuration. This basic simulation setup was further analyzed to estimate the latency of the experimental system, thereby providing additional insights into the latency in the simulator setup both with and without the experimental system.

The signal routing for the basic simulation mode is shown in Figure 2.14. This configuration involves the processing of a step input generated by the inceptor, which is initially handled by the SCM before being transmitted to the IC. Subsequently, the interface software forwards the signal to the aircraft model, prompting a step response defined by (2.1). This step response is then fed back to the interface software, which distributes it to the various components of the visual system, as visually indicated by the color highlighting in the figure. The signal is further send to the image generators, which render images for both the PFD and the simulated environment displayed on the TV screens. Additionally, the step response is sent to the MRC, where it is translated from UDP datagrams into ROS messages for rendering the image with Unity (Section 2.2.3) on the MR headset. This basic simulation setup was evaluated for both the MR and VR configurations of the MR headset, the TV screens and the PFD, resulting in four separate measurements for the end-to-end delay presented in Figure 2.16a. As the visual system elements do not provide real-time guarantees on the the processing delays, the end-to-end latency has been statistically analyzed. Therefore, at least 50 individual measurements were conducted for each configuration.

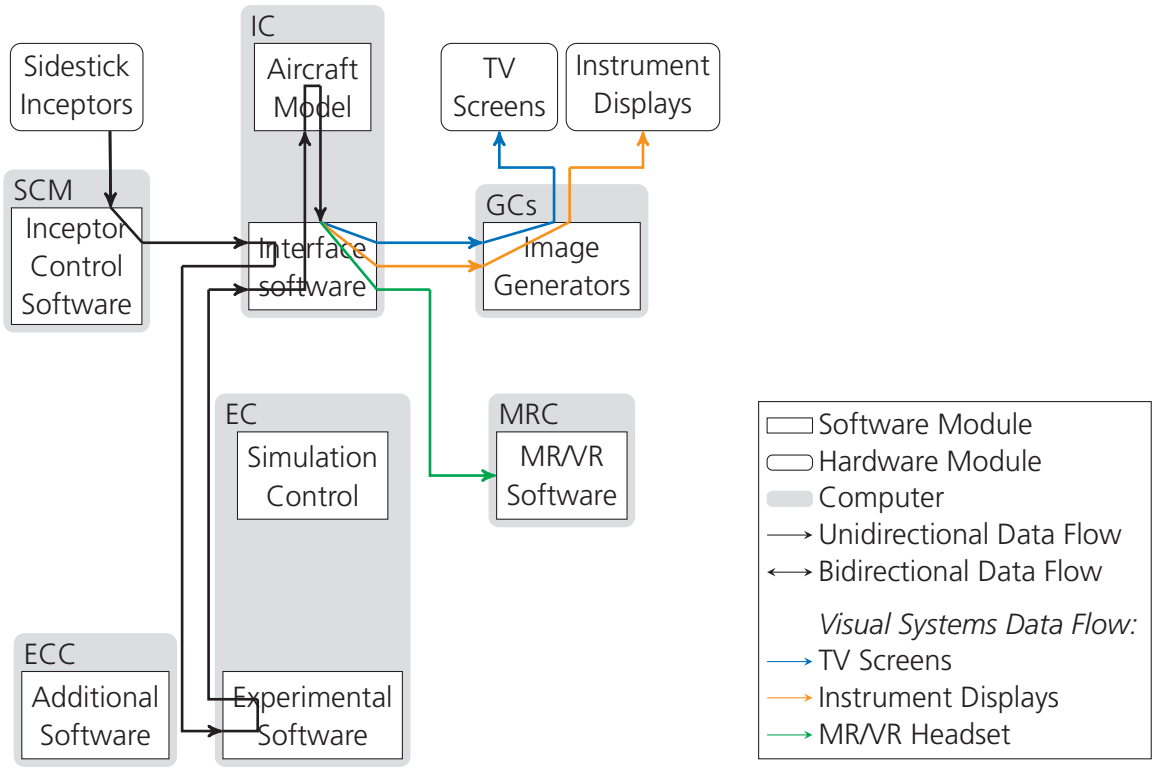
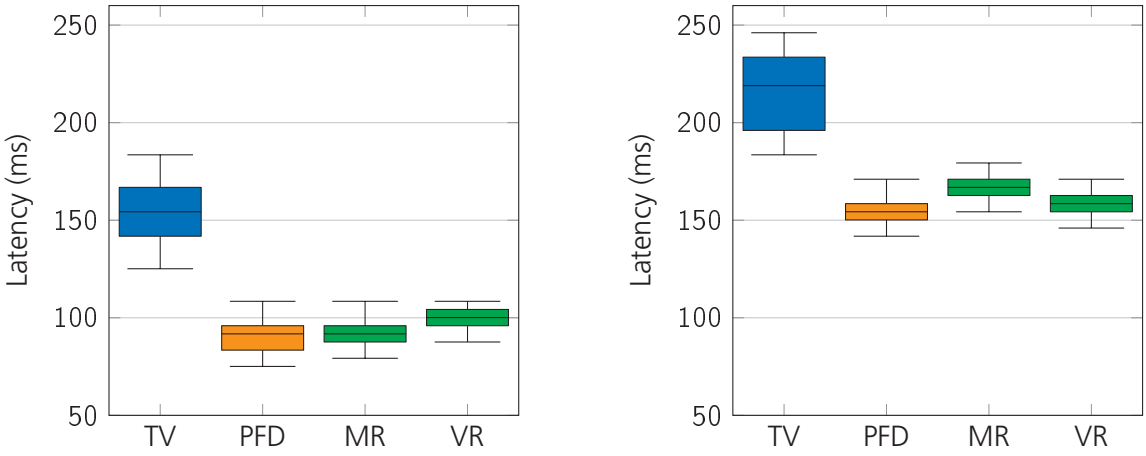


Figure 2.15: 2PASD Signal Flow With Experimental System

Notably, the latency for the TV screens has the highest value with a median of 154 ms. In contrast, the PFD and MR headset configuration show notably shorter time delays of 92 ms. The VR configuration demonstrates a slightly longer time delay by reaching 100 ms.
The comparison between the TV screen latency and PFD latency reveals that rendering the simulated environment requires at least 50 ms, considering both systems receive the step input simultaneously as the signal path is identical.

The TV screens exhibit significantly higher latency compared to the PFD and MR headset configurations. This disparity in time delay might be attributed to differences in display technologies or longer processing times of the AVES IG software driving the TV screens in comparison to the rendering software Unity. The TV screen configuration does also exhibit a higher jitter compared to the other configuration, as indicated by the longer whiskers in Figure 2.16a. This may be caused by the AVES IG software or may be attributed to the measurement method.
The TV screens display a complex scenery while the instrument displays as well as the MR headset do only show two distinct single-colored regions separated by the horizon. The measurement is based on the change of color of a specific pixel in the captured video images, therefore a more complex scenery may reduce the precision of measurements. For future evaluations, the TV screens should be configured to display the same artificial image as the other visual systems to increase the fidelity of the latency measurements.



(a) Time Delays in Basic Simulation Setup

(b) Time Delays With Experimental System

Figure 2.16: Time Delays of Both Setups

The second measurement configuration involves integrating the experimental system into the simulation (Figure 2.15). This configuration was used for the studies in this work, as the experimental system allows for utilization of flight control functions that, compared to the bare airframe model of the basic simulation setup, more closely resemble flying current helicopters (Section 3.1.2).
The SCM processes the input from the inceptor, which is then sent to the IC and routed to the EC. The experimental system, which in normal operation executes a flight controller algorithm, returns the signal unchanged for the latency measurement. This signal is then sent back to the IC and routed to the aircraft model. The aircraft model calculates a step response based on the given input according to (2.1) and the Interface Software distributes the response to the components of the visual systems (TV screens, instrument displays and MR headset) as in the basic configuration.
In the latency measurement configuration (Figure 2.15), the signal flow is the same for the TV screens, the PFD and the MR headset. This differs slightly from the signal flow in Figure 2.6, where the signal for displaying the image in the MR headset is transmitted directly from the EC to the MRC. The adjustment was made to enable comparability between the different displays. The results of the time delay measurements with the experimental system enabled (Figure 2.16b) show that the TV screens have the highest median time delay of 219 ms, the PFD showing a delay of 154 ms, while the Varjo headset exhibiting a delay of 167 ms in MR configuration and 158 ms in VR configuration.
The median end-to-end latency for all four signal paths is increased by 58 ms to 75 ms compared to the basic configuration, indicating that the experimental system introduces approximately 67 ms of time delay.

In contrast the first measurement, for which the experimental system was disabled, the time delay in the MR configuration is slightly longer than that in the VR configuration. Given that the signal path for VR and MR remains the same and is not altered by the

experimental system, it can be concluded that the measurement method or the analysis tool used to evaluate the camera images introduce some inaccuracy.
The end-to-end measurement only supplies relative delay times, lacking the capability to provide detailed information on a module-level basis. To further investigate and measure time delays on a module-level basis, the video-based setup is insufficient. A potential approach for more in-depth latency analysis involves modifying the software of the simulation components, specifically the interface software, to measure the elapsed time between sending a data packet to another module and receiving the response. This modification would facilitate identifying time delays of individual simulation modules.
The conducted measurements provide a valid baseline for optimizations, allowing for various measures to be applied to reduce latency and measure their effects on the overall latency. The latency measurement of the 2PASD setup provides valuable insights into the performance of the simulator in different configurations.

With regard to the investigations conducted in this research work and discussed in the upcoming chapters, the following key findings are derived, which should be considered in the evaluation of the results:

- The experimental system configuration did not meet the latency requirement of less than or equal to 100 ms specified in the CS-FSTD(H) for FFS level D. However, measurements demonstrated that this latency criterion can be achieved with the basic simulation setup.
- With the experimental setup that is being used in the upcoming studies, the latency is above 150 ms which may have an impact on the results.
- MR and VR latencies are comparable, so no major differences in the effects of latency can be expected when comparing the configurations.
- The TV screens have a significantly higher latency, which may have an impact on the results for the conventional simulation setup.

3 Potentials and Limitations of MR in Helicopter Simulation

This chapter presents two studies that aim at assessing the suitability of MR technology for helicopter pilot training. In Study 1, different implementations are evaluated by comparison of monitor-based display with a see-through headset within the 2PASD. The findings guide the selection of the most promising MR headset for the subsequent studies.
Thereafter, the selected MR headset has been integrated in a fully equipped EC135 cockpit. In Study 2, the MR headset is compared with the AVES's projection system. The study extends the flight tasks of the previous study by a malfunction procedure for increased pilot interaction.
The chapter concludes with a side-by-side comparison of all visual system configurations implemented in both studies across 2PASD and AVES, discussing key considerations for employment of MR technology, as well as its potentials and limitations in helicopter simulation.

3.1 Study 1: Comparison of See-Through and Monitor-Based Headsets

This section contains a comprehensive description of the experimental design, the task definition and the assessment methods used in Study 1. These elements also serve as the basis for Studies 2 and 3 and ensure a consistent approach to data collection and analysis. Parts of this chapter have already been presented in earlier work by the author's publications in [61, 62].

3.1.1 Experimental Setup

The primary goal of Study 1 is to evaluate a see-through and monitor-based headset, as well as explore various configuration options based on these technologies. The motivation behind the different setups is to explore the possibilities of MR implementation, which is a crucial aspect of addressing the research questions. This investigation shall provide valuable insights into the key considerations for implementing MR technology in a helicopter simulation environment (RQ1.1). By evaluating the different configurations, the most suitable MR headset technology (see-through vs. monitor-based) shall be determined for further usage in Studies 2 and 3.

The 2PASD's conventional visual system with five TV screens serves as the baseline for Study 1. The JVC HMD-VS1D is used for the see-through configuration, while the Varjo-XR3 Focal Edition is used as monitor-based headset. The technical characteristics (display technology, FOV, refresh rate and weight) of the different visual systems are listed in Table 3.1.
Utilizing the three visual systems, four different configurations are implemented for evaluation, see Table 3.2. The AR setup utilizes the JVC HMD-VS1D see-through headset, while the Varjo XR-3 is used for both MR and VR configurations. The 2PASD serves as a baseline reference for comparison purposes.

Table 3.1: Overview of Visual Systems.

Technical Specifications	JVC	Varjo XR-3	TV Displays
Display Technology	See-through	Monitor-based	Screens
Field of View (h × v)	120° × 45°	115° × 90°	148° × 48°
Refresh Rate	60 Hz / 70 Hz	90 Hz	60 Hz
Weight	1320 g	980 g	-

In the VR, AR and MR configurations, the visual environment is rendered with Unity (Section 2.2.3). Since the 2PASD does not provide a complete cockpit mock-up, parts of the cockpit are being displayed as rendered objects in the virtual world. For the different MR, VR and AR configurations, different mixtures of real and virtual worlds are required.

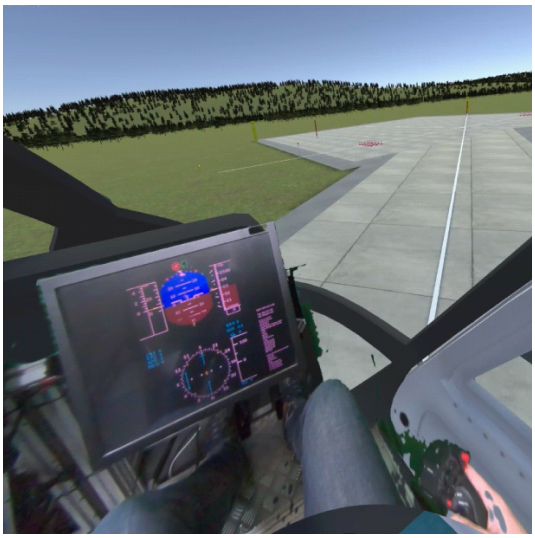
In the VR configuration, the whole cockpit with its interior and the pilot as an avatar is being visualized (Figure 3.1a). In the VR configuration, the movements of the control inceptors and the according arm and leg movements are animated. However, no body or hand tracking is being included. The flight tasks of the study (Section 3.1.2) require no interaction with the cockpit, as the pilots must have both hands on the controls at all times during the maneuver. Therefore, coupling the arm and leg movement to control inceptors is sufficient. For the MR configuration, a green box is built around the cockpit to use chroma keying (Section 2.2.5). In addition, tracking stations for head tracking and light sources for sufficient illumination are installed. In addition, for the MR configuration (Figure 3.1b), the virtual cockpit is adapted by eliminating the cockpit displays and the controls. Instead of the virtual cockpit instruments, the 2PASD cockpit mock-up is used including the controls and the displays for providing the PFD. For this blended reality chroma keying is used. Everywhere around the pilot seats, the 2PASD displays for the cockpit instruments and the controls a green box is installed. Wherever the color green is captured by the headset's cameras, the rendered image shows VR. Everywhere else the image of the camera is shown. No masking has been used for the MR configuration.

Table 3.2: Overview of Tested Configurations.

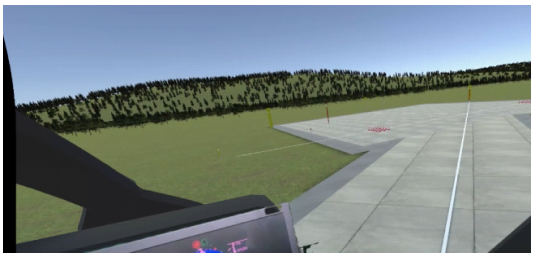
Configurations	JVC	Varjo XR-3	TV Displays
Conventional Setup (CON)	-	-	✓
Virtual Reality (VR)	-	✓	-
Mixed Reality (MR)	-	✓	-
Augmented Reality (AR)	✓	-	-



(a) 2PASD VR ©DLR. CC BY-NC-ND



(b) 2PASD MR ©DLR. CC BY-NC-ND



(c) 2PASD Representative for AR View ©DLR. CC BY-NC-ND



(d) 2PASD Conventional Setup (CON) ©DLR. CC BY-NC-ND

Figure 3.1: Visual System Configurations for Study 1

The JVC headset is used for the AR configuration (Figure 3.1c). The configuration is implemented in such a way that the VR is only displayed in the windows of the helicopter, everything else is transparent, i.e. the real world is visible. As a result, if the hands are within the defined window area, they are covered by the virtual images and are not visible to the pilot.
For the Conventional Setup (CON) configuration (Figure 3.1d), the TV displays show the same scenery as the headsets with identical visual cues, with the only difference that it is being rendered with the AVES IG. The five TV screens are arranged as follows: two displays are arranged vertically on the left and right and the fifth display is arranged horizontally in the middle, see Section 2.1.2.

3.1.2 Task Definition

The flight maneuvers are conducted in accordance with the Mission-Task-Elements (MTE) specification outlined in Aeronautical Design Standard (ADS-33) [63], which serves as a standardized framework for evaluating rotorcraft handling quality and flight performance. ADS-33 was originally designed to assess a helicopter's capabilities in a military context, but it has since become a widely accepted and utilized benchmark for handling qualities in rotorcraft research.

The flight tasks for the studies in this work utilize a subset of the well-defined flight tasks outlined in ADS-33, ensuring the results to be both repeatable and comparable. The pilots are tasked to conduct the hover and slalom maneuvers which introduce strict task boundaries, e.g. for the helicopter's position, velocity and altitude. Two sets of task boundaries or performance parameters are defined, with the compliance or violation of the boundaries determining the rated task performance to be classified as *desired*, *adequate* or *not adequate*, see also Section 3.1.3.

The **hover maneuver** should be initiated at a ground speed of 6 to 10 kts and with an altitude above ground of less than 20 ft. With the heading directed towards south, the pilot flies 45° relative to the heading of the helicopter to approach the hover point. As the pilot arrives at the hover point, they are tasked to hold the hover position for 30 seconds (Figure 3.2). In Table 3.3, the performance parameters for the maneuver are summarized. The hover maneuver has been chosen to study, among other aspects, the effects of the visual systems on the pilot's peripheral vision. In addition, a maneuver was sought that allows for evaluation of the visual systems with a virtually static image. Therefore, the hover maneuver was preferred over other MTE flight tasks, for example the hover turn. Furthermore, the hover maneuver represents a standard task that is routinely required during helicopter pilot proficiency checks.

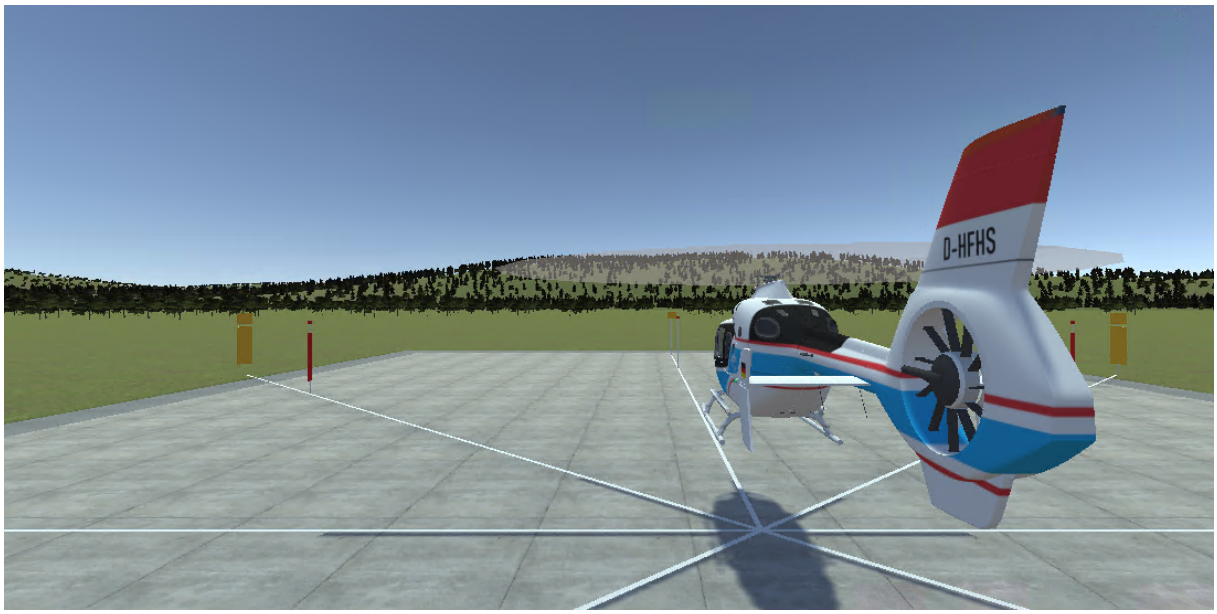


Figure 3.2: Mission-Task-Elements Test Course - Visual Cues for Hover Maneuver
©DLR. CC BY-NC-ND

Table 3.3: MTE Hover Performance Parameters based on [63].

Performance	Desired	Adequate
Attain a stabilized hover within X seconds of initiation of deceleration:	3 sec	8 sec
Maintain a stabilized hover for at least:	30 sec	30 sec
Maintain the longitudinal and lateral position within $\pm$ X ft of a point on the ground:	3 ft	6 ft
Maintain altitude within $\pm$ X ft:	2 ft	4 ft
Maintain heading within $\pm$ X deg:	5 deg	10 deg

Table 3.4: MTE Slalom Performance Parameters based on [63].

Performance	Desired	Adequate
Maintain an airspeed of at least X knots throughout the course	60 kts	40 kts
Accomplish maneuver below reference altitude of X ft:	Lesser of twice rotor diameter or 100 ft	100 ft

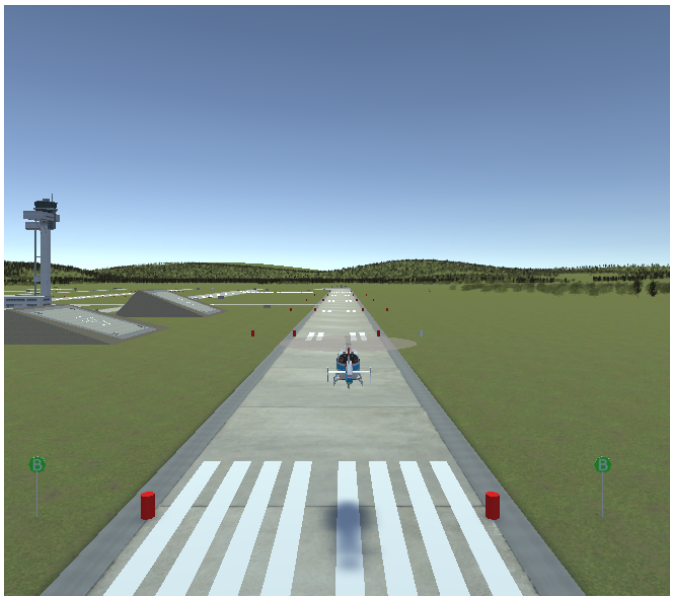


Figure 3.3: Slalom Test Course
©DLR. CC BY-NC-ND

The **slalom maneuver** is initiated with a stable forward flight at 60 kts ground speed. After passing the first gate, the pilots is tasked to maneuver through a slalom course comprised of four gates. For task completion, the pilot needs to fly four smooth turns at 500 ft interval (Figure 3.3). The performance parameters of the maneuver are listed in Table 3.4.
The slalom maneuver was chosen as it requires rapid changes of the helicopter's position and orientation. Compared to other MTE maneuvers, like the landing or pirouette task, it combines both high velocity close to the ground and large attitude changes. This allows for evaluating the visual systems in a scenario with highly dynamic imagery. In addition, the slalom maneuver was chosen to include a flight task that allows for evaluation of the tendency of the visual systems of causing simulation sickness (see velocity/direction changes in Section 1.1.3).

To perform these maneuvers, visual cues for each maneuver have been integrated into a MTE test course environment (Figure 3.2, Figure 3.3). Both the hover maneuver and the slalom maneuver were flown three times by each pilot in all of the configurations.

The simulation setup provides a flight controller for command augmentation and stabilization of the aircraft to increase handling qualities and reduce pilot workload. The flight control algorithms were selected to loosely resemble the flight control system of common utility helicopters like the EC135. Therefore, the flight controller provides the following modes: for pitch and roll axes, an Attitude Command Attitude Hold (ACAH) mode is enabled. In the yaw axis, rate damping is provided by a Stability Augmentation System (SAS), while collective inputs translate to Rate of Climb (RC) commands. This flight controller is used for all three studies in this thesis.

3.1.3 Assessment Methods

For the evaluation, different assessment techniques are applied. On the one hand, standardized subjective evaluation methods in the form of questionnaires are used. In addition, in the debriefing the pilots were asked some questions related to ergonomics of the headsets. On the other hand, for objective assessment, aircraft state data have been recorded to compare the flight performance of each pilot for the different configurations. Additionally, eye tracking data were recorded to investigate the gaze behavior of the pilots within the different configurations.
In the sections that follow, the various assessment methods listed in Table 3.5 are shortly presented and one method each is selected to evaluate presence, simulation sickness, flight performance and gaze behavior for the studies conducted in this thesis. This is followed by a detailed description of how the data are analyzed in this work.

Table 3.5: Assessment Methods

Subjective		**Objective**	
Presence	**Simulation Sickness**	**Flight Performance**	**Gaze Behavior**
Simulation Sickness Questionnaire	Presence Questionnaire (PQ, Witmer)	MTE	Fixation per Aera of Interest
Misery Scale (MISC)	Presence Questionnaire (SUS, Slater)	Predefined Flight Tasks	Duration of Fixation
Motion Sickness Assessment Questionnaire (MSAQ)	Igroup Presence Questionnaire (IPQ, Schubert)		Dwell Time, etc.

Presence

Definition of Presence:

In the literature, the sense of presence appears to be the key factor for the evaluation of virtual environments [64]. However, there are different definitions on presence and also several questionnaires to evaluate it. Witmer et.al. [65] define presence as "the subjective experience of being in one place or environment, even when one is physically situated in another".

Assessment of Presence:

Among various questionnaires for evaluating presence, the **Presence Questionnaire (PQ) of Witmer and Singer** is the most cited questionnare [66]. The PQ uses a seven-point scale and consists of 32 questions (see Appendix A.1).
Another presence questionnaire is **the SUS questionnaire** of Slater Usaoh and Steed [67]. This questionnaire consists of six questions that aim to evaluate presence in relation to real experiences (see Appendix A.2).
The **Igroup Presence Questionnaire (IPQ)** from [68] consists of 14 questions gathered from various sources, including the work of Witmer et al. and Slater et al., as well as original questions.

While the SUS and IPQ are more concise and allow for faster completion, the PQ offers a more detailed and comprehensive assessment of presence, allowing pilots to engage in more in-depth reflection on their experiences during evaluation. However, answering all questions in the PQ is time-intensive and as they are formulated to evaluate unspecific VEs, not all questions fit perfectly with helicopter scenarios in this study. Nevertheless, with a more detailed briefing of the questionnaire, it is possible to frame the questions for the study.

The **PQ from Witmer and Singer** was selected for the evaluations in this work. While it is more time-consuming and consists of numerous questions, it offers valuable insights into various aspects of the VE.

Subscales, Grading Scheme and Visualization Methodology for PQ:

Each question of the Presence Questionnaire is assigned to at least one of the following factors: Control Factors, Sensory Factors, Distraction Factors and Realism Factors , which have an influence on presence and are considered within different **subscales** according to Witmer and Singer: Involvement, Natural, Resolution and Interface Quality.

Involvement denotes a state of active engagement, where an individual's cognitive resources (attention and energy) are directed towards a specific activity or event that holds personal significance or relevance. This state is marked by a subjective experience of being deeply involved, invested or absorbed in the situation, reflecting a strong emotional and psychological connection.
Therefore the involvement subscale measures various aspects of users' experiences within a VE, including their perceived ability to control events and responsiveness of the VE to user-initiated actions, as well as their level of involvement in visual aspects and overall engagement with the VE.

In the context of VE, **Natural** refers to a sense of familiarity and consistency with real-world experiences. It describes interactions that feel intuitive and effortless, where the VE behaves in ways that are consistent with everyday life. The natural subscale assesses three key

aspects of users' experiences within a VE: the naturalness of interactions, the consistency of the VE with real-world experience and the intuitive feel of controlling locomotion through the VE.

The **Interface Quality** items evaluates two key aspects of users' experiences within a Virtual Environment (VE): the degree to which control and display devices facilitate or hinder task completion and the participants' ability to focus and concentrate on their tasks.

Two items that measured similar but distinct aspects of user experience within a VE do not group together with the other items. These items, which focused on the ability to closely examine objects and view them from multiple angles, are identified as **Resolution** items.

The resulting assessment of presence using PQ ratings is determined by calculating the arithmetic mean of the individual questions' ratings, resulting in values on a scale from 1 to 7. This method is applied for both the Total Score (TS) and the subscale ratings.
The PQ scale does not provide a meaningful indication of the minimum level of presence required in helicopter simulations. The reason for this is that the numerical value does not contain a concrete classification that could determine its suitability for the simulation context. To address this limitation, an additional grading scale is employed adapted from [69], which complements the original PQ questionnaire evaluation. This scale categorizes PQ scores based on thresholds: scores below 50 % (4 points) are considered as *unacceptable* and marked red, increments of 10 % above that point rank the score as *unsatisfactory* (light red), *marginal* (orange), *satisfactory* (yellow), *very good* (light green) or *excellent* (green).

Simulation Sickness

Cause of Simulation Sickness:

As introduced before with the use of VR and MR headsets, simulation sickness may occur (Section 1.1.3). Information from different sensory organs is used for the human sense of orientation. All sensory information are processed to determine the current position and movement in space. However, if the visual information creates a mismatch with the vestibular information, a sensory conflict occurs. If this conflict persists for some time, this can lead to simulation sickness [70].

Assessment of Simulation Sickness:

For the assessment of simulation sickness, various questionnaires exist were proposed.
The most widely used questionnaire is the **Simulation Sickness Questionnaire (SSQ)** [21, 70]. The SSQ [71] is a standard assessment method to quickly evaluate 16 physical symptoms which can occur during simulation studies. When applying the SSQ, participants are asked to rate their symptoms with the different graduations *none*, *light*, *moderate* and *severe* (see Appendix A.3).

Another questionnaire with similar purpose is the **Misery Scale (MISC)** [72]. This scale leverages the understanding that unpleasant symptoms typically precede others, such as dizziness, headaches, cold sweat and stomach discomfort, resulting in a condensed questionnaire, which distills the key symptoms into a concise format.

A third questionnaire, the **Motion Sickness Assessment Questionnaire (MSAQ)**, is used in [73]. The MSAQ contains 16 phrases, such as "I felt sick to my stomach", that can be answered on a scale from 1 (*not at all*) to 9 (*severely*).

The **SSQ** was selected for the study because it is quick to answer, provides specific information about which symptom has occurred and is commonly used.

Subscales, Grading Scheme and Visualization Methodology for SSQ:

The SSQ is composed of three distinct **subscales** that measure various aspects of simulation sickness symptoms. The **Nausea** subscale assesses gastrointestinal distress, including increased salivation, sweating, nausea, stomach awareness and burping. In contrast, the **Oculomotor** subscale evaluates visual and cognitive fatigue, such as fatigue, headache, eyestrain and difficulty focusing. Meanwhile, the **Disorientation** subscale measures balance-related symptoms, including vertigo, dizziness with eyes open or closed and blurred vision.
For the evaluation in this work, an additional **grading scale** is applied adapted from [74] where a grading scale is suggested which classifies SSQ ratings into five categories, *negligible*, *minimal*, *significant*, *concerning* and *bad*.
The SSQ score reaches from 0 to 236 score points. The questionnaire and the calculation of the results are presented in the Appendix A.3. To facilitate comparative analysis with the corresponding PQ diagram, the simulation sickness score axis was inverted and truncated at 35 to improve visual alignment. Furthermore, a similar color coding with same **colors scheme** was incorporated to enable comparison of the results with those obtained through the PQ assessment. Therefore, following applies: *negligible* (green), *minimal* (light green), *significant* (yellow), *concerning* (light red) to *bad* (red).

Flight Performance

In addition to subjective evaluation methods, pilots' flight performance during respective maneuvers is evaluated (see Section 3.1.2).

Definition of Flight Performance:

In this study, flight performance is defined in the context of flight task performance, where pilots' execution of predetermined maneuvers, i.e. hover and slalom, is evaluated using objective measures (e.g., position error, speed deviation) under the same environmental parameters, including visual cues, within a MTE test course.

Assessment of Flight Performance:

The ADS-33 [63] and the revised successor document, the MIL-STD-32742 [75], specifically by definition of the MTEs, provides a standardized, quantitative framework for objective assessment and comparison of pilots' performance within specific flight tasks. Under consistent flight parameters and visual cues, pilots are tasked with well-defined flight assignments.

For this work, an objective assessment of pilot's performance shall be carried out in order to support pilot's subjective ratings generated with questionnaires. The assessment of flight task performance as specified in ADS-33 is chosen, utilizing **Mission-Task-Elements** for evaluation.

The ADS-33 standard is primarily used to evaluate all aspects of helicopter handling qualities. The main objective of this thesis is to compare the influence of different visual system configurations (VR, MR, AR and CON) on a pilot's flight performance in order to understand the impact of each system. Therefore, this thesis specifically focuses on analyzing and comparing the individual flight performance of each participating pilot under different visual configurations, without focusing on the actual handling quality of the helicopter simulation.
The standardized MTEs (Section 3.1.2) provide a valuable basis for flight tasks that ensure the repeatability of specific flight maneuvers, which are crucial for consistent performance evaluation.
To assess the pilot's adherence to certain parameters such as altitude, heading and speed, aircraft state data can be analyzed retrospectively to determine whether these parameters were maintained within the required limits. Complying to ADS-33, each parameter can be classified as *desirable*, *adequate* or *not adequate*, providing an overall evaluation of how well the pilot performed their flight task within the different visual configurations.
The three ratings for the flight performance are used in the ADS-33 framework for handling quality evaluation. The rating system originates from the Cooper-Harper Handling Qualities Rating Scale, a decision-tree questionnaire, that asks pilots to judge how well the aircraft can sustain the required performance level under the current workload. The decision tree yields three levels that are defined as follows [76]:

- **Level 1:** The performance of the aircraft is satisfactory, no improvements are required and the flight characteristics are good enough to fulfill the MTE.
- **Level 2:** The aircraft has a deficiency that should be improved, but the deficiency does not prevent the mission. The flight characteristics are still adequate, although they may slightly increase the pilot's workload, slightly decrease mission fulfillment, or both.

- **Level 3:** The aircraft's performance deficiency must be corrected. The aircraft can still be controlled, but only with considerable additional effort, so that the MTE cannot be reliably completed.

These three levels derived from the Cooper Harper Scale are directly mapped to the task performance categories in [77]: Level 1 corresponds to *desired*, Level 2 to *adequate* and Level 3 to *not adequate*.

Eye Tracking

Relevance of Eye Tracking:

Visual information accounts for approximately 80 % of human perception, with substantial implications for situational awareness in aviation, particularly regarding aircraft orientation and air traffic [78]. Consequently, training programs for helicopter pilots focus on learning specific eye tracking patterns, such as scanning the external environment (out-the-window - OTW) to detect air traffic and frequently checking in-cockpit flight instruments (in-the-cockpit - ITC). In emergency situations, rapidly obtaining an overview of the issue through scanning cockpit instruments can be life-saving. The in-the-cockpit/out-the-window ratio (ITC/OTW ratio) provides a simple measure of how gaze behavior may be influenced by using different visual systems.

Eye tracking and gaze behavior in helicopter training have been studied by many authors to understand how pilots allocate attention and manage workload during various flight operations. Research has shown that pilot workload and attention allocation change constantly during challenging flight conditions [79], here specifically autorotation, and that experienced pilots develop distinctive eye movement patterns through extensive training, enabling rapid acquisition of critical information and quick reactions to carry out emergency procedures.
Moreover, research has explored the relationship between visual scanning patterns and cognitive states of helicopter pilots [80]. The analysis of visual scan data collected in simulated helicopter navigation tasks has shown that visual scan measures can provide insights into students' cognitive states during navigation tasks. Additionally, visualizing scan patterns can help instructors understand a trainee's approach and identify strategies employed.
Another study investigated the scan techniques of helicopter pilots during low-altitude high-speed flight [81], demonstrating that more experienced pilots develop different scan patterns. They spend more time looking at the ITC and less time looking OTW, which is crucial for maintaining required altitudes in low-level flight. The findings may support the development of guidelines for training programs and the design of head-up displays for helicopters.
Furthermore, research has examined the mental workload and visual scanning of helicopter

pilots during takeoff and landing simulations [82], identifying that expert pilots had significantly lower mental workload compared to novice pilots, especially in challenging conditions. The study also found that mental workload increased with task difficulty for both groups and that pilots made more visual fixations during cruising and longer fixations during landing. Experts made more instrument-focused visual fixations during cruising. The findings suggest that cockpit system design should prioritize reducing mental workload and guiding visual attention to improve safety.
In addition, a study compared pilot gaze behavior during a standard mission profile of a helicopter in real and simulated flight, finding general pilot gaze behavior characterized by looking OTW most of the time [83]. The study found differences between real and simulated flight in fixation duration, indicating a more complex information gathering process in the simulator.

In this thesis, it is assumed that varying simulator configurations affect gaze behavior, based on the relevance of specific gaze patterns for helicopter flight as reported by several authors and the investigated difference in gaze behavior between real-world and simulator flight. Therefore, eye tracking shall be analyzed in the studies for this work to investigate if MR has an impact on gaze behavior.

Eye Tracking Analysis:

Eye tracking analysis involves the examination of the gaze behavior of subjects, typically measured in terms of fixation or gaze durations, which refer to the length of time the subject spends focusing on a specific Area of Interest (AOI). In this study, several representative metrics have been identified as relevant for understanding the gaze behavior of pilots [84]:

- **Duration of fixation:** Longer durations suggest difficulties in extracting information or indicate appealing objects.
- **Fixation Count and Total number of fixation:** Both measure efficiency in information search.
- **Fixations per area of interest:** More fixations suggest higher perceptibility or significance.
- **Average fixation count/duration:** Reveals how long the average fixation lasted for.
- **Fixation rate:** Represents attention frequency.
- **Dwell time:** Measures time spent at a specific area.

The **area of interest** is chosen in this study to understand the gaze behavior of pilots during simulated flight, particularly in terms of whether they focus on ITC to perceive flight information from the flight instruments or OTW to, for example, scan air traffic or other visual cues. To gain an initial insight into the gaze behavior of pilots with different visual systems, the **in-the-cockpit/out-the-window ratio** is examined.

For the estimation of ITC/OTW ratio, the Focus Distance Distribution (FDD) method is being used based on the publication of the author of [85]. In contrast to other approaches, it takes the error of the eye tracking measurement system into account. The ITC/OTW ratio metric is based on the gaze distance, which is derived from the eye angles of both eyes by calculating the intersection of the eye normal vectors. Small errors in the eye angle measurement may result in large errors of the calculated gaze distance as the gaze distance is typically at least one order of magnitude larger than the distance between both eyes. To account for these errors, the FDD method calculates model distributions of gaze distances for different ITC/OTW ratios based on a normally distributed error of the eye angle measurement. By comparing each model distribution to the actually measured gaze distance distribution, the ITC/OTW ratio with the highest likelihood is determined.

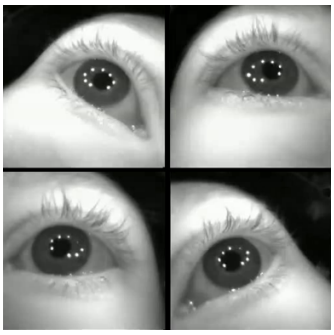
Eye Tracking Measurement Systems:

As eye tracking measurement system for the conventional simulators without wearing a MR headset, the Tobii Pro Glasses 3 are used (see Figure 3.4a). The Tobii Pro Glasses 3 utilize the Pupil-Center Corneal Reflection (PCCR) method, which is an eye tracking method that uses a near-infrared light source to create a reflection on the cornea, known as a "glint"(see Figure 3.4b). The reflection is then used in conjunction with pupil center detection to accurately track eye movements [86]. The Tobii Pro Glasses 3 feature eight infrared illuminators per eye, witch illuminate the eyes to assist the eye tracking sensor, and two cameras per eye to record eye orientation and movement. Additionally, the glasses are equipped with a scene camera, which captures the scenery in front of the user. The gaze tracking frequency is 50 Hz and given accuracy is 0.6° [87]. In addition, studies in [88] demonstrate that the Tobii Glasses 3 achieves an accuracy of approximately 0.96° sigma for a straight view to the target and even higher if the individual does not gaze directly ahead.

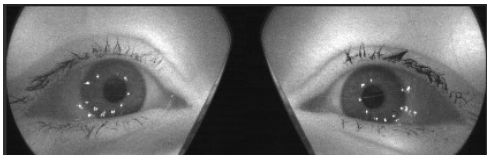
The Varjo eye tracking system employs a dual-camera approach, with one camera dedicated to tracking each eye independently (Figure 3.4c). This configuration enables the computation of a combined gaze ray using either binocular (both eyes) or monocular (single eye) input, if one eye is not tracked. The gaze tracking frequency is 100 Hz and the system is achieving accuracy down to one degree [89].



(a) Tobii Pro Glasses 3
©DLR.CC BY-NC-N



(b) Tobii Pro Glasses 3 Eye Cameras
©DLR.CC BY-NC-ND



(c) Varjo XR-3 Eye Camera
©DLR.CC BY-NC-ND

Figure 3.4: Different Eye Tracking Systems

3.1.4 Pilot Experience

Overall, three experienced helicopter pilots participated in the study (see Table 3.6). All were familiar with see-through and monitor-based headsets. Two of them are qualified as helicopter test pilots.
The results of the studies presented throughout this thesis is based on the evaluations conducted with these three pilots. For the remainder of this work, they will be referred to by Latin letters, as Pilots A, B and C.

Table 3.6: Pilot Experience.

Pilot	Total Flight Hours	Test Pilot	HMD Experience
A	6850 h	Yes	Yes
B	850 h	No	Yes
C	1500 h	Yes	Yes

3.2 Results of Study 1

This section presents the key outcomes for all tested configurations (MR, VR, AR and CON) from Study 1 conducted in 2PASD. Initially, the results of the PQ and SSQ questionnaires are analyzed. Then, the flight performance of all pilots in hovering and slalom maneuvers is presented. Subsequently, the flight performance of Pilot B with all configurations is discussed in detail.

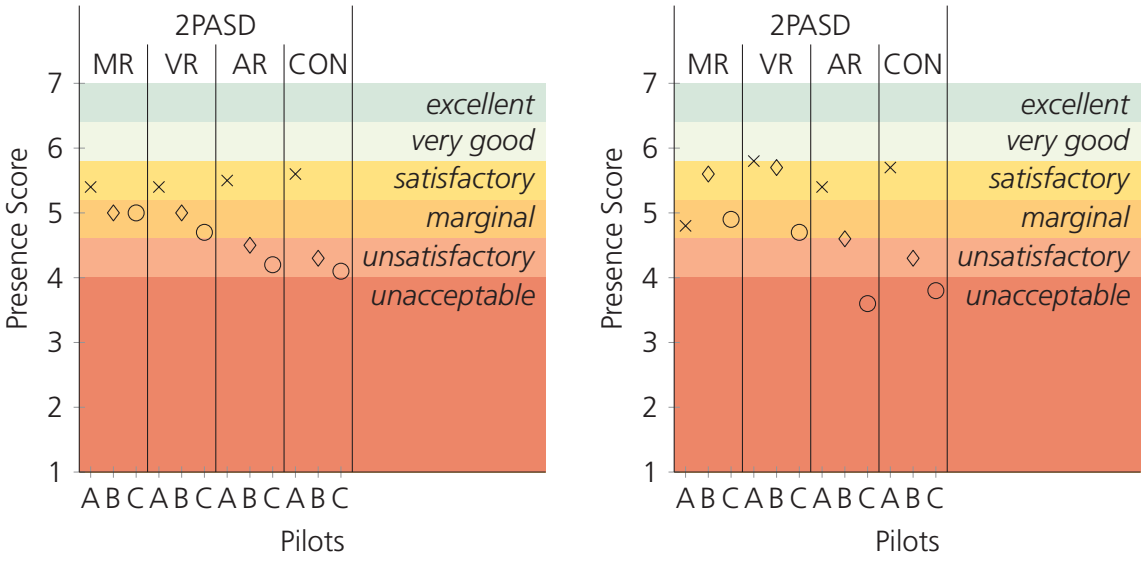
3.2.1 Results of Questionnaires

The following sections provide a comprehensive analysis of the questionnaire results obtained from Study 1. The PQ results of all three pilots for both hover and slalom maneuvers are presented and examined in detail based on the description in Section 3.1.3. In addition, the outcomes of the SSQ are reported, which was used to evaluate the level of simulation sickness experienced by pilots during simulated flight.

Results of Presence Questionnaire

The **PQ Total Score results for the slalom maneuver** (Figure 3.5a) show a distinct trend, with Pilots B and C exhibiting a preference for the MR configuration, followed closely by the VR configuration. The similar ratings of MR and VR configurations can be attributed to the fact that both utilize the same monitor-based headset and differ only in the display of real elements within the MR configuration. In contrast, pilot ratings for the AR configuration were lower, while the CON configuration received the lowest PQ score.
Pilot A's results deviated from this trend, with their PQ score not aligning with the other pilots' reported preference for the MR configuration. This discrepancy may either mark Pilot A's results as outlier data or suggest limitations in using PQ to assess helicopter simulation experiences. As Pilot A reported difficulties in understanding and answering the questionnaire, a more detailed briefing for the questionnaire could possibly improve its utility in this domain, as it was found to adequately capture the experiences of Pilots B and Pilot C.
For the PQ grading system, all pilots assigned *marginal* or *satisfactory* ratings to the MR and VR configurations. In contrast, the AR and CON configurations have a wider range of PQ scores, with some falling into the *unsatisfactory* category.
This indicates that there is room for improvement in all configurations in order to achieve the highest ratings of *excellent* or *very good*. The MR and VR configurations are closest to achieving this, but each configuration needs further development.

The **PQ scores for the hover maneuver** (Figure 3.5b) exhibit a slightly different trend compared to the slalom maneuver, with ratings across MR and VR configurations falling



(a) Slalom Maneuver **(b)** Hover Maneuver

Figure 3.5: PQ Total Scores (TS)

within the *marginal* to *satisfactory* range, consistent with previous results. However, the variability in scoring for AR and CON configurations is again more pronounced, with some scores falling into the *unacceptable* category.
Notably, the VR configuration receives higher scores than the MR configuration among Pilots A and B, contrasting the slalom results. The difference in perceived presence between the slalom and hover maneuvers may be attributed to the fact that the slalom maneuver involves fast-moving images, while the hover maneuver is characterized by a more static image. In addition, the distraction caused by chroma artefacts with the MR configuration can be stronger pronounced in fast-moving images and the constant change in the visual scene can exacerbate the distraction.

Further investigation of the impact of different configurations on presence is facilitated by examining the subscales (Section 3.1.3) presented in Figure 3.6.

The **involvement** subscale of the PQ assesses various dimensions of users' experiences within a VE, including perceived control over events, responsiveness of the VE to their actions as well as level of visual engagement and overall interaction with the VE.
The results show a clear trend supporting the results of the PQ TS, with MR and VR configurations receiving the highest involvement ratings (5.6 to 7.0) from Pilots B and C which is also supported by the pilots' comments that they felt more involved in the simulation in these configurations and that they were less aware they were sitting in a simulator.

The **natural** subscale assesses three key aspects of users' experiences within a VE: the naturalness of interactions, consistency of the VE with real-world experience and intuitive feel of controlling locomotion through the VE.

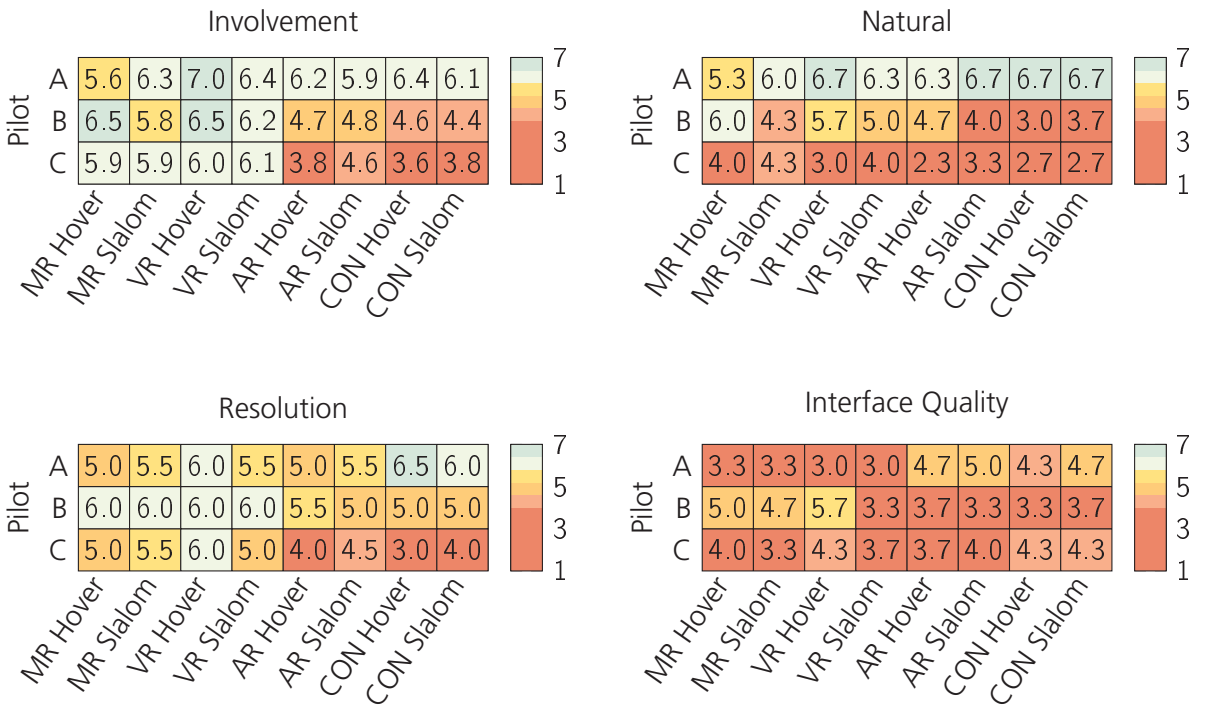


Figure 3.6: PQ Subscales

MR and VR configurations receive higher ratings, indicating that these configurations are perceived as more natural/immersive for Pilots B and C. However, the lower rating for natural compared to involvement and resolution ratings can be attributed to the fact that the pilots needed mor head movements to obey the VE than they would without the MR/AR headset. Additionally, artifacts due to chroma keying and reflections due to sub-optimal lighting can have a negative impact on the naturalness.

The **resolution** subscale evaluates the ability to closely examine objects and view them from multiple angles.
The results show that MR and VR configurations were overall assessed with higher scores, indicating that objects could be perceived more easily with the monitor-based headset's configuration. The AR configuration has a smaller FOV of 45°, which seems not sufficient for the pilots, as the external view is relevant for accomplishing the flight tasks. The TV screens (CON) do also not provide sufficient orientation in the VE, as the FOV is not adaptable like for the MR headset.

The **interface quality** subscale assesses two key aspects of users' experiences within a VE: the degree to which control and display devices facilitate or hinder task completion, and participants' ability to focus and concentrate on their tasks.
The results reveal poor ratings across all configurations, with many scores falling into the *unacceptable* range (< 4). Within the VR configuration, the pilots stated that the flight instruments were best readable, while the MR configuration caused the instrument to

appear blurred if increased head movement was involved. The AR configuration provided readable flight instruments, but the quality strongly depended on lighting conditions, as the display darkens the outside view. This highlights an area of concern that requires further optimization.

The sub-scales also allow for a more nuanced understanding of Pilot A's ratings. Notably, Pilot A's scores for involvement, natural and resolution are more or less constant across all configurations, while the inverse TS trend compared to the other pilots (Figure 3.5) is mainly caused by significantly lower ratings of the MR and VR in interface quality. As the pilot's feedback regarding presence did not significantly differ from the comments of Pilots B and C and Pilot A reported difficulties in understanding and completing the PQ, the validity of the pilot's ratings seems questionable.

Comparing the scores for both hover and slalom maneuvers between different configurations reveals no clear trend in the subscale scores regarding the flight maneuver.
Overall, the MR and VR configurations exhibit the highest potential for providing acceptable levels of presence in helicopter simulation. All configurations lack in interface quality, indicating to focus improvements in this area.

Pilots' Feedback on Presence

The pilots' feedback on presence aligns well with the PQ scores, demonstrating a distinct preference for certain configurations.

In particular, all pilots favored the **MR** configuration, as it offers a high degree of immersion and a realistic experience of the virtual world. The incorporation of real-world elements was a significant factor in their preference, as having their actual hands visible in the cockpit environment during MR simulation was particularly noted by the pilots as enhancing their overall presence in this configuration.

The **VR** configuration was marginally less preferred by the pilots compared to the MR configuration. They reported superior readability of instruments and fewer distracting artifacts during maneuvers relative to the MR configuration, which contributed positively to their experience. However, the pilots did report feeling disconnected from reality when utilizing the VR configuration. A contributing factor to this experience was the avatar (see embodiment in Section 1.1.3), suggesting that virtual representation of the pilots' bodies did not provide the same level of presence as the MR configuration.

Conversely, the **AR** configuration was deemed less immersive by all pilots due to the limited vertical FOV offered by the see-through display of the JVC headset. While the limited FOV is specific to the hardware being used and therefore not a characteristic limitation of AR technology, the compromised presence due to limited opacity of the VE poses a systematic deficiency of the technology. In addition, the pilots' arms are not visible in

certain areas of the VE, such as the window area, as the transparent area is statically fixed and limited to the cockpit interior. This restriction resulted in the pilots' perceiving the VE as less immersive.

Notably, the presence of the **CON** configuration was rated lowest among all configurations, with pilots uniformly reporting poor immersion or engagement during both maneuvers, due to its basic pilot station simulator setup featuring simple TV screens. The TV screen setup lacks depth perception, a critical aspect for immersive experiences. Furthermore, the individual TV displays' frames disrupted the outside view, creating visual clutter that impeded the overall experience.

The pilots' feedback on presence, in conjunction with the PQ scores, indicate that the MR configuration, closely followed by the VR, achieves the highest presence, suggesting that utilizing these configurations should have a positive impact on presence in helicopter simulation training.

Results of Simulation Sickness Questionnaire

Simulation sickness symptoms occurred rarely throughout the study. While Pilots A and B evaluated all configurations without encountering any symptoms, Pilot C experienced slight symptoms (*mild discomfort* and *slight dizziness (with their eyes open)*) on a single occasion, during evaluation of the MR configuration for the hover maneuver, resulting in a total score of 11 out of 236 possible score points. These symptoms occurred for a brief duration as Pilot C transitioned from low-speed flight to hover. The occurrence of unintentional rapid maneuvering caused the helicopter to roll to the side, resulting in fast-moving images of the surroundings displayed in the headset. Consequently, it is uncertain whether these symptoms were exclusively attributable to the MR configuration. It seems likely that similar symptoms could have been experienced in other configurations if Pilot C had performed the transition maneuver in a comparable manner. This observation is also consistent with the NATO guideline [22] that unexpected rapid movements can cause or exacerbate simulation illness symptoms (see predictability in Section 1.1.3).
Nonetheless, the experienced symptoms were transient and resolved quickly, thereby having negligible impact on the evaluation of the configuration's presence. Given that the pilots wore the headsets for approximately 15 to 20 minutes, further studies are warranted to investigate longer durations and assess the risk of simulation sickness developing over time, as prolonged exposure to VE exacerbates simulation sickness (see duration in Section 1.1.3).

3.2.2 Results of Flight Performance

The flight performance evaluation was based on performance parameters, which conform to the standards defined by ADS-33 [63], see Section 3.1.2. The primary objective was to compare deviations in flight performance achieved when using different visual configurations, such as MR, VR and AR, rather than evaluating the overall handling qualities (Section 3.1.3). A color-coding system was applied to match the performance assessment with the questionnaire rating scales: values within the green area indicate that the pilot met *desired* performance limits, data in the orange area represent achieving *adequate* limits and data in the dark red zone signify that the performance parameters were not met.

To provide an objective assessment of each pilot's best performance across different configurations, the highest score achieved in three consecutive runs is selected for a detailed evaluation.
For the comparative analysis in Section 3.6, where the results from Studies 1 and 2 are used to provide a comprehensive comparison of all configurations tested in the two distinct simulator environments of 2PASD and AVES, all individual runs are taken into account.

The results of the flight performance evaluation are presented in Table 3.7 for the slalom maneuver and in Table 3.8 for the hover maneuver. The tables list each pilot's flight performance for all MTE flight performance parameters across all tested configurations. For a comprehensive evaluation of the flight performance, Pilot B's results are discussed in detail. The analysis is conducted as described in Section 3.1.3.

Flight Performance of Slalom Maneuver

Table 3.7 lists each performance parameter (position (x,y), ground speed and altitude) of the slalom MTE for every pilot and configuration. For the slalom maneuver, no significant trend towards a specific configuration can be observed based on the overall flight performance assessment, as no color pattern is visible in the table.
Notably, all pilots achieved the same performance parameters in the CON configuration. The results indicate that all pilots experienced difficulties navigating through the gates when utilizing the TV displays as the visual system, resulting in falling into *not adequate* performance for the position.
Conversely, the pilots demonstrated more success flying through the gates achieving desired performance limits in the other configurations. In terms of ground speed performance, all pilots achieved either *adequate* or *not adequate* performance in all configurations. This can be attributed to the fact the maneuver was initialized with a ground speed of 60 kts, while falling below 60 kts will prevent from achieving *desired* performance.
The flight performance assessment for the slalom maneuver does not indicate a clear trend

favoring one specific configuration as the best option for a visual system in helicopter simulation training.

The flight path of Pilot B for the slalom task is depicted in on the right in Figure 3.7. Pilot B's results are analyzed in detail due to his limited experience with the ADS-33 MTEs, as he is not a test pilot and has lesser familiarity with all configurations. This selection minimizes potential biases in the evaluation as Pilot B should be more representative of the target group of pilot students compared to the Test Pilots A and C.
The top graph shows the helicopter's trajectory (x- and y-coordinates) through the slalom course, with the tracks for each configuration appearing similar at the beginning. The divergence of flight trajectories becomes apparent by the third turn, indicating that from this point onward, Pilot B faced increased challenges in maintaining ground speed and altitude within all configurations.
This trend is reflected in the middle graph, which shows variations in ground speed, and the bottom graph, which displays variations in altitude. The pilot described difficulties with orientation in the CON configuration, largely attributed to the limited view provided by the TV displays. Moreover, Pilot B noted that the lack of depth perception using the TV displays made it particularly difficult to estimate the helicopter's altitude and velocity accurately.

Table 3.7: Flight Performance of Slalom MTE in 2PASD

Config.	Parameter	Pilot A	Pilot B	Pilot C
MR	x,y	Desired	Desired	Not Adequate
	Ground Speed	Not Adequate	Adequate	Not Adequate
	Altitude	Desired	Desired	Desired
VR	x,y	Desired	Not Adequate	Desired
	Ground Speed	Adequate	Adequate	Adequate
	Altitude	Desired	Desired	Desired
AR	x,y	Desired	Not Adequate	Desired
	Ground Speed	Adequate	Not Adequate	Adequate
	Altitude	Desired	Desired	Desired
CON	x,y	Not Adequate	Not Adequate	Not adequate
	Ground Speed	Adequate	Adequate	Adequate
	Altitude	Desired	Desired	Desired

Flight Performance of Hover Maneuver

The flight performance of all pilots for each configuration during the hover maneuver is presented in Table 3.8.
All pilots achieved *adequate* or *desired* limits for the hover maneuver, except Pilot A, who struggled to maintain heading limits within both MR and CON configurations. Similarly, Pilot C faced challenges in maintaining heading with the CON configuration. The results show that, across all configurations, the pilots experienced difficulties in maintaining heading during the hover maneuver.
Similar to the results for the slalom maneuver, no clear trend is observed towards a single best configuration for achieving desired performance.

On the left side of Figure 3.7, the flight performance of Pilot B during the hover maneuver is illustrated, evaluating each configuration (MR, VR, AR, and CON) over a 30 second time period of maintaining hover at the designated hover point. The upper plots display x- and y-coordinates of Pilot B's flight position, while the third and fourth plot show deviation in altitude and heading, respectively.
For the MR, VR and AR configurations, Pilot B successfully maintained their position within desired limits throughout the hover maneuver. In contrast, with the CON configuration, Pilot B's performance was within adequate limits. Consistent ability to maintain desired altitude parameters across all four configurations (MR, VR, AR, and CON) is observed.
However, variability in heading within adequate limits is noted. Pilot B reported difficulties with maintaining orientation in the VE with the CON configuration, which is consistent with the findings for Pilots A and C.

The overall findings indicate that the CON configuration presents the most challenges for pilots in maintaining heading and orientation. The higher visual system latency, measured in Section 2.2.6, could have had a negative impact on performance, as for MTEs the pilots closely align their control inputs to the tasks' visual cues. However, the pilots did not explicitly report any problems with time delay. Also, no pilot-induced oscillations could be observed, which are a typical symptom of increased time delay. This suggests a limited impact of the system's latency.

Similar to the slalom maneuver, the hover task results do not show a clear trend in terms of which configuration, i.e. visual system, is most suitable for achieving desired performance.

Table 3.8: Flight Performance of Hover MTE in 2PASD

Config.	Parameter	Pilot A	Pilot B	Pilot C
MR	x Position	Adequate	Desired	Adequate
	y Position	Desired	Desired	Adequate
	Altitude	Adequate	Desired	Desired
	Heading	Not Adequate	Adequate	Desired
VR	x Position	Adequate	Desired	Desired
	y Position	Adequate	Desired	Adequate
	Altitude	Desired	Desired	Desired
	Heading	Adequate	Adequate	Adequate
AR	x Position	Desired	Desired	Adequate
	y Position	Adequate	Desired	Adequate
	Altitude	Desired	Desired	Desired
	Heading	Adequate	Adequate	Adequate
CON	x Position	Adequate	Adequate	Adequate
	y Position	Not Adequate	Desired	Desired
	Altitude	Desired	Desired	Adequate
	Heading	Not Adequate	Adequate	Not Adequate

Pilots' Feedback on Flight Performance

When using **MR and VR** configurations, pilots experienced brief blurring effects while moving their heads, increasing effort for changing gaze between ITC and OTW for checking flight instruments. This is consistent with the findings that motion blur contributes to visual degradation (Section 1.1.3).
The reduced FOV of the **AR** configuration was compared by pilots to wearing night vision goggles, requiring increased head movements to compensate for limited visual cues (see also FOV in Section 1.1.3).
The **CON** configuration with the TV display provides limited FOV visibility, which was a common issue among all pilots. This affects peripheral vision during the hover maneuver. The asymmetrical FOV due to off-center seating position in the side-by-side cockpit results in smaller horizontal FOV on the left-hand side, reducing visibility of visual cues on that side. In addition, the fixed FOV of the TV screen configuration contrasts with the MR headset, which provides a seemingly larger FOV by adapting the FOV to the orientation of the head.

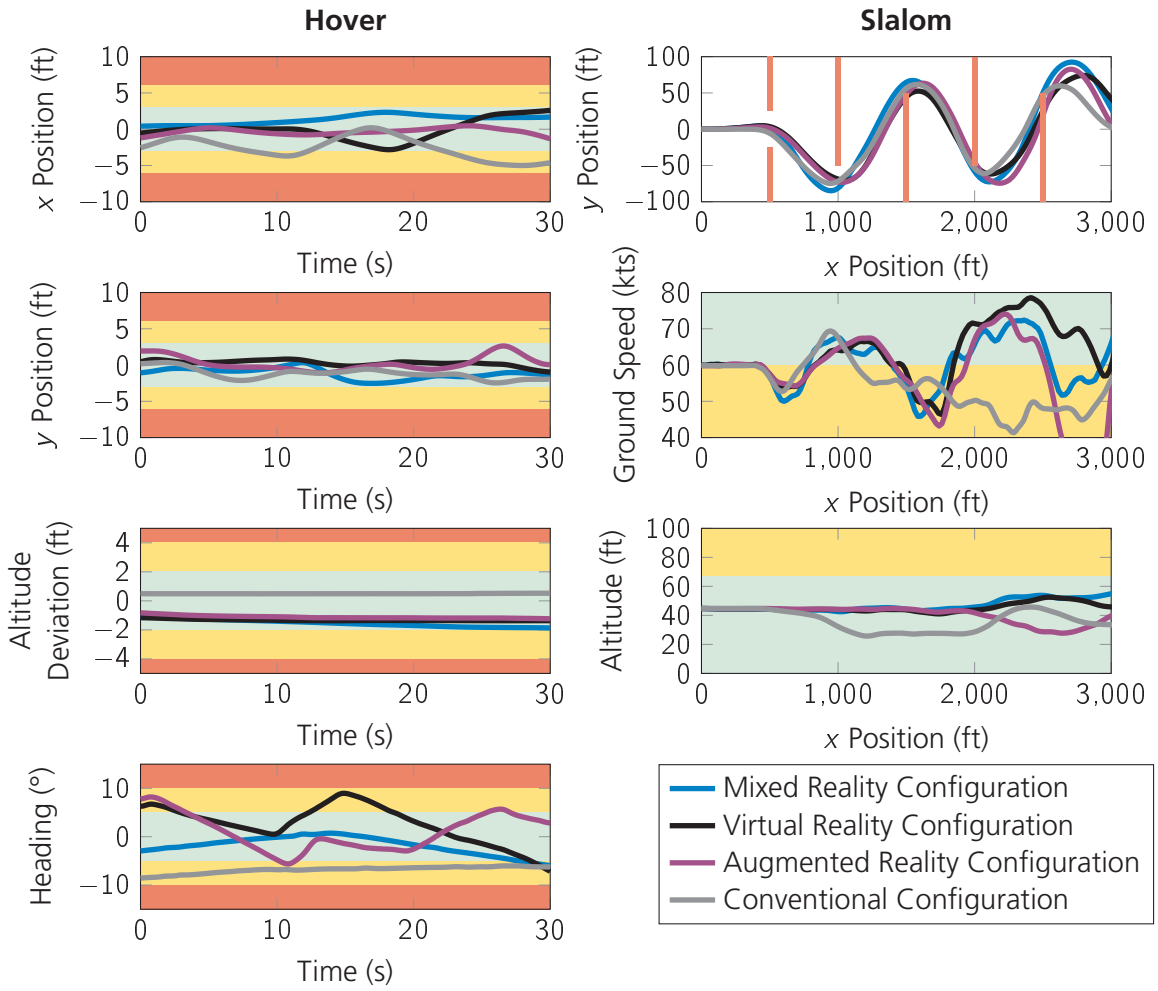


Figure 3.7: MTE Performance of Pilot B

3.2.3 Further Findings

Prior to presenting the condensed findings gathered for all tested configurations, an examination of the readability and comfort characteristics of each headset is conducted. During the debriefing session, pilots were asked to provide feedback regarding visual acuity and comfort levels experienced when using both monitor-based and see-through headsets.

Readability

The pilots were asked to read a visual acuity chart while wearing each one of the headsets on a standard A4-sized sheet with numbers in font sizes ranging from 10 points to 100 points. The results showed that both the monitor-based headset (Varjo) and the see-through headset (JVC) allowed the pilots to read all lines with ease.

In addition, the pilots were presented with a standard ICAO flight navigation map on a display mounted beside the primary cockpit displays. The pilots had to lean forward or read the display from a distance of approximately 75 cm. With the Varjo headset, it was challenging for the pilots to read radio frequency numbers printed on the map, as the image appeared slightly distorted. One of the pilots reported difficulties in reading blue marks and letters on the ICAO map and other information was partially illegible.
With the JVC headset, the ICAO map appeared more legible compared to the MR configuration using the Varjo headset.

Comfort

The pilots were asked to report their level of comfort while wearing the headsets. They reported no physical discomforts such as headaches, neck strain or eye fatigue. The weight of the headset was comparable to the helmet they wear during flight. However, the pilots noted that the headset might have had a greater impact if they had worn it for an extended duration, as the simulation study only lasted for approximately 15 to 20 minutes.

3.2.4 Conclusion of Study 1 and Selection of MR Headset for Further Studies

Based on the study's results and pilot feedback, the **MR** configuration appears to be the most promising visual system. Combining an MR headset with a helicopter mock-up has shown significant potential for helicopter pilot training, offering a highly immersive and realistic training experience. The inclusion of real elements can enhance pilots' sense of presence within the VE. The study's findings indicate that pilots' flight performance was comparable to the CON configuration.
However, the MR configuration faces some limitations, including image artifacts resulting from chroma keying, reflections on the PFD caused by additional light sources that can detract from the visual experience and image blurring during rapid head movements that can disrupt the MR experience.

The **VR** configuration also revealed potential benefits for helicopter pilot training when combined with a helicopter mock-up for haptic feedback. Moreover, the VR configuration exhibited fewer visual artifacts compared to MR as chroma keying is not needed. Challenges of the VR configuration for helicopter training include the necessity of rendering a detailed virtual cockpit with comprehensive flight instruments to provide a realistic simulation. Additionally, the disconnection from the real world can reduce immersion. Furthermore, accurate tracking of pilot movement necessitates the inclusion of hand or body tracking, in addition to head tracking, to ensure an immersive experience.

The **AR** configuration has shown some advantageous characteristics, including enhanced readability of flight instruments in well-lit environments and the preservation of pilots'

awareness of their surroundings. Notably, the virtual elements do not fully overlay the real world, maintaining a stronger (undesired) presence of the real environment. The see-through headset (JVC) used for the AR configuration provides a relatively small vertical FOV, which contributes to lower presence ratings and according to the pilots limits its suitability for conducting complex maneuvers compared to the Varjo headset.

Based on the results, the following approach was chosen:

The Varjo Headset was chosen for the following research work because the MR configuration established with the Varjo XR-3 headset was most preferred by the pilots. In addition, the configuration received high presence ratings and showed no significant deficiencies in flight performance compared to other configurations.

In order to effectively utilize simulation time for a comprehensive evaluation of the MR configuration in Study 2, the VR and AR configurations are no longer considered.
The results of Study 1 revealed some technical challenges associated with the MR configuration, including image artifacts from chroma keying, reflections on the PFD caused by additional lighting and blurred images resulting from rapid head movements.

For optimization of the MR configuration, masking/cutouts will be implemented in the cockpit region to minimize artifacts in this area caused by chroma keying.

Based on the findings from Study 1, the next step in this research work is to compare the MR visual setup to a flight simulator with a fully equipped cockpit and conventional projection system.
Therefore, the Varjo headset is integrated into the AVES in preparation of Study 2. The upcoming Section 3.3 describes the integration of the Varjo headset into the AVES and the optimization (Section 3.3) of the MR configuration through a hybrid approach using both cutouts/masking functions and chroma keying.
Section 3.4 then presents Study 2, including an analysis of cockpit interactions in a fully equipped cockpit for the MR configuration using the Varjo headset and the AVES visual system as baseline for comparison. The cockpit interaction is crucial for helicopter flight training as pilots need to learn to operate avionics systems, change radio frequencies and perform emergency procedures in the cockpit. The study aims to provide insights into the interaction dynamics that are essential for safe and efficient helicopter operations.

3.3 MR Integration Into Fully Equipped EC135 Helicopter Cockpit

For the integration of the Varjo XR-3 headset into the EC135 cockpit of the AVES (Section 2.1.1), the whole hardware (MR, head tracking stations, Mixed Reality Computer and light installation), was integrated in the cockpit. Two head-tracking stations were installed in the cockpit for the MR setup, one being located in the center above the cockpit panel and another at the top left next to the left cockpit door. In addition, the MRC computer used for rendering the MR image is integrated behind the pilot seats. Furthermore, additional lighting was installed to illuminate the cockpit interior, thereby enhancing the quality of the live image displayed within the headset.
Chroma keying shall be used for generating the MR environment, for that the dome is illuminated in green by the projection system (Figure 3.8). In regions where the color green is captured by the MR headset's cameras, the VE is displayed to the user. Everywhere else the real world captured by the headset's cameras is shown.
In addition to chroma keying, masking was applied to reduce artifacts and further increase MR image quality, see Section 3.3.



Figure 3.8: AVES EC135 Cockpit With Colored Projection Dome for MR Setup
©DLR. All rights reserved.

The MRC with the attached headset is connected to the AVES infrastructure. As the AVES simulation framework (Figure 3.9) is strongly related to the 2PASD framework (Figure 2.6), the integration effort was minimal. The datagram containing the aircraft state data is the

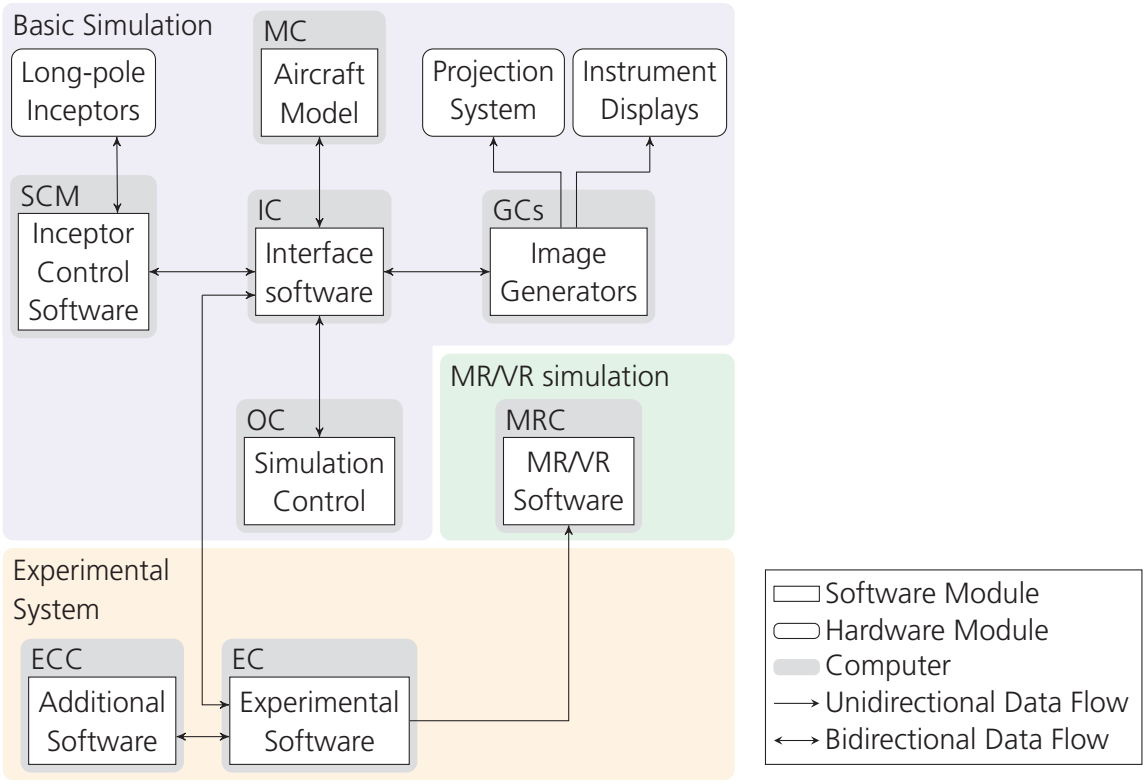


Figure 3.9: MR System Architecture in AVES

same for 2PASD and AVES, therefore no further modifications to the simulation network setup are necessary. The processing of the aircraft state datagram received by the MRC is the same as in 2PASD (Figure 2.7 in Section 2.2.2).

Optimization of MR Setup With Monitor-based Headset: A Hybrid Approach

For Study 1 conducted in 2PASD, chroma keying (Section 2.2.4) was used to create the MR environment. This was suitable for the initial study, as the presence of artifacts had little effect on the pilots' focus and without significant disruptions, the flight instruments remained accurately readable.
In Study 2, increased cockpit interaction shall be investigated, therefore the MR setup is further optimized. Firstly, the illumination setup is adapted by adding additional spotlights to prevent cast shadows. The other optimization is to add masks to the VE. Masking or cutouts define specific areas in the rendered image displayed in the MR headset, where the real world captured by the camera is displayed regardless of chroma keying. If this area is placed in the cockpit region, chroma keying artifacts are prevented in this area (Section 2.2.4).
In the fully equipped cockpit of the AVES EC135, green color is extensively used for flight instruments and the illumination of the cockpit. When using chroma keying only this



(a) AVES Mixed-Reality Without Masking
©DLR. CC BY-NC-ND

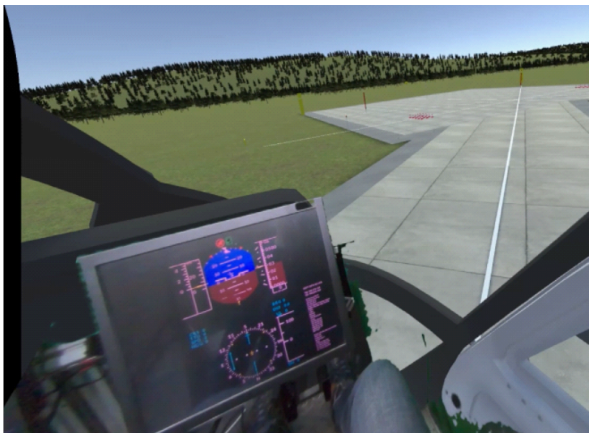


(b) AVES MR With Masking and Chroma Key
©DLR. CC BY-NC-ND

Figure 3.10: Combination of Chroma Keying and Masking for MR Environment in AVES

results in image artifacts, i.e. the rendered image shimmering through several segments of the cockpit, see Figure 3.10a. By combining masking and chroma keying, an increased MR image quality is accomplished, which could not be achieved by using exclusively one of the methods (Figure 3.10b,Figure 3.11).

Besides the AVES, the cockpit masks were created for the 2PASD as well. The inclusion of masking techniques leads to a significant reduction of artifacts and allows to precisely separate the display from the surrounding VE (Figure 3.10a).



(a) 2PASD MR Without Masking
©DLR. CC BY-NC-ND



(b) 2PASD MR With Masking
©DLR. CC BY-NC-ND

Figure 3.11: Combination of Chroma Keying and Masking for MR Environment in 2PASD

3.4 Study 2: Investigation of MR in a Fully Equipped Cockpit

This section provides an overview of Study 2 conducted in the AVES facility. Parts of this chapter have already been presented in earlier work by the author's publication in [90]. The focus of this study is comprised of two aspects.
Firstly, the study examines the pilot's interaction with the cockpit environment when using a MR headset in comparison to the conventional visual system (projection system), specifically during a malfunction procedure. The aim of this study is to understand the effect of using a MR headset as visual system on the pilot's interaction within a fully equipped helicopter cockpit.
Secondly, the MTE tasks previously used in Study 1 will be used in order to compare the different configurations of 2PASD and AVES. The purpose is to gain a comprehensive understanding of the usage of MR in various types of simulators.

3.4.1 Experimental Setup

In this study two visual systems are compared: the dome projection of AVES and the Varjo XR-3 Headset. The standard AVES configuration with the dome projection system serves as the baseline for comparing a conventional fully equipped simulator with a MR setup utilizing the Varjo headset (Table 3.9).

Table 3.9: Overview of Visual Systems

Technical Specifications	Varjo XR-3	Dome Projection
Visual System	Monitor-based	Projection
Field of View (h × v)	115° × 90°	240° × 90°
Refresh Rate	90 Hz	60 Hz
Weight	1320 g	-

The conventional visual system features nine high resolution projectors which span a FOV of 240° × 90° at a refresh rate of 60 Hz. The pilot view of the CON configuration is shown in Figure 3.12a captured by Tobii Pro Glasses 3 scene camera.

The MR environment is set up according to the description in Section 3.3. The projection dome is colored in green to allow for chroma keying, which is supplemented by cutouts in the cockpit area to improve overall image quality. This leads to the view in the MR headset as it shown in Figure 3.12b. The outside view displays the VE, where everything inside the cockpit is the real world captured by the MR headset's camera. The study setup enables a comparison between the dome projection system and MR headset visual



(a) AVES CON Setup ©DLR. CC BY-NC-ND

(b) AVES MR Setup ©DLR. CC BY-NC-ND

Figure 3.12: Visual System Configurations for Study 2

systems, allowing for a thorough evaluation of their respective capabilities and limitations in simulated helicopter flight.

3.4.2 Task Definition

Study 2 is comprised of three different flight tasks, see Table 3.10. For each pilot, the study begins with a familiarization flight, followed by each flight task first conducted with the MR configuration (Figure 3.12b) and afterwards with the conventional simulation setup with dome projection (Figure 3.12a).

Table 3.10: Overview of Tested Configurations

	Configuration	Flight Task
MR	Mixed Reality (Monitor-based Headset)	Familiarization Flight
MR	**Mixed Reality (Monitor-based Headset)**	**Malfunction Procedure**
MR	Mixed Reality (Monitor-based Headset)	Hover MTE
MR	Mixed Reality (Monitor-based Headset)	Slalom MTE
CON	**Conventional Simulator With Projection Dome**	**Malfunction Procedure**
CON	Conventional Simulator With Projection Dome	Hover MTE
CON	Conventional Simulator With Projection Dome	Slalom MTE

Familiarization Flight
Initially, a familiarization flight is performed to familiarize the pilots with the simulation environment and allow them to become accustomed to wearing the MR headset. Since the pilots were already familiar with the CON configuration of AVES and the time for the studies was limited, no familiarization flight was carried out with the CON configuration.

Malfunction Procedure
Within the Malfunction Procedure flight task, the pilots are asked to plan a cross-country flight with the GARMIN GNS430 flight navigation system. The GARMIN is located in the front area of the center console (see Figure 3.13, ①). The aim is to assess how well the pilot can interact with a typical cockpit system and also to evaluate the readability of displays. After completing the flight plan, the pilots are asked to depart from the airport and proceed according to their flight plan. En route during cruising flight, a generator disconnection (GEN DISCON) warning indication (Figure 3.13, ②) appears, triggering the pilots to execute the steps necessary according to the malfunction procedure to resolve the issue.

The GEN DISCON warning indication is related to a failure of the helicopter's electrical system and displayed on the CAD. The EC135 flight manual defines the following procedure to be executed in case of a generator disconnection event [91]:

1. Affected GEN switch - OFF
2. DC VOLTS, GEN AMPS and BAT AMPS - Check

If battery is discharged:

3. Electrical consumers - Reduce as much as possible
4. LAND SOON AS PRACTICAL

According to the procedure, the pilot's interaction consists of flipping off the switch of the affected generator (Figure 3.13, ③) and pressing the buttons (Figure 3.13, ⑤) of the VEMD (Figure 3.13, ④) on the right-hand side to change the displayed GEN field options to DC [V], GEN [AMPS] and BAT [AMPS] in order to monitor these parameters.

The simulated generator disconnection event causes the battery to discharge, preventing normal continuation of flight. Once the pilot assesses the situation, they are asked to decide how they wanted to proceed. Valid options could be to return to the airport of departure or continue the flight to the destination.
The scenario ends either with the landing at the departure airport or during cruise flight to the destination, depending on the pilot's decision. Due to time constraints, the flight to the destination airport is not completed.



Figure 3.13: Cockpit Instruments
©DLR. CC BY-NC-ND

MTE Flight Tasks

Following the Malfunction Procedure task, the pilots are asked to fly the MTE hover and slalom (see Section 3.1.2) as previously during Study 1. These tasks again were chosen to allow for comparison of visual system configurations evaluated in 2PASD during Study 1 and the new configurations evaluated in AVES.

3.4.3 Assessment Methods

For this study, the same assessment methods were used as in Study 1 (Section 3.1.3). To evaluate the presence of various system configurations, the PQ developed by Witmer and Singer was utilized. Simulation sickness symptoms were assessed using the SSQ from Kennedy. Additionally, flight performance was evaluated based on the ADS-33 criteria and eye tracking data was recorded for further analysis.

3.4.4 Pilot Experience

In Study 2 the same pilots participated as in Study 1 (Section 3.1.4). All of them were familiar with the AVES as well as see-through and monitor-based headsets from other studies. In addition, they were familiarized with the Varjo headset by the simulation runs conducted in Study 1.

3.5 Results of Study 2

In this section, the results of the pilot study conducted in AVES are presented. Initially, the subjective ratings provided by the pilots for the PQ and the SSQ are discussed for the malfunction procedure as well as for the MTE flight tasks. Subsequently, an analysis of the flight performance achieved by each pilot during hover and slalom maneuvers within the two configurations, MR and CON, is conducted.

3.5.1 Results of the Questionnaires

As in Study 1, for the assessment of presence and simulation sickness the questionnaires PQ and SSQ are used. The questionnaires are applied and evaluated as described in Section 3.1.3 for PQ and Section 3.1.3 for SSQ, respectively.
In the following sections, the questionnaire results are discussed for all conducted flight tasks, starting with the Malfunction Procedure.

Presence Questionnaire

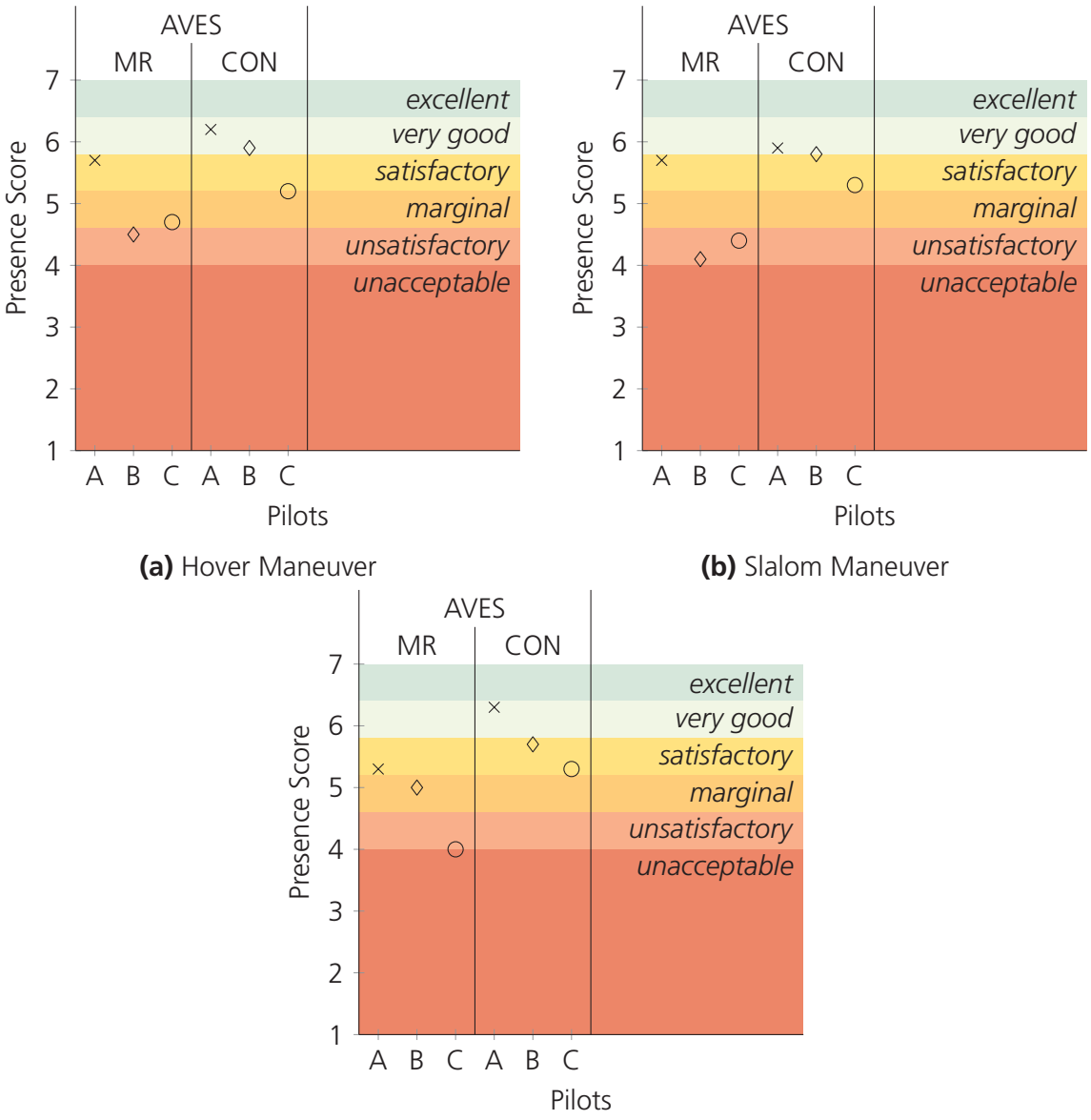
The total score for the PQ evaluation of the **malfunction procedure** performed for all three pilots is presented in Figure 3.14c. The results show, that the CON configuration resulted in higher presence ratings for all pilots. Each of the pilots rated the MR configuration at least one point lower than the conventional configuration, indicating a perceived lower presence with the MR configuration.
The pilots were able to perform the malfunction procedure with both configurations. They were able to operate the toggle switch for switching off the disconnected generator 1 without encountering any issues. It was also possible to toggle and read the displays for BAT, GEN and DC with both configurations. Visual navigation to the destination airport achievable as well, the pilots were able to determine their position using a channel and a highway as guiding lines.
The use of MR headsets as a visual system in the AVES, however, resulted in increased time required for pilots to enter destination airport coordinates into the navigation system, necessitating them to lean forward to read the information on its display. All pilots found direct visual access to the screen easier when not wearing the MR headset.
When the pilots were handed headphones to wear over the MR headset, the often missed the object while reaching out for it. This suggests that the MR headset impaired the pilots' depth perception and distance estimation, especially for objects close to them (see distortion and distance estimation error in Section 1.1.3). In contrast, the knobs of the flight instruments, which are usually about 70 cm away from the pilots, were operated with ease.

The PQ findings for the malfunction procedure are also reflected in Figure 3.14a and Figure 3.14b, which show the PQ ratings of all pilots for the **hover and slalom maneuvers**. A clear trend emerges, indicating that the presence is rated lower when using the MR headset not only for the malfunction procedure but also in general.

For a more detailed analysis, the PQ subscale results are discussed in detail (Figure 3.15):

The **involvement subscale** measures various aspects of a pilot's experience, including perceived control over events, visual engagement and interaction with the VE. The data supports the trend in the PQ total score results, showing that the CON configuration received higher involvement ratings (5.5 to 6.5) from the pilots than the MR configuration. This indicates that the pilots had a higher sense of control and involvement in the CON configuration. Interestingly, there were no significant differences in involvement ratings for all maneuvers in the CON configuration, regardless of whether the pilots interacted with the cockpit (malfunction procedure) or during the hand-on maneuvers (MTE tasks). This suggests that the CON configuration provides a consistent and engaging experience for pilots throughout the maneuvers. In contrast, the MR provides less consistent ratings, but at the same time does not show a trend across pilots.

The **natural** subscale assesses three key aspects, naturalness of interactions, correspondence of the VE to the real world and the intuitive feeling when controlling movement through the VE. In this study, the CON configuration received higher ratings, indicating that this configuration is perceived as more natural and immersive by all pilots. However, the pilots' customization and familiarity with the simulator play a significant role, as they are already more customized with the simulator's CON configuration than to using it with a MR headset. Furthermore, the restrictive setup in the narrow cockpit, including headtracking stations and closed cockpit windows, resulted in increased lagging of the images and artifacts, which could also affect the perceived naturalness of the experience (see also head tracking in Section 1.1.3). However, all pilots rated the naturalness for the malfunction procedure and the hover maneuver above 4 score points, which is still marginally acceptable. Nevertheless, it was shown that the slalom maneuver with the MR configuration was less awarded, demonstrating that fast moving images limit the perceived naturalness of the MR configuration.



(a) Hover Maneuver

(b) Slalom Maneuver

(c) Malfunction Procedure

Figure 3.14: PQ Total Score for all Three Maneuvers in AVES

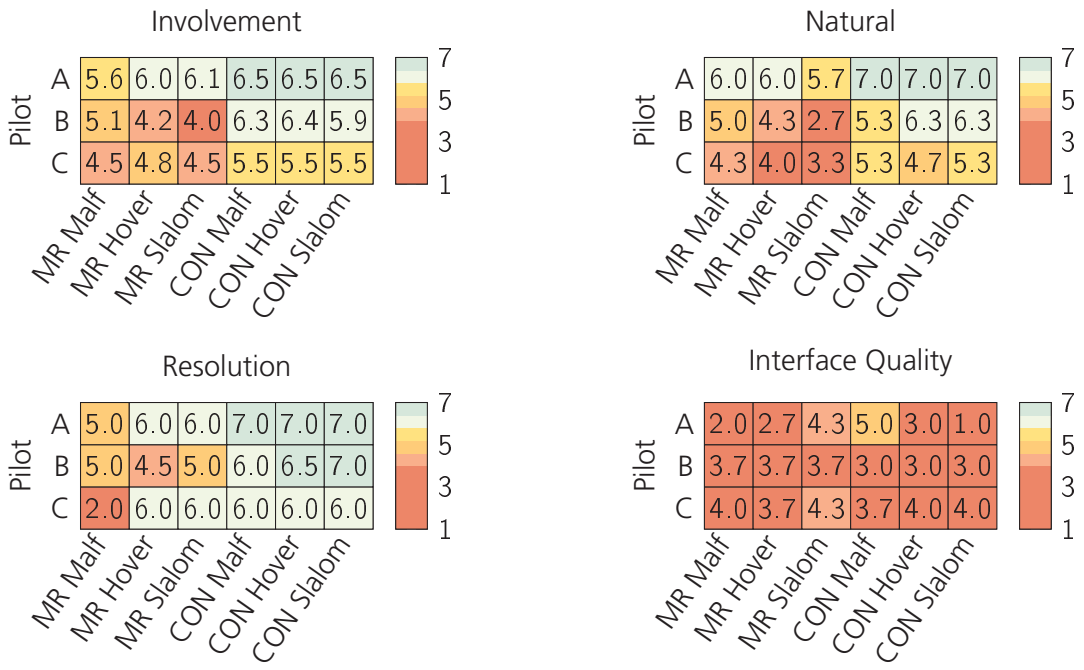


Figure 3.15: PQ Subscales

The **resolution** subscale evaluates the ability to accurately examine objects and view them from multiple angles. The results show that the conventional configuration received high ratings (6.0 to 7.0), indicating that objects were perceived easily with the projection system and within the real cockpit. The MR configuration was also awarded with high values of up to 6.0 by Pilots A and C, with the exception of the malfunction procedure. This was expected, as the malfunction procedure requires more work in the cockpit and the capturing of instruments and displaying them in the MR headset reduces resolution and legibility of the flight instruments.

The **interface quality** subscale assesses two key aspects of the user experience in a VE: the extent to which control and display devices facilitate or hinder the completion of tasks and the ability of participants to focus and concentrate on their tasks. The results display poor scores across all configurations, with many ratings falling into the unacceptable range. This highlights a problem area that requires further investigation and optimization, similar to Study 1 (Section 3.2.1). A possible cause for the poor results is that the pilots had difficulty applying the questionnaire's items to the helicopter simulation, as the questions were formulated in very general terms. This issue may be addressed through a more detailed briefing of the questionnaire.

Pilot Feedback on the Maneuvers

The results show that the pilots were able to perform the **malfunction procedure** with both the CON and MR configurations. However, the pilots preferred the CON configuration for several reasons.
First, with the CON configuration, the displays were readable more easily and therefore also the navigation system was easier to use for planning the route to the destination airport. Second, the pilots felt safer to use the route planning system and maintain a stable cruise flight with the CON configuration. Third, the peripheral vision of the CON configuration provided better situational awareness compared to the MR headset with its more limited FOV.

Two pilots chose to return to the take-off airport after the generator disconnect event, considering it their home base and knowing the personnel for the inspection was present. Pilot C stated that he did not use the navigation system for the return flight with the MR configuration, as they were familiar with the area and aware of his location. Adjusting the destination airport in the navigation system would be too time-consuming while wearing the MR headset. Without the headset, the two pilots easily changed the destination to the final airport.
The pilots also reported that they became more comfortable and confident in using the navigation system over time while in MR configuration. They found that their ability to enter data and use cockpit instruments improved as they became accustomed to the MR headset. This outcome aligns with the results of the study from [25] (see visual adaption in Section 1.1.3), in which people can adapt to a certain degree of visual displacement, but require more time for a task.

The pilots reported for the **hover and slalom maneuver**, that in MR configuration the flight instruments were displayed less sharp compared to CON configuration and blurred even more when they moved their heads. Additionally, they had to focus more to read numbers on the flight instruments or outside visual cues due to the limited FOV. The pilots were not able to detect movement of persons inside the cockpit even though they had a person sitting right next to them on the copilot seat. Usually, the pilots are aware of the movement of their crew member in their peripheral vision. However, with the MR headset, they needed additional head movement to scan their environment.
Especially in dynamic maneuvers, a fast scan of the flight instruments is crucial. The pilot stress that latency-free rendering is essential to perform the maneuver within limits, as any lagging would require corrective control inputs and increase risk of simulation sickness. They stated that they also minimized their gaze changes to reduce the risk of simulation sickness (see also head movements in Section 1.1.3).

The pilot feedback in combination with the PQ results clearly indicate that, not only for the malfunction procedure but also in general, presence is rated lower when using the MR headset.

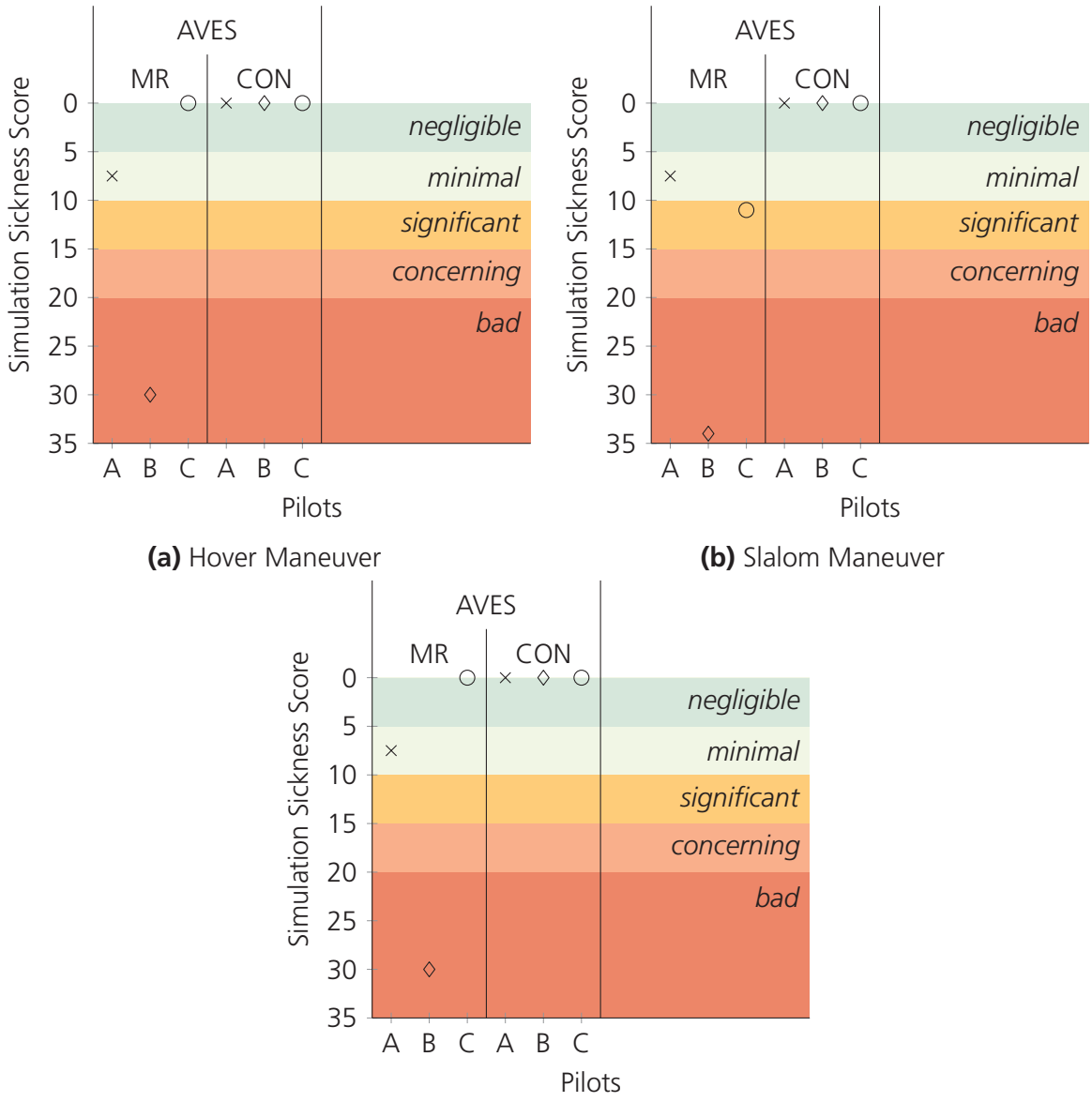
Simulation Sickness Questionnaire

The SSQ results for both configurations across all three maneuvers (Figure 3.16a, Figure 3.16b and Figure 3.16c) are shown in Figure 3.16. The pilots do not experience any simulation sickness results with the CON configuration. As all three pilots are well familiar with the AVES setup, they might be more accustomed to the CON configuration and less susceptible to simulation sickness symptoms.
In contrast, for the MR configuration, all pilots reported symptoms of simulation sickness. The severity of symptoms of Pilots A and C for the MR configuration ranges from no symptoms to *significant* based on [74]. The symptoms experienced by Pilot B are classified as *bad*, with a SSQ rating of 30 for the symptoms *slight fatigue*, *slight general discomfort*, *slight burred vision* and *slight dissiness (open eyes),* which were accompanied by neck pain while wearing the MR headset. Slight symptoms suffice for a configuration to be rates as *bad*, which, based on the rating system means, marks it as not acceptable for the simulation. However, it is worth noting that the initial calibration process for the SSQ scale from *negligible* to *bad* was conducted with military pilots, who may not be representative of the general population, as outlined in [92]. Therefore, the categorization scheme needs to be evaluated regarding its applicability for pilot students.

For a further in-depth analysis, the SSQ ratings are additionally evaluated using the subscales (Figure 3.17). The **nausea subscale** assesses gastrointestinal distress, including increased salivation, sweating, nausea, stomach awareness and burping. The **oculomotor subscale** evaluates visual and cognitive fatigue, such as fatigue, headache, eyestrain and difficulty focusing. Meanwhile, the **disorientation subscale** measures balance-related symptoms, including vertigo, dizziness with eyes open or closed and blurred vision.
As the susceptibility to simulation sickness varies between individuals, the pilots' results are discussed individually.

Pilot A showed no problems with symptoms associated with nausea (Figure 3.17). However, the pilot experiences difficulties focusing his eyes on certain information with the MR configuration, which is attributed to both the occumulator subscale and the disorientation subscale (Figure 3.17). The use of the MR headset causes the captured flight instruments and objects in the simulated environment to appear blurred and indistinct, so that the pilot must look closely to obtain the required visual information. In addition with the limited FOV, this leads to increased head movements, e.g. looking directly at the flight displays, as the could not be read in peripheral vision or a short scan, which can lead to general discomfort and eye fatigue (see also FOV in Section 1.1.3).

The higher ratings of the SSQ score within the subscales for **Pilot B** with the MR configuration can be partly attributed to difficulties with focusing. In addition, the pilot reported flickering images (see also screen flicker in Section 1.1.3), caused by issues with the head tracking setup. The image artifacts lead to dizziness and general discomfort and resulted in high ratings on the disorientation subscale for all flight tasks in MR configuration.



(a) Hover Maneuver

(b) Slalom Maneuver

(c) Malfunction Procedure

Figure 3.16: SSQ for all Three Maneuvers in AVES

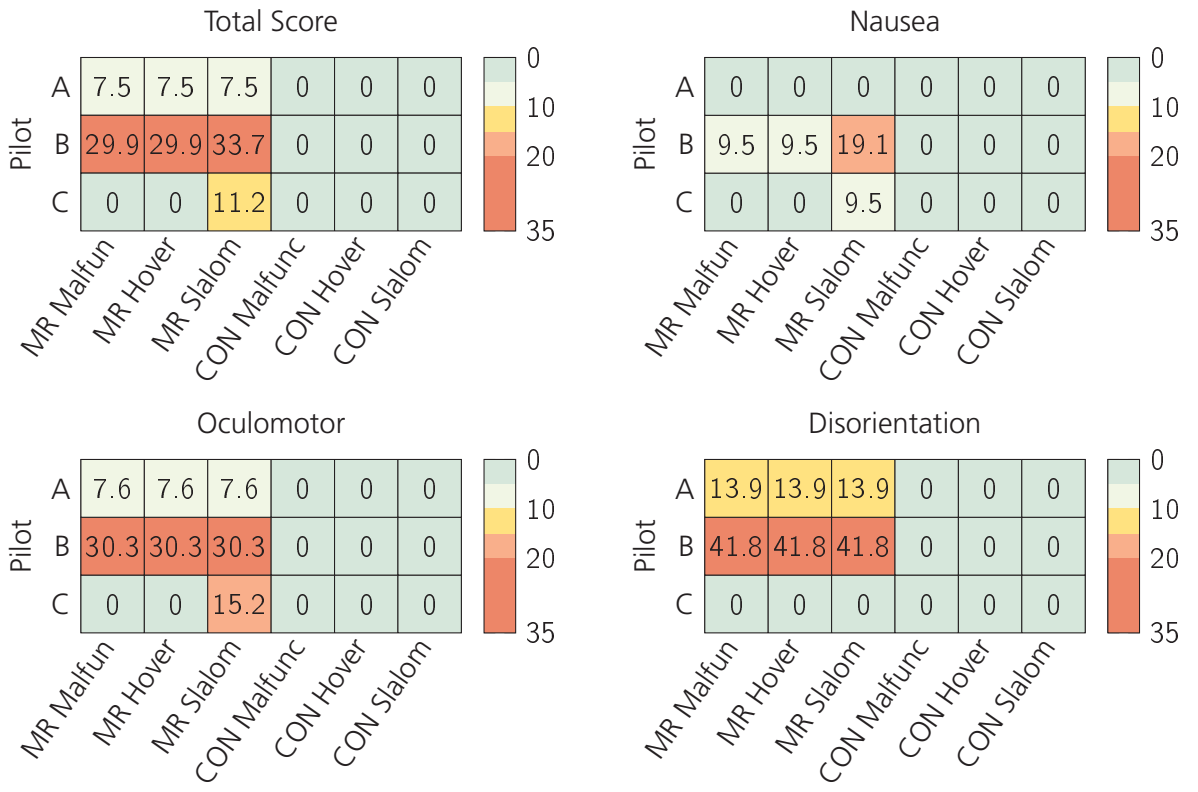


Figure 3.17: SSQ Subscales

The mild symptoms (Figure 3.17) of **Pilot C** are caused by a reboot event of the headset between the three runs of the slalom maneuver, resulting a blurred, distorted image displayed to the pilot for a brief moment. The symptoms faded quickly after the event and did not impose a strong effect on the pilot's well-being during the subsequent maneuver. This incident points out the importance of stable system operation to ensure that pilots' well-being is not negatively affected by system instabilities during training.

The symptoms reported by the individual pilots may be mitigated by optimization of the hardware and software setup. For instance, the placement of the tracking stations was limited by the compact cockpit, while sub-optimal lighting may have also affected the image quality of the camera.

In summary, the use of the MR headset configuration resulted in increased severity of simulation sickness symptoms compared to the CON configuration. A closer analysis of the MR data revealed that there is no significant difference in the occurrence of simulation sickness between the different conducted maneuvers. This suggests that the overall increase in symptoms is not due to specific maneuvers being more prone to simulation sickness, but rather a general effect of the MR headset configuration.

3.5.2 Flight Performance Results

In addition to the questionnaire results, the flight performance for slalom and hover maneuvers is evaluated. First, the performance of all pilots for the hover maneuver is analyzed. A detailed discussion of the results is then presented, with a focus on Pilot B's flight performance serving as a representative example. This is followed by a thorough examination of the slalom maneuver, where the performance of all pilots is assessed and a comprehensive discussion of the results is conducted, again with a focus placed on pilot B's flight performance as a reference point.

Flight Performance of Hover Maneuver

The performance parameters for the hover maneuver are listed in Table 3.11. As can be observed, each pilot achieved higher flight performance with the CON configuration compared to the MR configuration often achieving *desired* performance parameters. However, with the MR configuration pilots kept most performance parameters within *desired* and *adequate* limits.

Table 3.11: Flight Performance of Hover in AVES

Config.	Parameter	Pilot A	Pilot B	Pilot C
CON	x	Desired	Desired	Adequate
	y	Desired	Desired	Adequate
	Altitude	Desired	Desired	Desired
	Heading	Adequate	Desired	Desired
MR	x	Adequate	Adequate	Adequate
	y	Adequate	Desired	Desired
	Altitude	Desired	Adequate	Adequate
	Heading	Not Adequate	Adequate	Not Adequate

It should be noted that accurate head tracking is particularly important for this maneuver, as tracking errors will result in orientation errors of the image within the VE. A slight shift in the image can make a significant difference between *desired*, *adequate* or *not adequate* regarding the heading performance limits. By analyzing Pilot C's comments and virtual image recording, it was evident that the VE displayed in the MR headset was slightly misaligned in the heading angle axis, resulting in the pilot maintaining an offset heading angle and falling into *not adequate* performance parameter for heading.

In Figure 3.18 on the right, an example of the flight performance for hover maneuver of Pilot B is shown. In the two upper diagrams the x- and y-positions are presented for the duration of the 30 seconds of hover flight. The green color within the diagram indicates

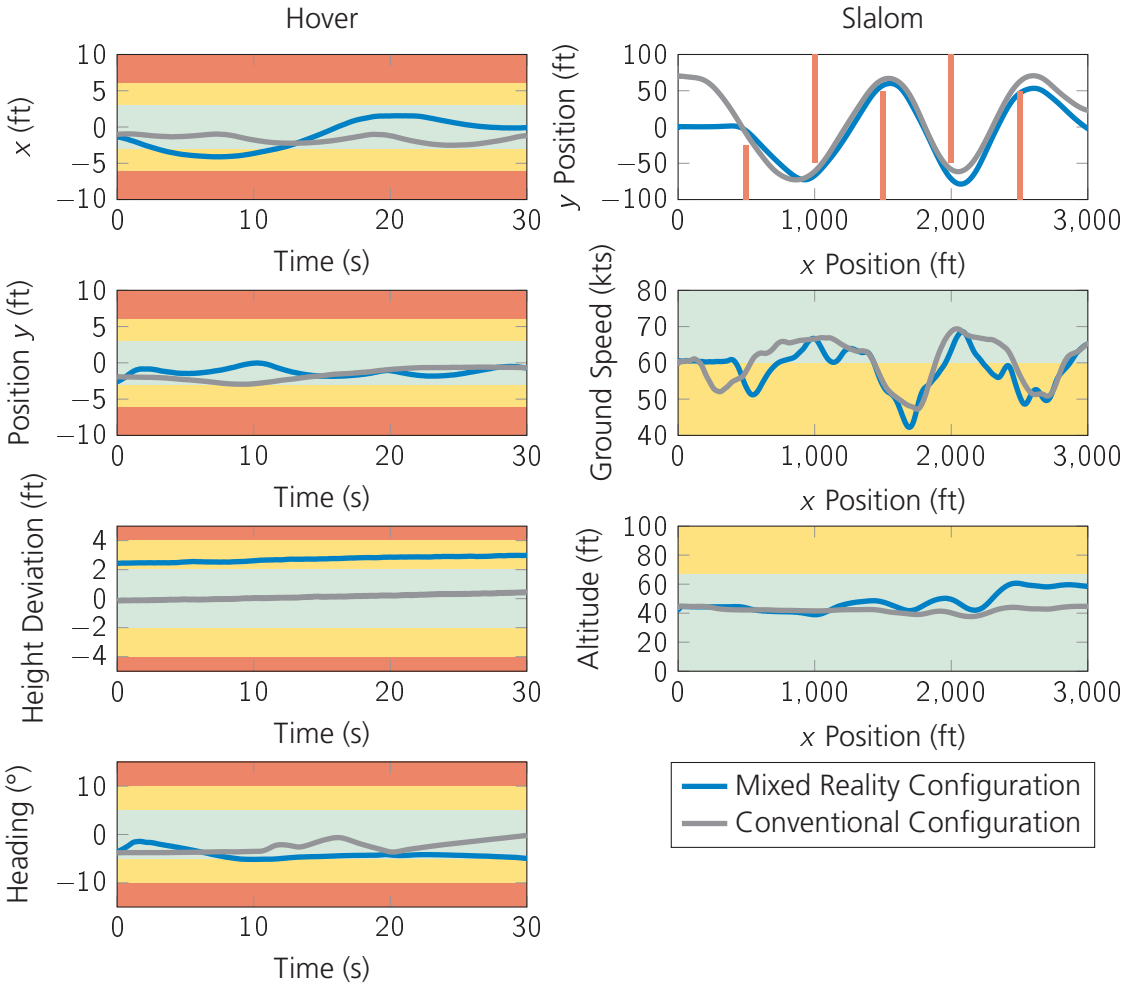


Figure 3.18: MTE Flight Performance in AVES of Pilot B

that the pilot's performance is within the *desired* performance limits. The yellow area represents *adequate* limits. For the MR configuration, the pilot maintains *adequate* limits for all parameters (x- and y position, height and heading (psi)). For the conventional configuration, the pilot meets *desired* performance for all parameters.

The flight performance data shows consistency with the findings gathered from the questionnaires. A correlation seems to exists between the pilots' evaluations of presence and their associated hover performances. When pilots rated their presence with the conventional configuration higher compared to the MR configuration, they demonstrated improved task performance for the hover maneuver. This potential correlation is further investigated in Section 3.6.2.

Flight Performance of Slalom Maneuver

The flight performance parameters for slalom maneuver is listed in Table 3.12. The table reveals that all pilots achieved adequate performance in the slalom maneuver with the CON configuration, with all pilots meeting *desired* performance for position and altitude parameters. However, they faced challenges maintaining speeds above 60 kts throughout the slalom course. The results indicate that pilots performed better with the conventional setup compared to the MR configuration, implying that visual disturbances such as flickering, blurred vision and lagging images may have negatively impacted not only on presence but also flight performance. One way to improve the experience would be to stream the cockpit instruments and navigation displays straight into the MR headset.
Figure 3.18 on the right side shows the flight trajectory of Pilot B for the slalom maneuver of Pilot B. The upper diagram shows the flight path. All pilots, including Pilot B, had difficulty navigating through the final gate within the MR configuration. They reported that flying with the conventional setup was easier because of the wider FOV. This allowed pilots to quickly scan their surroundings or recognize other cues in their peripheral vision. Pilot B maintained ground speed between 42 and 70 kts (middle diagram of Figure 3.18) achieving *adequate* performance. For both visual systems, Pilot B reached *desired* performance parameters regarding the altitude.

Table 3.12: Flight Performance of Slalom in AVES

Config.	Parameter	Pilot A	Pilot B	Pilot C
CON	x,y	Desired	Desired	Desired
	Ground Speed	Adequate	Adequate	Adequate
	Altitude	Desired	Desired	Desired
MR	x,y	Not Adequate	Not Adequate	Not Adequate
	Ground Speed	Adequate	Adequate	Adequate
	Altitude	Desired	Desired	Desired

Pilot Feedback on Flight Performance

During **hover maneuvers**, pilots reported several differences between using an MR headset and the conventional setup with projection system. The pilots noted that the MR headset provided accurate height estimations, matching the touchdown distance. However, they felt it was necessary to perform more head movements in the MR environment to maintain orientation and scan flight instruments. In contrast, the fully equipped cockpit with projection system allowed for a more fluent visualization of the outside view and also better readability of flight instruments, making it easier to perform the flight tasks and decreasing the amount of compensatory control inputs. The pilots also reported physical discomfort due to heat generated by illumination and the operation of the headset, despite cockpit ventilation operating during MR usage. In terms of overall immersion, the pilots subjectively perceived the MR environment as more immersive in the simulation scenario. Conversely, they felt more aware that they were sitting in a simulator cockpit when using the CON configuration.

During **slalom maneuvers**, the MR headset presented additional challenges: The pilots experienced lagging images and tracking issues with the MR headset, making it more difficult to provide smooth control inputs. Additionally, they reported orientation difficulties due to limited FOV and image latency. Furthermore, they had the impression of traveling at lower velocity compared to the AVES configuration.

In addition to their reports regarding the MR and CON evaluations, the two test pilots noted they had accumulated significant experience flying the MTE tasks in AVES. This extensive prior experience is likely to have influenced their task performance and preferences during the evaluation. Specifically, the pilots reported having a higher level of familiarity and proficiency with the CON configuration which may have contributed to their subjective assessment of its superiority over the MR configuration.

A comparative evaluation of all outcomes of Studies 1 and 2 is presented in the next section (Section 3.6), with the aim to develop a comprehensive understanding of the effectiveness of MR in helicopter simulation by comparing the results of each configuration within two different simulator environments (2PASD and AVES).

3.6 Comparative Analysis of Flight Performance for Both Studies

This chapter presents a comprehensive analysis of the results of all examined configurations on both simulators 2PASD and AVES for MTE flight tasks. First, a summary of all the tested configurations (MR, VR, AR and CON) is provided to give an overview. Subsequently, an analysis of the correlations between flight performance and subjective pilot ratings gathered using the PQ questionnaires is presented. For this purpose, a novel metric is developed to combine subjective results from the questionnaire with the objective flight performance results. By developing this novel metric, it can be investigated whether there is a correlation between the objective and subjective results in order to draw additional conclusions for an evaluation of the quality of the MR headset as a visual system in helicopter training simulation. The results of the SSQ were not analyzed for a correlation with flight performance, as few to no symptoms occurred during the flight in the simulation, so that the data were deemed not sufficient for such an analysis. Following the discussion of the correlation, the eye tracking evaluation with focus on the ITC/OTW ratio is presented for all configurations.

3.6.1 Overview of all Evaluated Configurations

For this work, six different configurations have been evaluated to assess the use of MR for helicopter simulation training (Figure 3.19) in both simulators 2PASD and AVES. The baseline for comparison are the **2PASD in CON configuration (Figure 3.19c)**, a fixed-based pilot station with TV Displays and the **AVES in CON configuration (Figure 3.19a)** with its projection dome. These two simulators were used to demonstrate the versatility of MR, enabling the following tested configurations:

MR configuration in AVES (Figure 3.19b): As with AVES in CON configuration, the fully equipped cockpit is available, only the OTW view is virtually displayed, while the inside of the cockpit is captured by the monitor-based headset's cameras.
MR configuration in 2PASD (Figure 3.19d): The 2PASD is a simple pilot station featuring flight instruments on the displays in front of the pilot to provide the flight information. Therefore, the VE fills a larger portion of the image, additionally displaying the helicopter structure. The display instruments, pilot seat and pilot themselves are captured by the monitor-based headset's camera.
VR configuration in 2PASD (Figure 3.19f): For the VR configuration the setup is fully virtual, with the pilot being replaced by an avatar.
AR configuration in 2PASD (Figure 3.19e): For the AR configuration, the see-through headset was used, which only displays the VE as an overlay to the real environment, with some areas of the display being see-through.



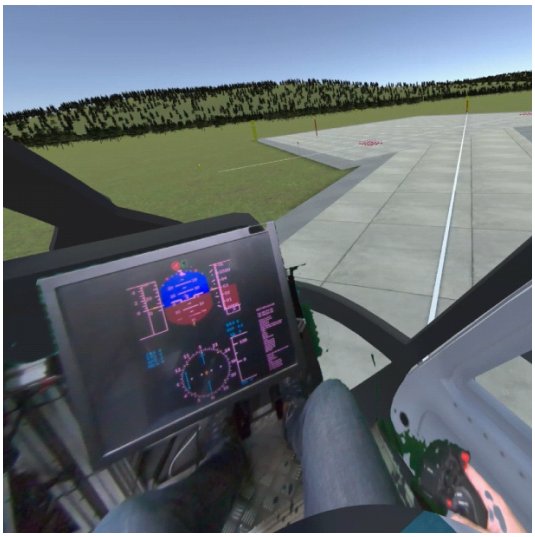
(a) AVES CON



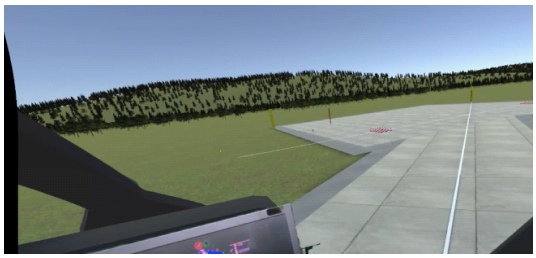
(b) AVES MR



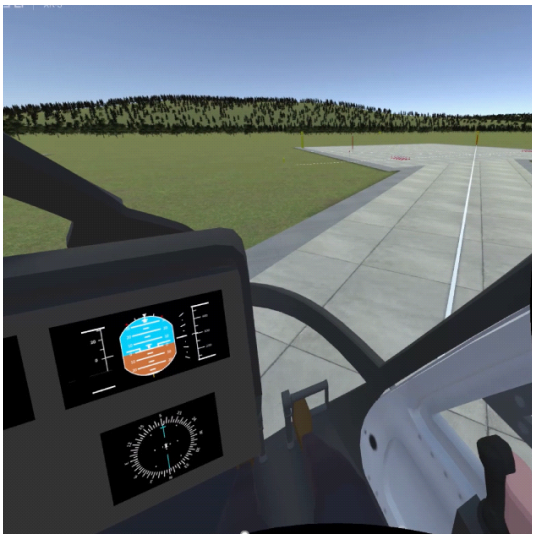
(c) 2PASD CON



(d) 2PASD MR



(e) 2PASD Representative for AR View



(f) 2PASD VR

Figure 3.19: All Tested Configurations ©DLR. CC BY-NC-ND

3.6.2 Evaluation of all Configurations Regarding PQ in Relation to Flight Performance

In this section a comparison of the results of the MTE tasks from Study 1 and Study 2 is presented for all configurations (Figure 3.19). It is investigated whether there is a correlation between the presence ratings and the pilots' flight performance. The MTE results (*desired*, *adequate* and *not adequate*) are converted to numerical values to allow a comparative analysis with the PQ. This approach enables an analysis of how presence affects pilots' flight performance, which in turn provides a more comprehensive understanding of the visual system.
A detailed examination of the correlation for the SSQ and the flight performance has not been carried out, as it is already apparent that there are too few non-zero points to determine a significant correlation as only few symptoms occurred during the two studies.

Correlation of Presence and Flight Performance

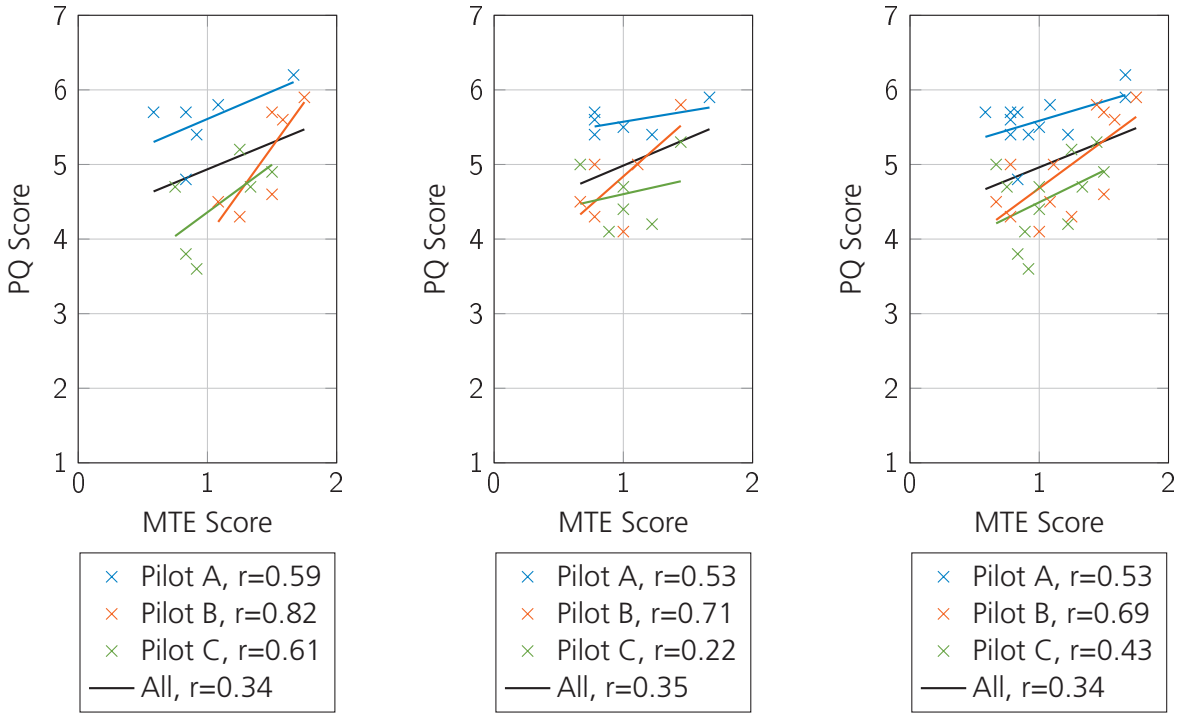
To compare the PQ ratings with the flight performance results, all flight performances are evaluated based on whether the performance parameters (position, height, heading, etc.) for *desired* and *adequate* were achieved. The flight performance boundaries are translated into numerical values to allow for further analysis: If the pilot did not comply with either *adequate* or *desired* performance limits, a zero (0) is assigned for the parameter. If the pilot complied with the *adequate* limits for the parameter, a one (1) is awarded. A two (2) is assigned for achieving the *desired* parameter. An average flight performance score is calculated from the different performance parameters for each run (position, altitude, heading). Since each pilot flew three times per configuration, the values are averaged again over the three runs. The final value obtained represents the overall flight performance achieved by each individual pilot for the respective simulator configuration. A mapping function is developed that allows for transferring the numerical flight performance rating to the presence score value range. This allows to investigate potential correlations between both metrics using a regression analysis.

Two approaches are employed to determine the relationship between the presence score (1 to 7) and the numerical MTE ratings (0 to 2), see (Table 3.13): plain linear mapping and piece-wise linear mapping of MTE scores to the presence score value range.
For the plain linear mapping, each rating scales minimum and maximum values are mapped to each other, resulting in *not adequate* (0) being mapped to PQ 1, *adequate* (1) being mapped to PQ 4 and *desired* (2) being mapped to PQ 7.
The boundaries for the piece-wise mapping are based on the categorization of acceptability adapted from [69], where presence above 70 % is deemed *acceptable*, above 50 % is classified as *marginal acceptable* and under 50 % is considered *not acceptable*. The results

Table 3.13: MTE Score and PQ Score Mapping

Flight Performance	MTE Score	PQ Score (Linear).	PQ Score (Piece-wise Linear)
Not adequate	0	1	4
Adequate	1	4	5.2
Desired	2	7	7

are mapped as following: *desired* corresponds to the presence score of seven (7), *adequate* to presence score of 5.2 and *not adequate* to PQ of one (1).
To evaluate both approaches, the Pearson Correlation Coefficient of each linear regression between the PQ score and the MTE score is calculated, providing insights into the relationship between presence and flight performance. The Pearson correlation coefficient provides a normalized measure for the linear correlation between two data sets by calculating the covariance of both sets and dividing the result by the product of both data sets' standard deviations. This yields a value range for the coefficient of -1 to $+1$. A value of ± 1 indicates a perfect linear correlation between two data sets, while a value of 0 implies no linear linear dependency between the data sets.
For the plain linear regression, the mapped values and regression lines are visualized in Figure 3.20. There are three diagrams in total, one diagram each for the hover (Figure 3.20a) and slalom maneuvers (Figure 3.20b) and a third diagram showing both maneuvers combined (Figure 3.20c).
The Pearson correlation coefficient r is calculated for each pilot using the six data points. In addition, an overall regression line and the corresponding r value are calculated across all pilots and configurations.
All diagrams show a trend towards a correlation between the PQ values and the MTE values. If the PQ value increases, the MTE value also increases. As can bee seen, the overall Correlation Coefficient r_A is low (0.34 to 0.35) compared to each individual pilot's correlation coefficient. This can be attributed to the fact that the ratings differ between pilots, visible as offsets and different regression line gradients in the diagrams. This difference between pilots is a well-known phenomenon in subjective ratings and referred to as inter-pilot variability. The diagram Figure 3.20 emphasizes this variability. The different regression lines (especially blue and green, but also red with a higher gradient) show similar trends, however, they are shifted differently in the PQ score (y-axis). Since this work does not focus on comparing pilots flight performance between each other, pilot bias is not eliminated. If inter-pilot variability was considered, normalization would typically be required to account for differences between pilots and ensure the accuracy of results.
The r values for the different maneuvers exhibit significant variability across pilots. Pilot A demonstrates a medium correlation coefficient in both hover and slalom, with values of r ranging from 0.53 to 0.59. Conversely, Pilot B displays a high correlation in hover (r = 0.82) and a good correlation in slalom (r = 0.71), with an overall correlation of r= 0.69 for both maneuvers. Pilot C, on the other hand, exhibits a good correlation in hover (r = 0.61)

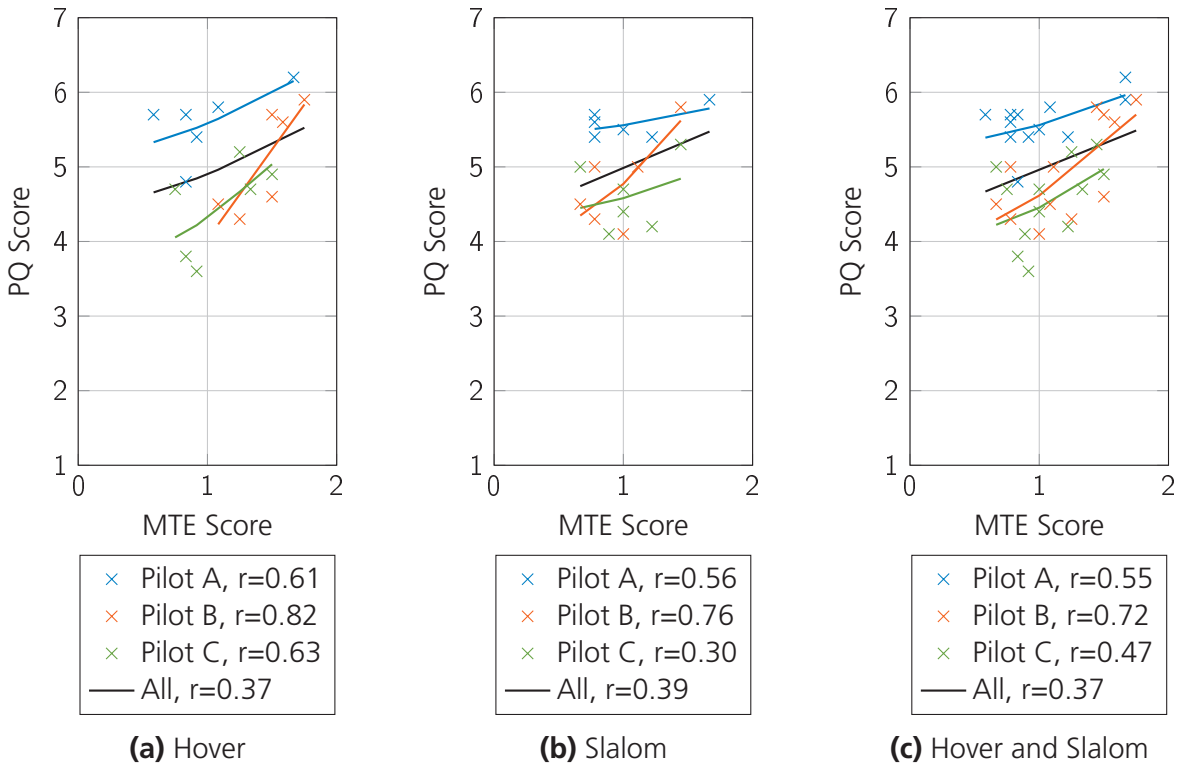


(a) Hover **(b)** Slalom **(c)** Hover and Slalom

Figure 3.20: Linear Regression

but a poor correlation in slalom (r = 0.22).
In addition to the plain linear regression, Figure 3.21 displays the correlation for piece-wise linear regression. The regression lines exhibit a discontinuity, resulting from the intentional application of piece-wise linear mapping in this analysis. Pilot B shows the highest correlation between presence ratings and flight performance in both the hover (Figure 3.21a) and slalom maneuvers (Figure 3.21b). r for Pilot B is 0.82 for hover, 0.76 for slalom and 0.72 for both maneuvers combined (Figure 3.21c). In contrast, Pilot A's r's ranges between 0.55 and 0.61, indicating a weaker relationship between presence ratings and flight performance. With Pilot C, again larger differences between both maneuvers can be observed regarding the correlation, with r = 0.63 for the hover and r = 0.30 for the slalom maneuver equally. As for the plain linear regression, the variability of ratings between pilots is observed, resulting in overall r falling into the range 0.37 to 0.39.

In order to compare the correlation coefficients for both mapping approaches, the complete data are presented in Table 3.14. The table shows that the plain linear mapping exhibits lower correlation across all pilots and configurations, while the piece-wise linear mapping based on the categorization of acceptability displays a higher correlation. This outcome suggests that the inclusion of acceptance categorization in the mapping is beneficial for the combined assessment of presence and flight performance.



(a) Hover **(b)** Slalom **(c)** Hover and Slalom

Figure 3.21: Piece-wise Linear Regression

Table 3.14: Comparison of Correlation Coefficients Between Mapped MTE Scores and PQ Scores for Different Regression Approaches

Mapping Method	Pilots	r of Hover	r of Slalom	r of Hover + Slalom
Plain Linear Mapping (LM)	A	0.61	0.56	0.55
	B	0.82	0.71	0.69
	C	0.61	0.22	0.43
	All	0.34	0.35	0.34
Piece-wise Linear Mapping (PWM)	A	0.61	0.56	0.55
	B	0.82	0.76	0.72
	C	0.63	0.30	0.47
	All	0.37	0.39	0.37
Difference of the Mapping Methods (PWM - LM)	A	0.02	0.03	0.02
	B	0.00	0.05	0.03
	C	0.02	0.08	0.04
	All	0.03	0.04	0.03

Analysis of Presence and Correlated MTE Ratings

For the comparative evaluation of presence across all configurations, both the PQ ratings and the correlated MTE ratings are analyzed. For the correlation of MTE to PQ scores, the regression with piece-wise linear mapping is applied.
The overall results are displayed in Figure 3.22 for the hover maneuver and Figure 3.23 for the slalom maneuver. PQ scores are displayed in black, while the MTE Scores are marked in blue.

An examination of PQ and MTE scores show that the **AVES CON** configuration was assessed with a higher PQ scores than the **AVES MR** configuration. Moreover, the pilots adhere better to the specified task limits, thereby supporting the assumption of a correlation between the PQ and MTE scores.
For the various 2PASD configurations, this trend is not as evident as for AVES. For Pilots B and C, the PQ scores exhibit a similar gradient over different configurations as the MTE scores, indicating a correlation. For Pilot A, the gradients over the configurations does not align between PQ and MTE scores. However, the validity of Pilot A's presence scores has been questioned, as they contradict the results of the other pilots and Pilot A reported difficulties completing the PQ (Section 3.2.1). Based on this observation, the MTE task performance might also be used to validate presence scores reported by the PQ, utilizing the observed correlation between the two metric.

Presence and Flight Performance Results for Hover Maneuver

The AVES CON configuration generally shows the highest presence and best flight performance for the hover maneuver. The MTE score is within *adequate* performance parameters (average of all three runs) for all pilots and the PQ score falls into the range from *satisfactory* to *very good*. These results support the pilots' comments that from all configuration the fully equipped cockpit is the most preferred one. According to the pilots the reason for this is the large FOV of the projection dome, providing a higher level of awareness of the helicopter orientation (Section 3.5) and co-pilot's actions. In addition, wearing no headset attributes to higher comfort.
The MR configuration's performance in AVES was limited due to the closed cockpit which caused head tracking issues and artifacts caused by reflections on the cockpit windshield. However, the PQ scores of AVES MR and 2PASD MR/VR configurations all fall within *marginal* to *excellent* except for a single rating of Pilot B. This indicates that both MR and VR configurations show the potential to close the gap with the AVES CON configuration, when extrapolating the current state and rate of development of MR/VR technology.
The 2PASD AR and CON ratings have reached *unsatisfactory* and *unacceptable* levels in certain cases, particularly affecting Pilots B and C. Nonetheless, the pilots were able to comply with the task limits much better than the presence ratings would suggest.

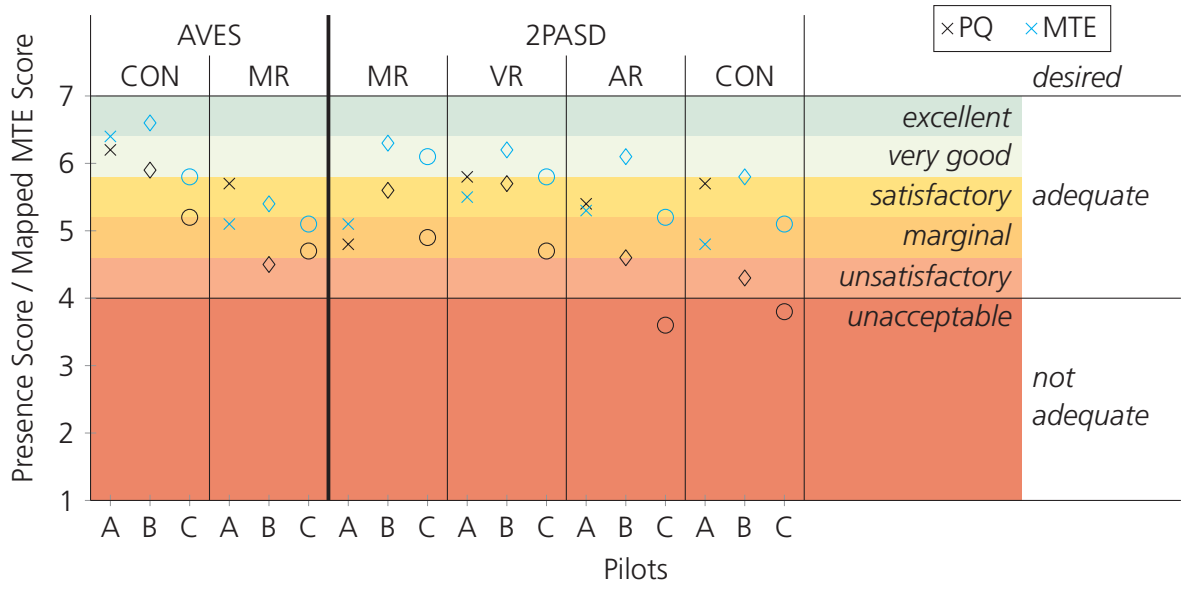


Figure 3.22: PQ and Mapped MTE Scores for Hover Maneuver

Presence and Flight Performance Results for Slalom Maneuver

For the slalom task, a correlation between flight performances and presence is recognizable as for the hover task, see Figure 3.23. The AVES CON configuration again is rated with higher presence scores than the AVES MR configuration. This correlates with the flight performance where the MTE score is higher for the CON configuration.
Regarding the MTE scores, the pilots achieve the highest ratings for the AVES CON configuration with the scores lying within *adequate* parameter limits. As the pilots are more familiar with the AVES long pole control inceptors and overall simulation environment than the 2PASD with its side sticks, here also training and customization may have an influence on the results.
The presence scores as well as the MTE scores of the 2PASD MR and VR configurations fall, as for the hover task, in the *marginal* to *satisfactory* categories supporting the argument that these configurations show potential for pilot helicopter training.
The scores of the AR and CON configurations in 2PASD also are similar to the hover maneuver results (*unsatisfactory* to *satisfactory*), making these two configurations the least suitable in comparison. For the AR configuration, the most frequently commented issues are the small FOV and the real world shimmering through the VE. With higher opacity of the display and an increased FOV, the configurations potential for use in flight simulation could be increased.

The MTE scores achieved for the slalom maneuver are generally lower compared to the hover task. This may be explained by the initial flight condition: in the slalom runs the simulation was started with the helicopter traveling at 60 kts of air speed. The boundary for desired task performance is to maintain an air speed of 60 kts or above. As a consequence, a slight deceleration would result in lower task performance. Starting with an increased

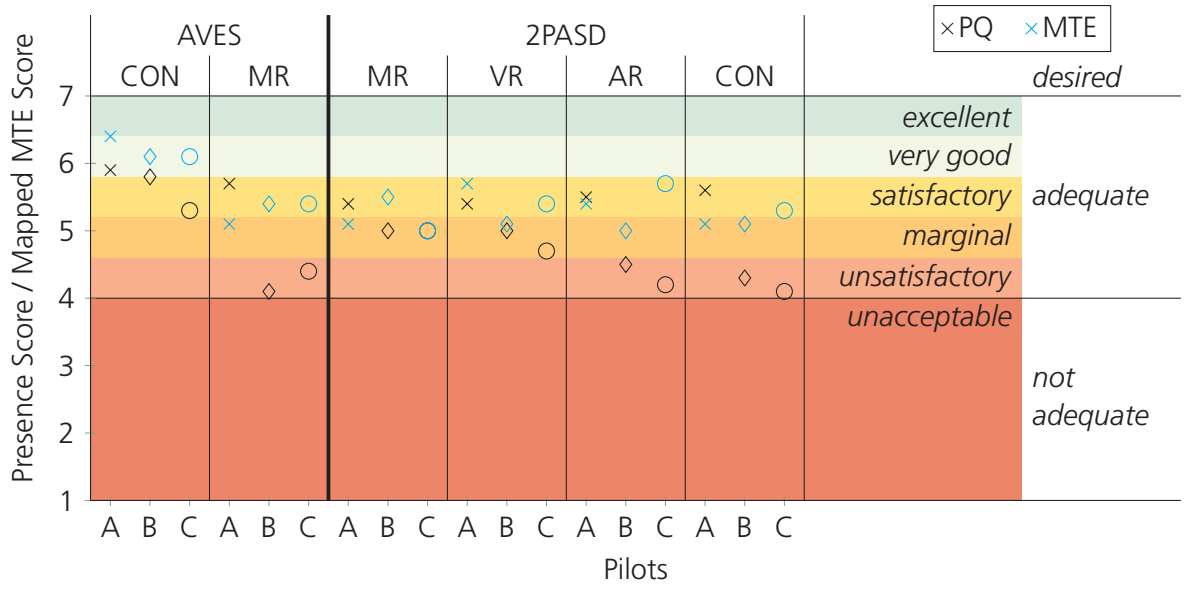


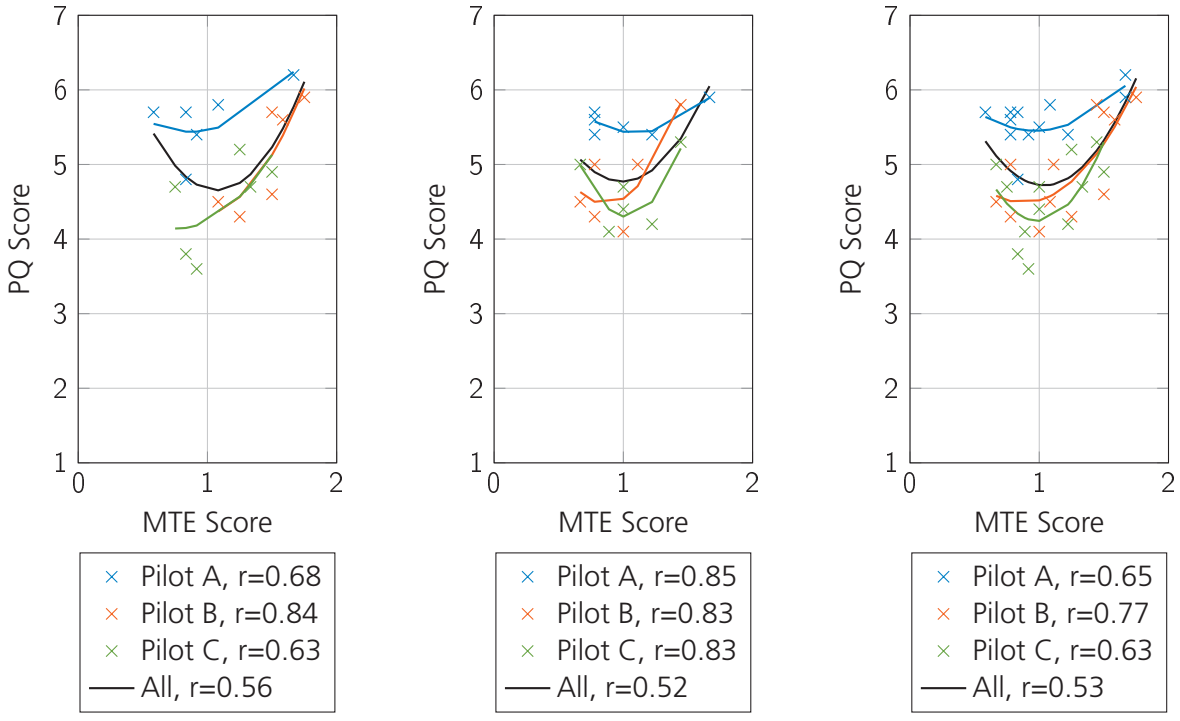
Figure 3.23: PQ and Mapped MTE Scores for Slalom Maneuver

initial velocity might have resulted in higher MTE scores as it would have decreased the workload in accurately maintaining air speed.

The correlation between presence and flight performance should be further investigated to determine whether the visual display system is effective for training helicopter pilots in simulation.
A comprehensive approach that incorporates both subjective experiences (captured through questionnaires) and objective data on flight conditions will facilitate a more nuanced understanding of the relationship between presence and flight performance.

Outlook for the Correlation Analyses

Besides the plain linear mapping and the piece-wise linear mapping presented in the sections above, a third mapping function was evaluated for analyzing a potential correlation between MTE flight performance scores and presence ratings. This was based on the observation that the presence ratings seem to not monotonically increase with MTE score. Therefore, a second order polynomial regression was applied to the data, see Figure 3.24. The plots show that for most datasets the regression exhibits a distinct minimum of presence at a MTE rating of approximately 1. The regression does also generally yield higher correlation coefficients compared to the plain linear and piece-wise linear mapping. As there are only few data points (six per pilot and flight task), a higher correlation coefficient may also be attributed to over-fitting as the degree of regression function is increased. However, the regression functions' shapes with the minimum occurring approximately at the same MTE score throughout the datasets does indicate a general trend.



(a) Hover **(b)** Slalom **(c)** Hover and Slalom

Figure 3.24: 2nd Degree Polynomial

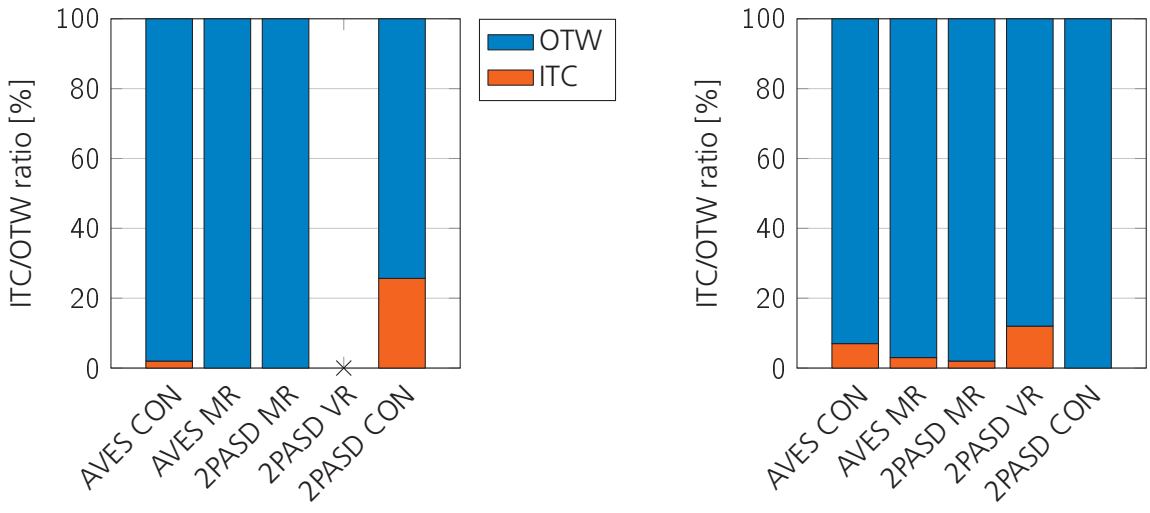
The regression indicates that minimum presence occurs for *adequate* task performance and that there is an increase in presence for decreasing as well as increasing task performance. One could conclude that increasing presence by increasing image quality may have a negative impact on task performance. High image quality can be characterized for example by high resolution, high level of detail and large FOV. However, high image quality can affect flight performance as it presents a large amount of information at once and pilots have to process this accordingly, resulting in a higher workload (see visual realism in Section 1.1.3). At the same time, small inaccuracies, such as the movement of grass not matching wind direction or rotor downwash of the helicopter, can distract the pilot. Such inaccuracies may explain why higher presence ratings could lead to a decrease in flight performance. Further research could be conducted to determine the ideal image quality for MR, which could potentially provide insights into the necessary quality of images for MR.

Regarding simulation sickness, only few symptoms occurred during the conducted studies. To clarify whether simulator sickness influences flight performance, a more comprehensive study is necessary. Longer flight maneuvers, which account for the impact of time on symptom development, should be examined. Moreover, developing a universal score that enables comparison of subjective and objective data on flight performance in relation to presence and simulation sickness would be beneficial and could be taken into account for future research.

3.6.3 Eye Tracking Evaluation

For the eye tracking evaluation, the ITC/OTW ratio is chosen. As described in Section 3.1.3, this metric is suitable for providing first insights on the gaze behavior of the pilots within the different configurations and respective visual systems. For this evaluation, the Focus Distance Distribution (FDD) method was applied as, in contrast to other approaches, it takes the error of the measurement system into account, see Section 4.4.5.

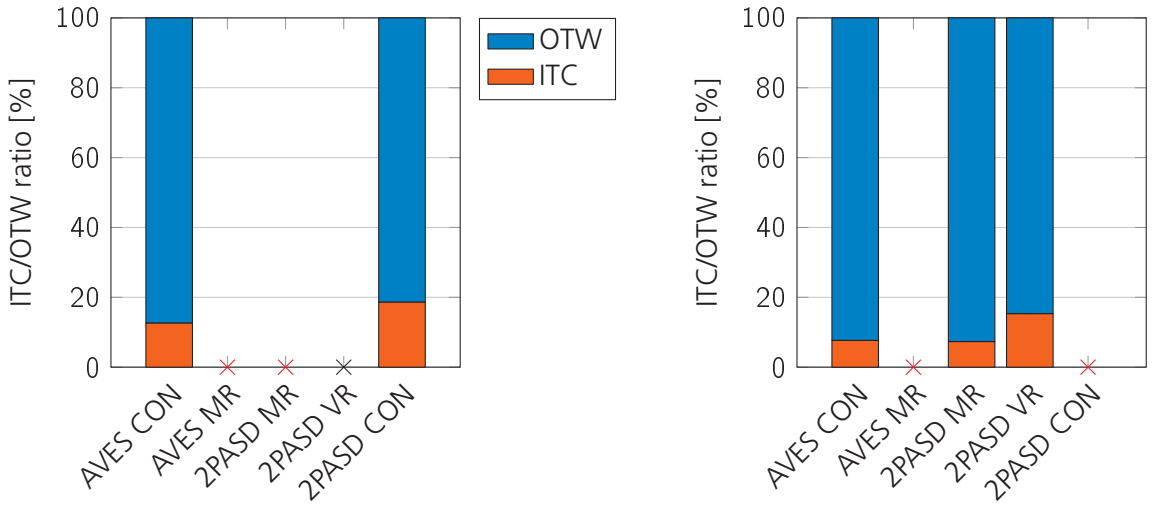
The ITC/OTW ratios derived from the data recorded for hover and slalom maneuvers are plotted in Figure 3.25 and Figure 3.26, respectively. The data only cover Pilots A and C, as Pilot B is wearing spectacles which are not compatible with the eye tracking system of the Varjo headset and additional lenses for the Tobii Glasses with different visual acuities were not available during the studies.
For some of the simulation runs of Pilots A and C, no eye tracking data could be recorded although the measurement systems' hardware and software was properly set up. The missing data is indicated by ×. Furthermore, some recorded data inhibited issues from inadequate calibration, as evidenced by data showing 100 % ITC for a simulation run, contradicting with pilot comments stating they primarily looked out the window during the flight tasks. These erroneous datasets (indicated by ×) were excluded from analysis as the necessary correction of the recorded gaze distance could not be reliably determined. This issue highlights the need to conduct calibrations with multiple points ITC and OTW before each simulation run to correct gaze distance offsets in the data during post processing.
For the remaining valid eye tracking data, the resulting ITC/OTW ratio was calculated as the mean of the three runs per pilot and configuration. The data for the hover task is plotted in Figure 3.25 for Pilot A (Figure 3.25a) and C (Figure 3.25b). The flight task requires looking OTW for visual cues, leading to predictable results: pilots focus primarily OTW. The pilots



(a) Pilot A **(b)** Pilot C

Figure 3.25: ITC/OTW Ratios for Hover Task

reported that they only glimpse, if necessary, to the flight instruments as they can receive all information to maintain a stable hover position from the MTEs visual cues.
Comparison of the different configurations indicates that pilots do look ITC the least for the MR configurations. This aligns with the pilots comments that the instruments appear blurred when they move their head, taking additional time to focus and read instruments. Furthermore, additional head movement was necessary to read the cockpit instruments compared to other configurations due to the limited FOV. With the cockpit instruments streamed into the MR headset (VR configuration), the instruments' image quality is increased and a higher ITC/OTW ratio is observed for Pilot C.
The ITC/OTW ratios for the slalom task are depicted in Figure 3.26 for Pilot A (Figure 3.26a) and C (Figure 3.26b).



(a) Pilot A **(b)** Pilot C

Figure 3.26: ITC/OTW Ratios for Slalom Task

Due to the recording and calibration issues, half of the data points are missing, prohibiting to investigate differences between CON and MR/VR for Pilot A. For Pilot C (Figure 3.26b) however, again the ITC/OTW ratio increased for VR compared to MR. The decrease of ITC for MR indicates that overall image quality affects pilot gaze behavior. Therefore, characteristics as FOV, image blur and focusing time and their relevance for the specific flight task need to be taken into account when applying MR technology for flight simulation.

3.7 Specific Implementations of MR in Helicopter Simulation

This chapter presents the results of the investigation into the implementation of MR and AR headsets in two distinct helicopter simulator environments. The following sections aim to highlight setup characteristics of MR and the implications of using MR in the two different simulation environments of 2PASD and AVES. In addition, the impacts of using monitor-based and see-trough headsets for helicopter simulation shall be discussed.
Based on the findings, key considerations for implementing MR in helicopter simulation are derived. They are categorized into use case considerations on the one hand and implementation considerations on the other hand.
The categorization and discussion of the considerations lead to the answering of Research Question (1.1).
The MR headsets used in research studies represent the current state of technology at that time. However, as MR technology continues to advance with improvements in display technology, tracking systems and software capabilities, the overall user experience can be significantly enhanced.

3.7.1 Setup Characteristics of the Evaluated VR, MR and AR Configurations

In this work a common simulation setup was created, that has been used for both the monitor-based and the see-through headsets to implement all tested visual system configurations. The setup includes head tracking, additional light sources and a high-performance computer for 3D rendering.
The head-tracking system ensures that the virtual environment is updated in real-time as the pilot moves their head, maintaining an accurate representation of their surroundings in regard to position and orientation. Optimal lighting is crucial for an immersive MR experience, especially to ensure that the real world is clearly visible. Additionally, a computer with high 3D computing capability is required to ensure rendering at high frame rates, allowing for a responsive and immersive MR experience.

To establish the **MR configuration** in 2PASD and AVES, a monitor-based headset (Varjo) was used. It provided an immersive experience by combining elements from the virtual world and the real world, with the real world being captured by the headset's scene cameras. A green screen setup for chroma keying was used in this thesis to display the VE in the headset. The green color indicates areas in the image where the virtual elements are to be overlaid with the real environment, creating a fusion of virtual and real elements. This prevents real-world objects from being obscured with virtual elements in contrast to masking/cutout techniques (Section 2.2.4).

In 2PASD, a stand with green fabric, serving as the green screen, was set up around the mockup. In the AVES, the projection dome was illuminated in green using the projectors to create a green screen on the projection dome.

For the **VR configuration**, the same monitor-based headset (Varjo) has been used as for the MR configuration. It was configured to display only VE allowing the pilot to view virtual objects and information. Therefore, a fully virtual helicopter environment was created, including a helicopter interior and fully equipped cockpit with flight instruments, to provide a realistic and immersive training scenario. Furthermore, the movements of the pilot's virtual avatar were animated based on the measured control inceptor displacement. This setup eliminated the necessity for hand-tracking while providing a simple setup with increased immersion for the evaluated hands-on maneuvers.

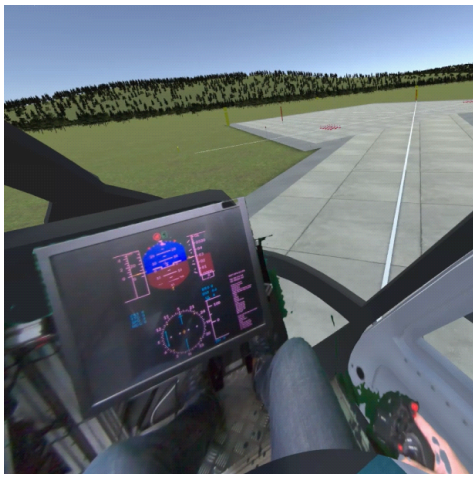
The **AR configuration** exhibits some hardware requirement similarities to the MR and VR configurations, including head tracking and high 3D computing capability. However, due to the see-through display of the utilized JVC headset, different requirements apply. To ensure optimal performance, a closed mockup with non-reflective textures is advantageous to prevent undesired reflections and image artifacts caused by real objects outshining the virtual image. Lighting is also important for the legibility of the cockpit instruments. To ensure optimal visibility and immersion with the see-through headset, it is recommended to set the display brightness of instruments to a high level. Additionally, proper internal lighting is crucial, but at the same time it must be carefully applied to minimize reflections that could decrease image quality of the see-through headset.

3.7.2 MR Implementations in Two Different Simulation Environments (2PASD and AVES)

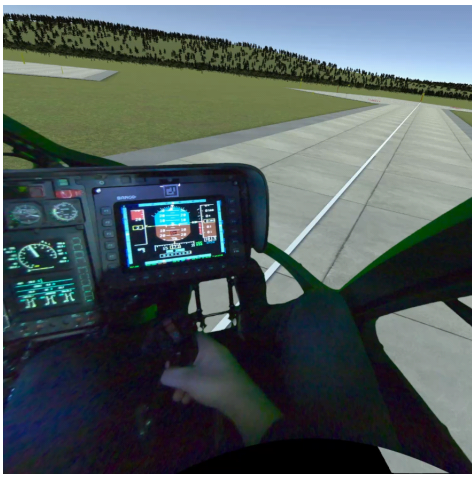
The implementation of MR in two distinct helicopter simulation environments, 2PASD and AVES, has been explored in this work. The key findings are summarized as follows.

Combination of the Virtual and Real World Components: The MR configuration in the 2PASD features a more prominent virtual component, as it not only displays the external view but also renders the helicopter cell and interior parts, thus occupying a higher share of the VE, see Figure 3.27a. In contrast, the MR configuration in the AVES features a fully equipped cockpit, thereby providing a higher proportion of real-world components, see Figure 3.27b.

Cockpit Interaction: The 2PASD only employs simple displays to provide essential flight information, such as Primary Flight Display and First Limit Indicator. For this reason, it is mainly used for hands-on maneuvers, as no cockpit interaction, e.g. pressing operating knobs or switches, is possible.



(a) 2PASD MR ©DLR. CC BY-NC-ND



(b) AVES MR ©DLR. CC BY-NC-ND

Figure 3.27: MR Configurations

Conversely, the MR configuration in the AVES enables complete interaction with the cockpit including haptic feedback as the mock-up resembles a fully equipped EC135 cockpit, which is beneficial for cockpit familiarization.

Space Requirements: The MR configuration in the 2PASD requires little space, as the mock-up is relatively small. In addition, it allows for optimal positioning of head tracking systems and light installations.
In contrast, the AVES MR configuration necessitates a minimum space equivalent to the size of the original helicopter cockpit.

Head Tracking and Illumination Positioning: In AVES the closed cockpit restricts head tracking and illumination positioning, which can negatively impact the immersion. It is essential to consider additional hardware placement and design when using a fully equipped cockpit like AVES. To minimize artifacts within the VE, it is essential to consistently replicate the lighting conditions employed during the initial setup. This helps to ensure that the color separation for chroma keying remains constant throughout the simulation.
In 2PASD, however, sufficient spatial flexibility allowed for easier integration of head-tracking and lighting systems, which leads to a more immersive and disruption-free VE.

3.7.3 Differences in Application of Monitor-based and See-Through Headset

The comparison between the monitor-based headset (Varjo) and the see-through headset (JVC) reveals distinct differences in terms of hardware, readability, immersion, interaction with the cockpit and comfort.

Hardware: Monitor-based headsets require the use of a green screen in conjunction with chroma keying, necessitating additional hardware.
In contrast, see-through headsets demand non-reflective surroundings to prevent the real world from shining through the VE.

Readability: Monitor-based headsets capture the real world through cameras and merge it into the VE. For displays which are not rendered as virtual elements, the readability and overall image quality depends on the resolution of the headset's cameras. To minimize image artifacts and frame rate drops, accurate head tracking and illumination are required. These measures may be supported by masking/cutouts in the cockpit area to further increase image quality or even stream directly flight instruments into the MR headset.
See-through headsets, on the other hand, utilize a see-through display, allowing pilots to perceive the original cockpit, resulting in cockpit instruments being readable more easily. However, a see-through display darkens the outside view due to the utilized display technology. As a result, bright flight instruments are required to enhance readability.

Presence: Overall VE experience differs between the two headset types. Monitor-based headsets provide a more immersive experience, with pilots feeling more present in the VE. This is caused by the fact that pilots are able to see their own hands and are not aware of the surroundings of the simulation facility.
See-through headsets, on the other hand, increase pilots' awareness of their surroundings, allowing them to see through the display and observe the real world. Therefore, the pilots were more aware of the respective simulation facility. In addition, body parts are not visible in areas where the VE is shown as an overlay in the headset display, mirroring the behavior of using masks within the monitor-based headset, where virtual objects obstruct the real world (Figure 2.9a).

Interaction with the Cockpit and FOV: Due to their restricted FOV both headset types impact the gaze behavior of pilots. All pilots reported that they had to increase head movements to obtain all the important flight information and to survey airspace. The FOV does also influence interaction with the co-pilots, as with the limited FOV pilots are not always aware of their co-pilot's actions.
The use of monitor-based headset can lead to visual distortions, which makes it difficult to estimate the distance of objects that are in the immediate vicinity of the headset (close to the head), in contrast to problem-free interaction with cockpit instruments that are located at a distance of approximately 70 cm. However, the headset's focus creates difficulties in monitoring display within peripheral vision. For the see-through-headset, the pilot's view is too restricted by the headsets' small vertical FOV of 45 degree. This strongly degrades the broad FOV provided by the large cockpit windows of helicopters.

Comfort: No physical symptoms in regard to the weight of the headsets have been reported throughout the studies. However, the pilots estimate that wearing the headsets for extended periods of time might have a negative impact on comfort.

Key Considerations Regarding the Application of MR in Helicopter Simulation

The different implementations of MR established for the evaluation studies provide insights into effectively implementing MR in helicopter simulation using various technologies. Based on the gathered knowledge several factors were identified, which should be taken into account when applying MR technology. These different factors are categorized into two areas of consideration.

The **Use Case Considerations** address relevant factors regarding the application of an individual simulator. These considerations include factors with regard to presence, interaction with cockpit, gaze behavior and readability. This shall allow for an assessment of which factors are of importance for a specific application depending on the training tasks to be performed with the respective simulator.

The **Implementation Considerations** address technical characteristics of MR setups and provide guidelines in designing and optimizing MR simulators. These considerations include factors regarding headset type selection, environmental and technical setup, hardware limitations and reduction of image artifacts.

3.7.4 Answer to Subordinate Research Question (1.1)

The identification and categorization of the most important considerations allows for answering Research Question (1.1):

(1.1) What are the key considerations for the application of MR headsets (here see-through and monitor-based headsets) in helicopter simulation?

To effectively integrate MR headsets into helicopter simulation, two areas of consideration must be taken into account: use case and implementation considerations.

Use Case Considerations

Presence

- Monitor-based headsets provide higher immersion (pilots see their hands, less awareness of real-word surroundings).
- See-through headsets increase situational awareness of real world (pilots see real-world surroundings), which may reduce immersion in the VE.

Interaction with Cockpit

- Monitor-based headsets: Cockpit instruments must be captured and streamed to the MR headset display, leading to potential image blur caused by head movement and reduced resolution. Alternatively, displays may be rendered as part of the VE and directly streamed to the MR headset display.
- See-through headsets allow for direct visibility of cockpit instruments but may darken the outside view, requiring bright displays for sufficient readability.
- FOV limitations (especially the JVC's 45° vertical FOV) restrict pilot perception of co-pilot actions and peripheral airspace, increasing the need of head movements.

Gaze Behavior and Readability

- Typical scan patterns are disrupted, requiring increased head movement to gather flight information.
- Monitor-based headsets' image quality of real-world cockpit objects depends on their camera resolution
- See-through headsets suffer from reflections on real-world objects and have brightness issues.
- Stabilization of helicopter can be challenging due to limited FOV and increased focus times for reading instruments.

Implementation Considerations

- **Headset Type Selection**
  - Monitor-based headsets require additional hardware to provide green screens for chroma keying but offers higher immersion.
  - See-through headsets require non-reflective environments and bright displays but provide direct real-world visibility.
- **Environmental and Technical Setup**
  - Lighting must be optimized to avoid reflections and image artifacts.
  - Head tracking precision is critical to avoid orientation errors and maintain high frame rates of the VE.
  - Space constraints may negatively impact hardware placement (especially in fully equipped cockpits like AVES).
- **Hardware Limitations**
  - FOV limitations of current MR headsets impact gaze behavior.
  - High 3D computing performance is required to maintain high image frame rates and minimize display latency.
  - Wearing MR headsets for extended time periods may negatively affect comfort.
- **Reduction of Image Artifacts**
  - Monitor-based headsets' image artifacts can be minimized using a hybrid approach that combines chroma keying and masking/cutouts.
  - Monitor-based headsets displaying flight information on virtual instruments in the VE increases image quality and decreases artifacts compared to capturing and displaying real-world instrument displays.
  - Adjustment of the simulator setup to reduce shadows on the green screen and avoiding reflective surfaces can minimize image artifacts.

3.8 Potentials and Limitations of Using MR in Helicopter Simulation

This section summarizes the main findings of the two studies, focusing on the application of MR technology in helicopter pilot training. The limitations and potentials of MR in helicopter simulation are identified based on the visual system configurations evaluated in 2PASD and AVES. From these findings, overall potentials and limitations of MR simulators are derived.
The MR headsets used in the studies represent the state of technology available at the time. MR headsets are subject to ongoing improvement, including advancements in display technology, tracking systems and software capabilities, which can significantly impact the MR experience.

3.8.1 Summarized Pilot Feedback and Further Observations of Configurations in 2PASD

The pilots preferred the **MR configuration** over the VR configuration, which were both established using the Varjo headset. They felt present during the MR simulation and less disconnected from the virtual world compared to the VR configuration. The inclusion of real-world elements, such as visible hands, significantly contributed to their preference. The Varjo headset's use of cutouts and chroma keying blended both worlds smoothly without generating opacity issues, as observed within the AR configuration. The pilots also reported better depth perception and height estimation with the MR in comparison to the 2PASD CON configuration.
However, the pilots' feedback reveals that brief blurring effects occurred when moving their heads, adding to the challenges of changing their gaze direction between ITC and OTW for checking flight instruments. In addition, the readability of flight instruments was decreased with the Varjo headset. The pilots needed to deliberately focus on the displays to read the required information.

The pilots reported improved readability of instruments and reduced distracting artifacts during maneuvers with the **VR configuration**. However, they felt disconnected from VR. The pilots' feedback suggests that a body or hand tracking is necessary, as they stated that interacting with the simple avatar feels less immersive.
To further enhance the VR configuration, animating the avatar with body tracking can be beneficial, providing pilots with a more realistic and engaging experience. Additionally, providing effective haptic feedback, for example with haptic gloves, when pilots interact with cockpit buttons can contribute to increased presence. In addition, a more detailed virtual cockpit with fully equipped flight instruments is essential for helicopter flight training.

For the studies in this research, the JVC headset was employed for the **AR configuration**. Pilots reported that the virtual world did not seamlessly overlay the real world in this setup, resulting in the real environment being more prominent than the virtual. In addition, the perceived readability of displays varied between pilots, with factors such as ambient brightness affecting the sharpness of the image and readability.
To optimize the AR configuration, a closed mock-up with non-reflective textures could potentially eliminate undesired reflections and enable objects from the real world to blend more smoothly into the image. Additionally, special attention should be payed to the illumination level, as it impacts the readability of cockpit instruments. Another limitation of the JVC headset is its small vertical FOV. It was found to reduce the presence ratings for the configuration. Another deficiency decreasing presence is given by body parts being cut outside the cockpit area, as the VE overlays real-world object in the area of the outside view.
All pilots reported that the JVC headset provided a less immersive experience and was less suitable for conducting the evaluation maneuvers compared to the Varjo headset.

Pilots reported poor immersion and engagement with the **2PASD CON configuration** during all maneuvers. The TV displays lacked depth perception and individual TV display frames disrupted the outside view, creating visual clutter that disturbed the overall experience. The CON configuration with the TV displays provides a limited FOV, affecting peripheral vision during the maneuvers. The asymmetrical FOV due to off-center seating position in the side-by-side cockpit arrangement results in smaller horizontal FOV on the left-hand side, reducing visibility of visual cues on that side. In addition, the fixed FOV of the TV display configuration contrasts with the MR headset, which provides a seemingly larger FOV by adapting the FOV to the orientation of the head.
However, the CON configuration provides a simple setup that employs TV screens as visual systems. This design eliminates the need for additional devices such as head-tracking or green screens, making it a relatively simple solution.

In general, no systematic disparity in **flight performance** was observed for the different visual systems or maneuvers in 2PASD. This indicates that the visual system configurations did not meaningfully influence pilot performance in this simulator.
No notable symptoms of **simulator sickness** were reported throughout the studies, with the exception of the dynamic flight maneuver. This suggests that, for exposure periods of 15–20 minutes, the occurrence of simulator sickness seems to be low.
Regarding **gaze behavior**, a slight variation between configurations was detected. Notably, pilots tended to minimize head movement with the MR configuration, likely to prevent image blurring.

3.8.2 Summarized Pilot Feedback and Further Observations of Configuration in AVES

Within all the configurations evaluated, the **AVES CON** configuration with its fully equipped cockpit was preferred by all pilots and received the highest presence ratings. The pilots stated that this configuration provided the largest FOV, allowing them to maintain awareness of the helicopter's orientation at all times and also observe their co-pilot's interactions through the peripheral view which was limited with the MR headset configuration. The flight instruments were clearly readable and the absence of the headset's added weight contributed to enhanced comfort.
However, utilization of the projection dome introduces significant infrastructural requirements. Not only is it physically large, but it also weighs several tons (10 t in the case of AVES with motion platform), necessitating a substantial building with a robust foundation to house the simulator.

The use of **MR** headsets in a fully equipped cockpit presents several challenges. In comparison to the fully equipped cockpit with projection dome, the pilots stated that the image in the MR headset occasionally flickered or lagged which arose due to suboptimal head-tracking in the fully equipped cockpit and artifacts caused by the cockpit windows, contributing to a less immersive experience than using the projection dome as the visual system. The pilots reported that the flight instruments were displayed less sharp and blurred even more when they moved their heads with the MR headset during hover and slalom maneuvers. Additionally, they had to deliberately focus to identify values on the flight instruments or visual cues outside. The pilots were not able to detect movement of persons in the cockpit even though they had a person sitting on the co-pilot seat.
However, the MR configuration in a fully equipped cockpit presents several advantages. A significant benefit is that it does not require a projection dome, thereby reducing the need for large and expensive infrastructure. This compact solution allows for a more space-efficient setup, while also enabling the possibility of building a portable, flexible simulator. The MR configuration provides pilots with access to original cockpit instruments, which facilitates familiarization with the real helicopter's layout and functionality (e.g., knowing the position of displays with crucial flight information, feeling haptic feedback from knobs and switches). The blending of virtual and real-world elements can be easily adapted through software for rendering the VE, allowing for flexible configurations. Additionally, the use of video-captured images enables pilots to see themselves within the simulation, which enhances presence compared to VR configurations.

General observations in the AVES indicated that the MR configuration led to a slight decline in pilot flight performance, a trend that was also reflected in the presence ratings. Specifically, higher sense of presence correlated with increased flight performance.
Incidents of **simulator sickness** were caused by the AVES MR setup, likely due to head-tracking inaccuracies (insufficient distance between the pilot's head and the MR headset due to the compact cockpit design), as well as visual artifacts, such as flickering caused by reflections on the cockpit windows.
A noticeable difference in pilot **gaze behavior** was observed in AVES. Pilots glanced more frequently at the cockpit instrumentation in the CON configuration compared to their gaze behavior in the 2PASD.

3.8.3 Answer to Subordinate Research Question (1.2)

On the basis of the previous chapters, results of Study 1 (Section 3.2), results of Study 2 (Section 3.5) and the comparative analysis (Section 3.6), RQ (1.2) can be answered:

(1.2) What are the potentials and limitations of using MR headsets for helicopter pilot simulation training compared to other visual systems (TV screens and dome projection system)?

Potentials of MR in Helicopter Training

- **Enhanced Presence**
  - Pilots reported feeling more present in MR than VR due to visible hands and real-world blending.
  - MR bridges the gap between fully virtual VR and the real world, reducing the feeling of being disconnected pilots experienced in VR.
- **Adaptability**
  - Virtual elements (e.g., flight instruments) may be flexibly rendered together with virtual and real-world components for tailored training scenarios.
  - MR allows pilots to interact with original cockpit controls, improving familiarity with helicopter layout and haptic feedback (e.g., knobs, switches).
- **Advantages Compared to TV Displays**
  - Unlike TV-based configurations, MR eliminates visual clutter from screen frames and provides a more seamless OTW.
  - MR headsets adjust the pilots' FOV based on gaze direction, unlike TV displays, which statically limit peripheral vision.
- **Flight Performance**
  - There is no impact on flight performance compared to TV screens and no major impact compared to a fully equipped cockpit with dome projection.
  - The results show a correlation between the presence experienced by pilots and their flight performance.
- **Space Efficiency**
  - MR configurations require less space and reduced infrastructure than traditional projection domes (e.g., AVES' 10-ton dome).
  - MR enables portable and modular simulators (e.g., 2PASD's setup with MR headset, green screen and head tracking).

Limitations of MR in Helicopter Training

- **Instrument Readability and Visual Clarity**
  - Blurring and flickering during head movements with monitor-based headsets due to tracking delays or window reflections in cockpits may occur.
  - The reduced readability of instruments with the monitor-based headset forces pilots to focus on displays, unlike the effortless visual perception of information without MR headsets.
  - In contrast to VR or dome projection systems, see-through and monitor-based displays are highly dependent on ambient light, with cockpit and display brightness affecting display sharpness.
- **FOV Restrictions**
  - MR headsets provide a smaller FOV (hardware constraint), reducing presence and peripheral awareness compared to projection domes.
  - With see-through displays, virtual overlays cut off real body parts in the image (e.g., hands near windows), disrupting natural interactions.
  - The limited FOV can lead to different gaze behavior.
- **Technical and Ergonomic Issues**
  - Head-tracking latency causes flickering or lag in MR image, making it feel less immersive.
  - Blurring, latency and FOV limitations can contribute to simulation sickness.
  - Wearing headsets for extended periods of time may cause fatigue, whereas projection domes offer enhanced comfort (no headset required).

In conclusion, MR for helicopter pilot simulation training offers both potential advantages and limitations compared to conventional visual systems like TV screens or projection systems. While MR configurations provide flexibility, reduced infrastructure requirements and enhanced immersion, they also have limitations such as restricted FOV, reduced readability of flight instruments, simulation sickness and ergonomic considerations. While MR headsets cannot yet fully match the immersion and comfort of fully equipped cockpits with projection domes, they show promise to enhance or complement helicopter pilot training.

4 Evaluation Assessment Methods for MR Helicopter Simulation

In previous studies (Section 3.1 and Section 3.4), short-duration flight maneuvers were employed to evaluate key aspects of MR technology. Given the numerous configurations involved, standardized short maneuvers lasting no more than five minutes with an additional malfunction procedure were flown, resulting in about ten minutes of simulated flight time. While this approach proved suitable for initial MR evaluation, it was clear that training on a simulator would necessitate longer periods of engagement. Consequently, this chapter focuses on the development and evaluation of extended-duration flight tasks designed to cover a broader spectrum of training scenarios. The primary objectives are twofold: firstly, to develop and refine MR-based flight maneuvers that comprehensively meet the training needs of pilots and secondly, to evaluate the long-term usage of MR technology in a simulator through longer, more complex flight tasks. In addition, the pilots were asked to give feedback on two different presence questionnaires to validate the previous choice of using the PQ of Singer and Witmer.

This chapter initially presents an overview of the development of flight tasks, outlining the essential key considerations and objectives determining its design. The following section presents Study 3, which is based on the publication of the author of [93], in which pilots evaluate the developed tasks and and give an outlook on their feasibility for future studies. The simulation setup and the evaluation methods used are then explained before the results of this study are presented.

4.1 Development of Flight Maneuvers for Evaluation of MR-based Visual Systems

The proposed flight tasks are based on the idea of utilizing maneuvers that are flown in helicopter pilots' proficiency checks. The proficiency check must be completed by the pilot together with a flight examiner once a year in order to renew their license. This can be done in flight or in a qualified simulator. During this flight, pilots must perform both standard maneuvers and emergency procedures. This means that a wide range of maneuvers need to be demonstrated. Future simulators should therefore include all the systems and functions that these flight maneuvers require. For the evaluation of MR, two flight tasks were created: Arial Work and Ground Work flight tasks.

4.1.1 Aerial Work Flight Task

The intention of the Aerial Task is to allow pilots to get familiar with the simulator before moving on to more challenging maneuvers. Therefore, it starts with standard procedures like normal take-off, flying standard curves and performing steady climb or descent flight. The idea is that, at first, no complex maneuvers are demanded, but rather ones that must be accomplished in a proficiency check. Next, pilots are tasked to perform a quick-stop maneuver, where the helicopter needs to decelerate quickly and come to (almost) hover, increasing the difficulty level. Then, an emergency procedure should be performed during approach, however, in this study, only a normal approach and landing will be simulated, as the used simulator does not offer the functions and systems to simulate an emergency procedure. From the described considerations and task elements, the following flight was incorporated (Figure 4.1):
The first flight phase starts with standard procedures. The pilot starts from an airport on a runway performing a normal take-off procedure. Once at 1500 feet altitude, the helicopter should hold steady level flight at 90 kts and keep the speed during cruise flight. The pilot is then asked to perform a 180° right turn with a 30° bank angle. Thereafter, the pilot should climb to 2000 ft altitude at 500 ft/min rate of climb while maintaining constant heading. Subsequently, the pilot should perform a 180° left turn, again with 30° bank angle. After leveling off, the pilot should then descend to 1500 ft altitude at 500 ft/min descent rate with constant heading. Once the helicopter is stable in forward flight at approximately 90 kts, a quick stop should be performed while maintaining altitude and decelerating to a speed of 20 kts or less. Finally, a left turn of approximately 180° should be performed and an approach to the airport should be attempted. The pilot should report when the airport is in sight. As the 2PASD is not able to simulate emergency procedures, the pilots subsequently shall perform a normal landing as a replacement. The landing completes the flight task.
For the study in this thesis, an additional objective was to identify which emergency procedures were suited for integration into the task. Therefore, after completing the task, the pilots were asked which emergency procedures would be suited to be included to terminate the Aerial Work Task.

4.1.2 Ground Work Flight Task

The Ground Work task aims to combine various ground maneuvers into a single task. At the same time, this task should be precisely defined so that pilots can perform the task repeatedly, which makes it easier to compare the results within different configurations. In addition, one aspect for the definition of this tasks is also to evoke different gaze behavior compared to the Aerial Work task.
Whereas the maneuvers close to the ground should require more OTW views, as the outside visual references are necessary to hold the helicopter stable, the cruise flight with given

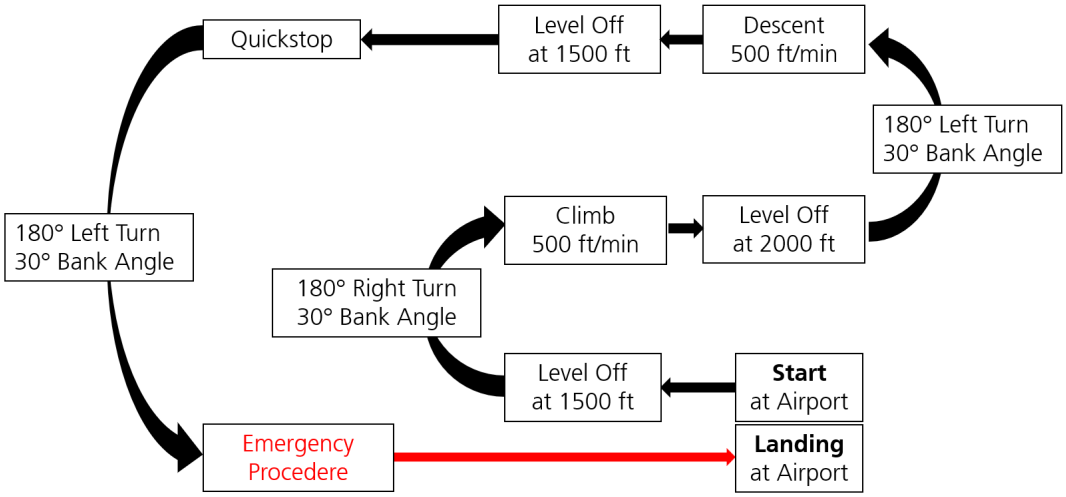


Figure 4.1: Aerial Work Task ©DLR. CC BY-NC-ND

guidelines such as climb rate, turn rate, heading, etc. requires frequent ITC looks to survey the flight instruments. With that in mind, the idea is to look for close to ground maneuvers which are also part of the MTE tasks of ADS-33 [63]. As there was already a MTE test course available (Figure 5.3), it was used to provide the pilots with the visual cues necessary for conducting MTEs. With this test course as basis, the following Ground Work Task were defined:

At first, the pilots start with the hover MTE maneuver (see "Start" in Figure 4.2 and also Section 3.1.2). Thereafter, the pilot flies along the taxiway to the lateral repositioning task by turning left and then right on the taxiway. When the green sign with a "B" is on the right, the pilot should turn heading to the north, the visual cues are now in sight. If stable, the pilots starts with the lateral repositioning to the right until a similar green sign with an "A" is in sight (see Figure 4.3). After recovered to stable hover, the pilot may start the repositioning to the left. After reaching the green sign "B", the pilot should turn the nose of the helicopter to the west and begin taxiing to the runway to perform a right-hand traffic circuit. The task is finished by landing on the runway.

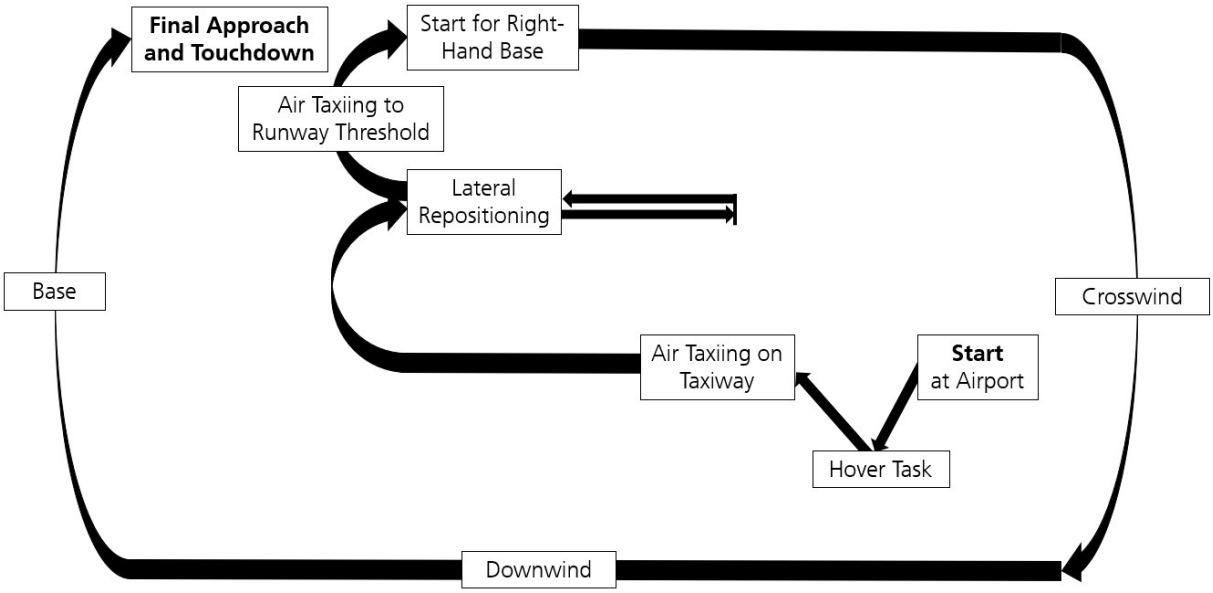


Figure 4.2: Ground Work Task ©DLR. CC BY-NC-ND

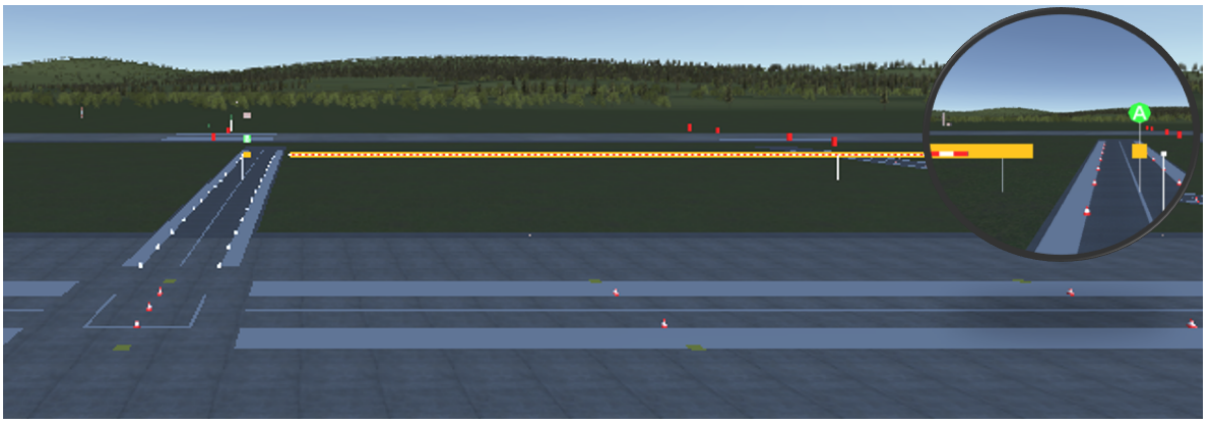


Figure 4.3: MTE Lateral Repositioning ©DLR. CC BY-NC-ND

4.2 Study 3: Proposing a Standard Maneuver for Evaluation of New Visual Systems for Helicopter Simulation

The main objective of Study 3 is to assess the suitability of the developed flight tasks (Section 4.1) for the evaluation of different visual systems and simulator configurations. Two separate tasks were developed, each lasting approximately 15 minutes. This design allowed pilots to evaluate individual flight aspects with questionnaires immediately after completing each flight task, when they could still clearly recall and assess their experience. Study 3 was conducted in the 2PASD (Section 2.1.2) just like Study 1 (Section 3.1). The flight tasks were performed in both the conventional 2PASD configuration and the MR configuration. The flight tasks were evaluated with the Simulation Sickness Questionnaire (SSQ), as well as with the SUS questionnaire. Although the SUS has not been used in previous studies, it was presented to the pilots in order to evaluate the previous usage of PQ from Witmer and Singer and determine which questionnaire provides a more accurate and reliable measure of presence. In addition, eye tracking data with the Tobii Pro Glasses 3 were recorded for ITC/OTW ratio evaluation.
In the following sections, the experimental setup for the study is described together with the configurations in which the flight tasks were performed. This is followed by a brief overview of the applied assessment methods.

4.2.1 Experimental Setup

Study 3 was conducted in the 2PASD (Section 2.1.2). The 2PASD was chosen as it offers greater robustness of the MR configuration due to its flexible setup, which allowed for more accurate positioning of the green screen, head tracker and illumination compared to the narrow AVES cockpit. This experimental setup provides a more immersive and realistic virtual environment experience, with fewer visual artifacts and head tracking issues (e.g., misalignment) compared to the AVES MR configuration from Study 2 (Section 3.4).
The conventional setup of the 2PASD with its visual system comprised of five TV screens serves as baseline for evaluating the flight tasks. For the MR configuration, the Varjo headset (Section 2.1.3) was used as the visual system as in previous Studies 1 and 2.
To facilitate the use of chroma keying, a green box was built around the cockpit (Section 2.2.5). As in Study 1 tracking stations for head tracking and light sources were installed. In contrast to Study 1, the MR configuration differed as the masking functions were used in combination with chroma keying (Section 3.3) to minimize image artifacts, for example on the displays of the flight instruments. Also additional light sources were added to illuminate the green screen more evenly and avoid shadows, creating an optimal environment for video capturing through the Varjo headset.



(a) Conventional Configuration
©DLR. CC BY-NC-ND



(b) Mixed Reality Configuration
©DLR. CC BY-NC-ND

Figure 4.4: Configurations of Study 3

4.2.2 Task Definition

The individual flight tasks are described in detail in the previous chapter (Section 4.1). The test matrix covering the two flight tasks in both configurations is presented in Table 4.1. At first, the pilots were granted some time for a familiarization flight, where they freely may test different maneuvers getting costumed to the Varjo headset. Thereafter, they proceeded with the Aerial Work task followed by the Ground Work task both in MR configuration (Figure 4.4b). Subsequently, they flew the both tasks also with the CON using the 2PASD's visual system (Figure 4.4a).

Table 4.1: Test Matrix of Study 3

	Configuration	Flight Task
MR	Mixed Reality (Monitor-based Headset)	Familiarization Flight
MR	Mixed Reality (Monitor-based Headset)	Aerial Work
MR	Mixed Reality (Monitor-based Headset)	Ground Work
CON	Conventional Simulator With TV Screens	Aerial Work
CON	Conventional Simulator With TV Screens	Ground Work

4.2.3 Assessment Methods

For the evaluation of the study different assessment methods are used (Section 3.1.3). Simulation sickness is assessed using the SSQ, presence is evaluated by applying the SUS questionnaire and gaze behavior is analyzed utilizing ITC/OTW ratio. The flight performance is not considered here, as the primary objective is assessing the flight tasks.

The SUS was used in this study to investigate if it is more suitable for assessing presence in a MR helicopter simulation environment compared to the previously used PQ. In order to obtain feedback on the two questionnaires from the pilots, both questionnaires were presented to them after the first task in order to obtain comparative feedback. For time reasons, however, only the SUS was subsequently completed for all flight task in the study. The SUS consists of six questions that aim to evaluate presence in relation to real experiences, specifically assessing how much the VE feels like a real or physical space (Appendix A.2). It is shorter than the PQ developed by Singer and Witmer and has a different approach as it encourages the pilots to compare their current VE experience to real life experiences.

In order to record eye tracking data for the assessment of ITC/OTW ratio the pilots were wearing the Tobii Pro Glasses 3 (Section 3.1.3) in the CON configuration.

4.2.4 Pilot Experience

In Study 3 the same pilots participated as in Study 1 (Section 3.1.4) and Study 2. The pilots were already familiar with the Varjo headset and the different visual configurations from the previous studies.

4.3 Results of Study 3

In this chapter, the results of Study 3 are presented. Initially, the results of the SSQ for the two flight tasks (Aerial Work and Ground Work) are discussed regarding the impact of longer flight duration on the pilots' well-being. Subsequently, the feedback on the SUS is presented and its applicability is discussed. Thereafter, the eye tracking data of the Tobii Glasses 3 are analyzed for both tasks to identify whether the two flight tasks cause different gaze behaviors in the pilots, as the tasks were designed in such a way that they evoke different gaze behavior patterns, which makes it possible to investigate differences in the pilots' gaze behavior. Finally, general feedback of the pilot on the two developed flight tasks is presented.

4.3.1 SSQ Results of Flight Tasks

One of the objectives of Study 3 is to investigate the impact of longer flight tasks on simulation sickness. Therefore, the pilots reported their simulation sickness symptoms after each task using the SSQ. Flying close to the ground is very demanding and, based on feedback from many pilots and also from various studies (see visual scene density and altitude above terrain in Section 1.1.3), it is also more prone to cause simulation sickness symptoms and high workload in simulation compared to other flight tasks. Therefore, the SSQ ratings should provide information on whether specially the Ground Work but also the Aerial Work task is suitable for evaluation of new visual systems. In Figure 4.5 the simulation sickness scores of all pilots are presented for the Aerial Work (AW) and the Ground Work (GW) Tasks.

Pilot B started with *negligible* symptoms (score 4) after the first flight task and experienced a increase to *minimal* symptoms (SSQ score 7) throughout the rest of study. The pilot explained that longer flight tasks resulted in more weight-related discomfort because of the MR headset and also associated sweating, which lead to headache and fatigue symptoms.

Pilots A also showed an increase in simulation sickness symptoms for the consecutive flight tasks when using the MR configuration. They struggled more with the MR setup during Ground Work Task, which can be attributed to the fact that flying close to the ground is a challenging and potentially simulator-sickness-inducing scenario.

Pilot C did not experience any simulation sickness symptoms throughout the study. This can be attributed to the fact that the pilot has several hours of experience with VR, not only in flight simulation but also for recreational use. Furthermore, susceptibility strongly varies between individual humans.

The simulation sickness were generally lower for the CON configuration, aligning with pilots' comments that they are more accustomed to the conventional setup. Nevertheless, they experienced some progress with customization of the MR headset. For example, the pilots reported that they moved their heads more often in order to obtain all required

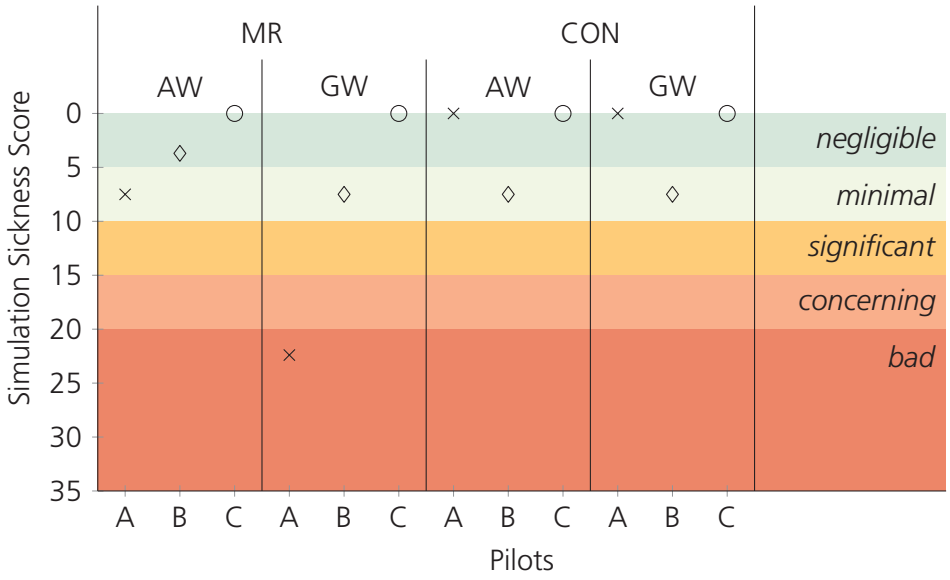


Figure 4.5: SSQ Scores for Aerial Work (AW) and Ground Work (GW)

information, which could not be seen at a glance with the MR headset due to the limited FOV and the associated reduced peripheral vision.

Comparison of the results to the analysis of Study 1 (Section 3.2), where almost no simulation sickness symptoms occurred,shows there was a notable increase in SSQ scores across Pilots A and B, likely due to longer flight tasks. During Study 1, Pilot C reported mild symptoms, but these were attributed to an unintentional dynamic roll maneuver during the deceleration phase as the helicopter approached the hover point. This incident was an isolated event caused by increased speed on approach leading to an unintentional sharp deceleration. In Study 3, Pilot C reported no simulation sickness symptoms. The results of the study indicate that each pilot experiences the simulation environment individually and is differently susceptible to simulation sickness. These differences are attributed to factors such as maneuvers, daily health status and adaptation or habituation to the technology used in the simulation.

Regarding the grading score the severity of symptoms falls mainly in the range between *minimal* and *significant*. However, the maximum score of Pilot A of 22 for the Ground Work Task in MR configuration is categorized as *bad*, which is beyond *unacceptable*. Here the grading score based on [74] should be questioned, as Pilot A was able to accomplish the task and did recover from the symptoms quickly. For future qualification the grading score needs to be evaluated for categorizing the MR simulation regarding simulation sickness symptoms and their impact on the training of pilots.

Presence Questionnaires (SUS)

The SUS (Appendix A.2) was presented to the pilots to obtain feedback on the questionnaire's applicability to evaluate presence in an MR helicopter environment. The SUS questions were perceived by the pilots as allowing for primarily binary responses ("yes" or "no") rather than allowing for nuanced responses. Specifically, question five of the SUS was deemed lengthy and challenging to comprehend due to its broad description. Additionally, similar questions within the questionnaire raised concerns by the pilots about what specific aspects were being measured within each question.
Due to pilot feedback indicating that the questions were posed in such a way that they were most likely to be answered with "yes" or "no" and that it was difficult for them to provide nuances in between their ratings, the resulting data was not considered significant and was therefore excluded from further analysis.

4.3.2 Eye Tracking Evaluation of Flight Tasks with the Conventional Setup

One aspect that influenced the development of the two flight tasks (Aerial Work and Ground Work) was the objective that different gaze behavior patterns would be evoked in pilots by the flight tasks, allowing for the examination of varying ITC/OTW ratios to investigate differences in pilot gaze behavior. This goal has been achieved, as shown in Figure 4.6, which presents the evaluation of eye tracking data recorded using the Tobii Glasses 3. Only eye tracking data from the CON configuration was used for the flight task evaluation, because the MR configuration was found to affect pilots' gaze behavior, as discussed in Study 2 (Section 3.6.3).
The diagram reveals that during Aerial Work Task in CON configuration, both pilots looked more ITC compared to the Ground Work task. This is consistent with pilot comments explaining that the requested flight parameters, such as speed, climb rate and heading, made the flight resemble an IFR rather than a VFR flight.
Notably, during Ground Work Task, the pilots tended to look outside more often to stabilize the helicopter in flight, which aligns with the ITC/OTW ratio for Ground Work task. This pilot behavior is consistent with their comments that they are using the given MTE visual cues and additional cues in the VE for stabilizing the helicopter in flight.
The results underline that the two developed flight tasks are capable of inducing varying gaze behavior and are therefore well suited to be utilized for investigating the effects of visual systems on pilot gaze behavior.

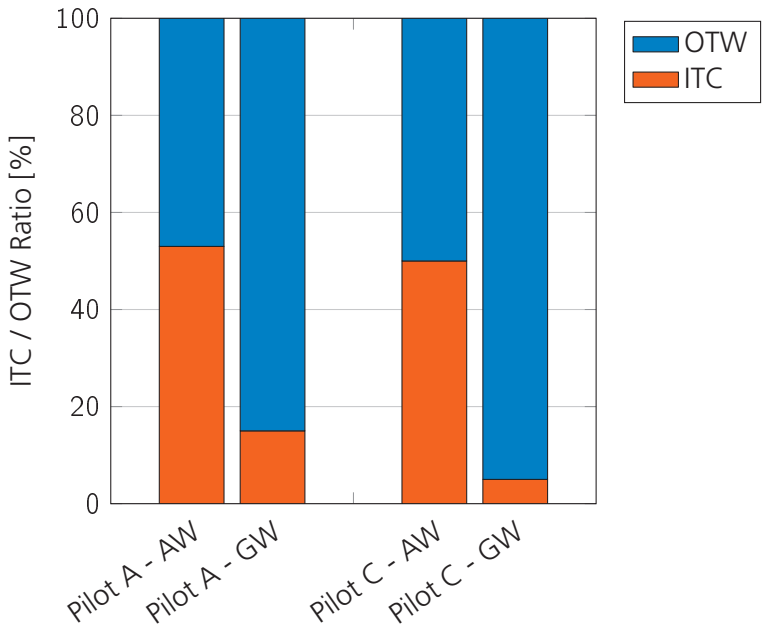


Figure 4.6: ITC/OTW Ratio for Trainingflights With Tobii Pro Glasses 3

4.3.3 Pilot Feedback on Aerial Work Task

Thereafter, the pilots were asked to suggest an emergency procedure which could be included in the Aerial Work tasks for future studies in industrial settings.

The pilots were asked to provide feedback on the Aerial Work flight task. The flight task begins with a normal take-off, followed by standard procedures for familiarization. The pilots reported that the time taken to familiarize themselves with the MR simulator was appropriate, neither too long nor too short. Subsequently, they executed a quick stop maneuver, which reportedly increased their workload, but would spare them the capacity to then perform an emergency procedure instead of a normal approach and landing on the runway.
To find a suitable emergency maneuver to append to the Aerial Work task in future evaluations, the pilots were asked to suggest typical emergency procedures they perform during training. They recommended scenarios such as autorotation, tail rotor failure and One Engine Inoperative (OEI). These procedures were assessed based on their comments and relevance to industrial simulation settings (see Section 5.2).

- **Autorotation:** The test pilots explained that they had not flown a simulator that could realistically simulate autorotation yet. Since the flight dynamics and engine model significantly impact the quality of the autorotation simulation, this emergency procedure was not chosen for future evaluation.

- **Tail Rotor Failure:** This emergency procedure can only be simulated in a safe environment, as the loss of tail rotor thrust can have catastrophic consequences and

cannot be simulated in real flight. Although the simulation enables risk-free practice, it does not offer the possibility of comparing the simulation with real conditions in order to assess the quality of the simulation. It was therefore excluded from the initial evaluation of the simulator in an industrial environment.

- **One Engine Inoperative:** The OEI procedure requires interaction with the cockpit and can be performed in real flight. According to the pilots, the procedure is resembled well in standard training simulators. Since the evaluation of MR should include interactions with the cockpit, the OEI emergency procedure provides a well-suited maneuver for the given purpose.

Based on the assessment of different emergency procedures, the OEI emergency procedure is chosen for the evaluation of MR simulator in an industrial setting (Section 5.2). However, depending on the focus of future evaluations, it may be substituted by other emergency procedures such as autorotation and tail rotor failure to tailor the flight task to specific requirements.

4.3.4 Pilot Feedback on Ground Work Task

The pilots generally reported that the most challenging maneuvers in the simulator are those close to the ground, where height estimation is difficult due to the limited OTW view of a simulator and low-resolution environmental visualization. Small details like moving grass, missing texture variations on taxiways or missing visual cues (such as wind flags) further complicate maneuvering. Despite these conditions, all pilots were capable to fulfill the given tasks. They reported that the Ground Work Task is definitely more challenging than the Aerial Work Task, which is intended by the study design. The task was described as demanding but feasible at the same time with both the 2PASD conventional setup and the MR setup. Nevertheless, the MR caused more simulation sickness symptoms, see Section 4.3.1.
The visual cues of the test course are useful in supporting the pilots to successfully complete the tasks. However, this visual information is primarily known by test pilots due to their education, whereas other pilots may not be as familiar with these type of cues as they are part of the ADS-33 standard which is not included in the usual pilot training or operation. Therefore, the pilots recommended a comprehensive briefing in advance of flying the Ground Work task.

4.3.5 Conclusion of Study 3

In summary, the pilots were all capable to complete the two tasks with the conventional setup and also MR headset. They stated that all elements of the flight tasks are also part of either daily flying routines or the proficiency check which has to be accomplished by pilots each year to renew their license. Therefore, the developed tasks indeed reflect actual flight training procedures. The tasks also effectively induced varying gaze behavior and are therefore well-suited for investigating the effects of visual systems on pilot behavior. The pilots stated that the task sequences were well designed to strike a balance between demanding maneuvers and familiarization periods, ensuring a good mix of challenge and comfort.

4.4 Discussion of the Applicability of Assessment Methods

This chapter evaluates the suitability of the assessment methods for MR helicopter training simulation used in Studies 1 - 3. Different aspects are considered to assess the applicability of each method. First, the ability of the pilots to work with the assessment methods (ease of use) is considered. Second, the quality and accuracy of the resulting data in conjunction with the grading scale and/or categorization scheme provided by each method is evaluated. Finally, suggestions for adaptions and improvements are derived in order to provide application remarks and recommendations when applying the individual assessment methods for the evaluation of MR simulation.

4.4.1 Assessment with Standardized Questionnaires

As described in Section 3.1.3, following questionnaires were used throughout this thesis: For evaluating presence, the Presence Questionnaire (PQ) of Singer and Witmer [65] was used in Studies 1 and 2, while SUS from Usoh, Slater and Steed [67] was applied in Study 3 to validate the choice of using the PQ.
For the evaluation of simulation sickness, the Simulation Sickness Questionnaire (SSQ) of Kennedy et. al. [71] is used in all three studies.
The discussion of each assessment method is divided into four categories: ease of use, accuracy and quality of the data, the grading scale and/or classification scheme, application remark and recommendation for future use.

Presence Questionnaire by Singer and Witmer

The PQ (Appendix A.1) is a frequently used questionnaire to measure the test subject's experience of "being present in a VE, even when physically located elsewhere" [65]. It consists of 32 questions that assess various factors influencing presence, classifying them as Control Factors, Sensory Factors, Distraction Factors and Realism Factors.

- **Ease of Use:** As the PQ contains very general questions for the evaluation of the VE, but not specifically for the helicopter evaluation, the PQ was explained in detail during each study's briefing. For example, the question "How natural was the mechanism which controlled movement through the environment?" [65] for this application relates to the pilot's control inceptors.
  Nevertheless, the intent of some of the questions remained unclear, resulting in the pilots checking back while filling out the questionnaire. In addition, the PQ is comprised of many questions, ultimately increasing reporting time. On the other

hand, pilots reported that the questionnaire allows for evaluation of various facets of a VE by reflecting different aspects of the experience.

- **Accuracy and Quality of the Data:** The results of the PQ ratings gathered in the studies are consistent with the general comments of the pilots. In addition, the results also show a correlation with flight performance (Section 3.6.2). Further investigation into this connection could offer valuable insights into the virtual environment's ability to replicate real-world conditions, thereby enhancing the training efficacy of a helicopter simulation (Section 4.4.3).
- **Grading Scale and Categorization:** In addition to the Total Score of the PQ, the subscales assess various dimensions of users' experiences within a VE, including control, naturalness, resolution and interface quality. These subscales aim to evaluate individual aspects of a VE in order to provide a better understanding of the perceived presence. The PQ score visualization adopted from [69] includes an additional adjective grading scale to assess usability of a system based on PQ scores. A score below 50 % of the PQ score is deemed *unacceptable*, while scores above 50 % are rated in 10 % increments as *unsatisfactory*, *marginal*, *satisfactory*, *very good* and *excellent*. The feedback from the pilots was found to match the classification. However, a larger number of pilots and studies are necessary to obtain a more precise statement about the significance of the adjective grading scale.
- **Application Remark:** The PQ features generic questions to assess any kind of VE. For the studies conducted in this thesis, the pilots were briefed how to apply the generic questions to the specific study.
- **Recommendation:** *The PQ is suitable for evaluating a subject's presence during a MR helicopter simulation. To ensure a more specific statement and easier completion, a comprehensive briefing is advised.*

Presence Questionnaire by Usoh, Slater and Steed

The SUS questionnaire (Appendix A.2, [67]) consists of six questions that aim to evaluate presence inside a VE in relation to real experiences.

- **Ease of Use:** The SUS questionnaire was used in Study 3 to gather feedback from pilots on its effectiveness in evaluating presence in a MR helicopter environment. The approach of this questionnaire attempts to capture the entire experience by including both real and simulated holistic scenarios. However, the evaluation of helicopter simulations in this research currently focuses narrowly on the pilot's actions in the cockpit and neglects preceding and subsequent events such as briefings, pre-flight checks and debriefings. The pilots found it difficult to compare the simulated experiences with real procedures. The pilots noted that the distinction between being in a

building beforehand and entering deliberately entering a simulator on the one hand, and performing perform pre-flight checks and subsequently entering the helicopter on the other hand, was substantial. However, the conducted studies mainly focused on the in-flight time when the pilots are already sitting in the cockpit and performing flight maneuvers. Therefore, the pilots were not able to effectively evaluate the simulation as they were aware that they were in a simulation environment from the start. The pilots also pointed out a critical point: The questionnaire did not address important aspects of evaluating the simulation, particularly its limitations and its impact on training effectiveness and immersion. By neglecting these crucial factors, the questionnaire was considered insufficient for the evaluation of the MR simulation.

- **Accuracy and Quality of the Data:** Based on pilot feedback, the data collected using the SUS was not considered significant and was therefore excluded from further analysis. This decision was made because the pilots' responses were heavily clustered either at the lower or higher end of the Likert-scale, rather than providing nuanced answers.
- **Application Remark:** The questions should be thoroughly discussed during the briefing, as the they have a strong visual component and therefore require a lot of imagination by the pilots.
- **Recommendation:** *For the research focus of helicopter simulation training the PQ of Singer and Witmer is preferable and should be chosen over the SUS questionnaire.*

Simulation Sickness Questionnaire

Simulation sickness can occur when using MR headsets due to a mismatch between visual and vestibular information (Section 3.1.3). The severity of symptoms typically increases with the time spent in a sensory mismatch condition. The assessment of simulation sickness can be conducted using the SSQ, which is a common technique for evaluating 16 physical symptoms during simulations. Participants rate their symptoms on a scale of *none*, *light*, *moderate*, and *severe* (Appendix A.3, [71]).

- **Ease of Use:** Pilots were able to quickly and easily complete the SSQ with little explanation required.
- **Accuracy and Quality of the Data:** The questionnaire covers typical symptoms resulting from simulation sickness. The pilots' general comments also validated the SSQ results, supporting their accuracy. The SSQ explicitly asks about symptoms, which, in combination with debriefing, can help uncovering causes for simulation sickness in the simulation setup. For example one pilot stated that the symptoms headache and fatigue, which were rated as *slightly*, were caused by the weight of the headset.

- **Grading Scale and Categorization:** The SSQ score ranges from 0 to 235. To transfer the numerical range into comprehensible ratings, a grading scale from [74] was added, which supplies adjective ratings ranging from *negligible* to *bad*. It was found that this grading scale appeared to be too strict as even slight symptoms are categorized as *bad*, leading to simulations being classified as unacceptable when only a few symptoms are present. [92] suggests that this categorization may not be applicable to the broad mass of student pilots and may be too stringent. This is consistent with the results of the studies, which show that symptoms lasting only a short time before disappearing resulted in being rated as *bad*. Further research is needed to validate the grading scale and determine whether it needs to be adapted for MR helicopter simulation, especially in regard to extended simulation periods.

- **Application Remark:** It is necessary to investigate more closely which severity of symptoms is acceptable for the pilots and how wearing the headset during longer simulation training affects the pilot. The grading score needs to be validated for future evaluation to allow for meaningful results in future qualification.

- **Recommendation:** *The use of the SSQ is recommended to get feedback of the pilots' well-being during the MR simulation and in combination with a debriefing allows for identification of causes of simulation sickness symptoms (e.g., headache due to the weight of the MR headset).*

4.4.2 Assessment of Flight Performance

The evaluation of flight performance for different visual systems using standardized tasks such as those of the ADS-33 provides a basis for comparing deviations in flight behavior when using different visual systems.

- **Ease of Use:** The trained test pilots were familiar with the visual cues of the MTE test course. The non-test pilot was also able to use them in their intended manner for stabilization of the helicopter. Nevertheless, it still needs to be investigated if for all non-test pilots the visual cues are beneficial or if they cause too high workload if followed too rigorously, i.e. seeking *desired* task performance. Nevertheless, a comprehensive briefing can support the understanding of the flight tasks and the corresponding visual cues.

- **Accuracy and Quality of the Data:** The standardized MTEs serve to simplify the pilots' understanding of their flight tasks and to create comparable data for assessing the impact of the visual system on flight performance. The narrow scope of the task elements and the visual cues ensure that different pilots follow similar flight paths, resulting in a more comparable database of flight data compared to other, less specific flight tasks.

- **Grading Scale and Categorization:** The ADS-33 provides a performance classification scheme using three ratings of *not adequate*, *adequate* and *desired*. The scheme is widely used for handling qualities analysis and has been proven effectively over many years. For the evaluation of visual systems for simulation, the handling quality is not the primary focus. Instead, the flight task performance in different configurations was assessed and compared with each other to allow for identifying the visual system's impact on flight performance.
- **Application Remark:** The use of standardized flight tasks enables comparability, allowing the effects of using MR headsets in helicopter training simulation to be accurately measured and ensuring that the results are reliable and reproducible across different studies. When used, the pilots should be sufficiently briefed to fully understand the individual flight task's objectives and visual cues.
- **Recommendation:** *Using pre-defined flight tasks, such as those from ADS-33, are recommended for comparing different visual systems, as it produces reliable and comparable flight performance data.*

4.4.3 Correlation of PQ With MTE Flight Performance

A novel metric for evaluating a novel vision system, such as the MR headset, has been created by linking flight performance assessments gathered with MTE ratings to presence assessments collected with PQ ratings (Section 3.6.2).

- **Ease of Use:** The proposed correlation metric is based on data collected with the PQ and MTE flight maneuvers. Regarding the ease of use of these metric, see Section 4.4.1 and Section 4.4.2, respectively. The derivation of correlation is conducted post-hoc, therefore not consuming additional time during studies. The mathematical operations required for mapping MTE ratings to the PQ scale and performing a regression to determine the correlation can be executed automatically by a suitable software. In this thesis, an evaluation script was developed in MATLAB for that purpose.
- **Accuracy and Quality of the Data:** The MTE score has shown a correlation with the PQ score, indicating that it can be used to combine both subjective (reported presence) and objective (flight performance data) results.
- **Grading Scale and Categorization:** To derive the correlation metric, the flight performance levels defined in ADS-33 as *not adequate*, *adequate* and *desired* are first translated to numerical values 0 - 2. The numerical values are then mapped to the PQ's value range 1 -7, allowing for straightforward comparison of individual presence ratings to flight performance results. The mapping in this thesis was conducted using a piece-wise linear mapping, however other approaches may be implemented.

- **Application Remark:** The correlation metric has demonstrated potential in evaluating visual systems to be used in helicopter training simulator qualification. It would be beneficial to further investigate merging both scores (PQ and MTE) into one single score, which could be explored in future studies. Additionally, incorporating the SSQ into the PQ / MTE score could provide a more comprehensive evaluation tool. To further investigate the correlation, the following aspects can be explored:
  - By using both metrics and applying the correlation, the validity of the reported PQ ratings can be assessed using the objective flight performance to identify invalid PQ scores provided by pilots struggling with the questionnaire.
  - If presence and flight performance are correlated and the nature of the correlation is well-known, it may be possible to quit using PQ in future evaluations in favor of only using the objective flight performance metric, thereby saving time.
- **Recommendation:** *Correlating MTE scores to PQ ratings already provides information about the user experience and pilot performance of the MR simulation and is therefore recommended to be used for combining subjective and objective analyses.*

4.4.4 Assessment Using Proposed Flight Tasks

The utilization of standardized, predefined flight tasks over an extended period of time, together with MTEs allows for comparison of flight performance differences between visual systems and across pilots. By employing the same tasks for all assessments, an accurate evaluation of the impact of MR visual systems on the simulation of helicopter pilot training can be obtained.

- **Ease of Use:** The discussion of ease of use for MTEs was already provided in Section 4.4.2. In summary, comprehensive briefing and some customization may be required for pilots unfamiliar with the flight tasks to execute them properly.
  The other elements of the proposed flight tasks consist of standard maneuvers included in standard proficiency checks, which every pilot is expected to perform during their flight training.
- **Accuracy and Quality of the Data:** Providing the pilots with precise instructions on flight parameters like velocity, altitude, and climb rate allows for collecting comparable data over extended periods of time. The combination of heterogeneous flight profiles ensures prompting various pilot behavior, regarding for example their control strategies and gaze behavior, resulting in comprehensive assessment scenarios.
  In addition, the flight state data recorded for task elements like MTEs may be utilized for flight performance evaluation, creating objective datasets suitable for comparative analyses.

- **Grading Scale and Categorization:** Three stages of flight tasks cover a comprehensive range of aspects related to helicopter training, including:

  1. **Hands-on flying tasks** (hovering and slalom maneuvers) to assess pilot performance under demanding flight conditions.
  2. **Cockpit interaction using malfunction/emergency procedures** with the MR headset, evaluating usability and accessibility of cockpit instruments while wearing the device.
  3. **Ground and Arial Work tasks** that simulate realistic training scenarios in helicopter pilot training (Section 4.1).

  For each of the stages, different grading scales might apply. For the MTE flight tasks, the flight performance assessment of ADS-33 can be applied (Section 4.4.2), while for malfunction procedures the time needed for detection of cues and interaction with the cockpit might be evaluated.

- **Application Remark:** The proposed flight tasks aim at covering both typical helicopter flight as well as demanding specialized maneuvers to provide comprehensive insights into the characteristics of MR setups. Based on the targeted use case for a specific simulator, it might be sufficient to focus on distinct aspects of flight. For example, the performance in demanding maneuvers might be of less importance for a MR simulator intended for training of cockpit procedures only.

- **Recommendation:** *It is strongly recommended to utilize long, predefined flight tasks in addition to short task elements for the evaluation of MR in helicopter simulation. Extended flight tasks with heterogeneous scope of training may better uncover factors such as susceptibility to simulation sickness, while also covering various aspects of training more comprehensively compared to short, homogeneous task elements.*

4.4.5 Assessment With Eye Tracking

Eye tracking data can provide valuable insights into changes of pilot gaze behavior in simulation due to varying visual systems. Based on the raw gaze data recorded with eye tracking hardware a variety of metrics can be derived, see Section 3.1.3. For the studies in this thesis, area of interest was assessed utilizing the ratio between in-the-cockpit (ITC) and out-the-window (OTW) gazes. For the determination of ITC/OTW ratios, the Focus Distance Distribution method was used.

- **Ease of Use:** For recording of eye tracking data, specialized hardware needs to be employed. For generating baseline data in conventional simulators pilots need to wear eye tracking glasses, while in MR setups, the utilized headset may allow for low-effort collection of eye tracking data. However it was found in the studies, that

data collection did not always work reliably, resulting in the necessity to calibrate the system for each pilot, which consumes time.
The FDD method applied for extraction of ITC/OTW ratios is computationally expensive compared to other approaches e.g., using a static threshold for distinction of ITC to OTW, however as the analysis is carried out after completion of a study, this does not effect ease of use.

- **Accuracy and Quality of the Data:** The two eye tracking systems used in this thesis both provide a spatial accuracy of the measured gaze angles of approximately 1°. While this may be acceptable for metrics directly using the gaze vectors, the limited accuracy results in large errors for derived eye tracking metrics that are based on 3D projections of the gaze angles, as for example the ITC/OTW ratio. Analysis tools should account for the measurement errors, this is why the FDD method was chosen for evaluation of ITC/OTW ratios.
- **Grading Scale and Categorization:** Based on raw eye tracking data, a variety of derived metrics and associated grading scales can be applied. In this thesis, ITC/OTW ratio was chosen as it is crucial for understanding pilots' scanning patterns. It is especially valuable for evaluation of MR configuration as it allows to identify if the setup impacts the pilots' behavior regarding reading cockpit instruments and scanning for air traffic.
- **Application Remark:** Recording of eye tracking data in the conducted studies was partly unreliable resulting in data with offsets or no data at all. It is advised to confirm robust operation of the eye tracking setup in advance. In addition, calibrations should be conducted by asking each pilot to focus several points with known positions to be able to compensate for potential bias errors in the recorded data during post-processing.
- **Recommendation:** *Analysis of eye tracking data is strongly recommended due to the substantial insights provided by the data. To ensure robust data recording, the setup should be thoroughly calibrated. Based on the specific use case the user should carefully select which assessment metric to use from the multiple options that are provided by raw eye tracking data. The ITC/OTW ratio was found useful as its output allows for easy interpretation and provides valuable insights. However, for a in-depth analysis, more sophisticated approaches may be used, e.g. utilizing object recognition to identify which individual instrument a pilot is focusing at a time.*

4.4.6 Answer to Subordinate Research Question (1.3)

Based on these findings, the third subordinate Research Question (1.3) can be answered.

(1.3) How can new systems, such as MR headsets, be evaluated to assess the quality of a visual simulation?

In this thesis, various assessment methods for the evaluation of a MR headset's impact on simulation quality were investigated. Based on the results, the following **Effective Methods for Assessment of MR Simulation Quality** were identified:

- **Subjective Questionnaires:** The standardized questionnaires PQ and SSQ gather pilots' feedback on their experience with the MR simulation. They allow for comparison of subjective ratings between pilots and configurations, while their subscales in conjunction with pilot comments enable to identify specific aspects of an VE that require adjustments.
- **Flight Performance Evaluation:** Assessing how well pilots perform specific flight tasks or missions in the simulated environment provides an objective measure of simulation quality. The use of well-defined maneuvers increases repeatability and enables direct comparison between different simulation configurations.
- **Proposed Flight Tasks:** In this thesis, specific flight tasks for the evaluation of MR visual systems were developed. They combine standard procedures known from flight training with demanding and standardized maneuvers. The tasks have proven effective in prompting a higher bandwidth of pilot behavior compared to isolated tasks, allowing for a more comprehensive assessment of the simulation setup.
- **Combination of Subjective and Objective Metrics:** Combining assessments from subjective and objective metrics provides a more comprehensive assessment of flight performance in correlation with presence and allows for validating individual metrics' results against each other. In addition, if the correlation between two metrics is well-known, issues in reporting or data collection may be identified. For example, if a pilot's reported presence ratings do not match their flight performance, the pilot might be struggling with the presence questionnaire.
- **Eye Tracking Analysis:** Analyzing gaze patterns can reveal valuable insights into how pilots interact with the simulated environment. In addition, comparing eye tracking data between different simulator configurations can uncover differences in gaze behavior and the impact on the training of gaze patterns.

5 Overall Assessment and Knowledge Transfer

The scientific knowledge gained by answering the subordinate research questions forms a solid basis on which the main research question can be answered in the following. Subsequently, the transfer of knowledge generated in this thesis for the development of a MR in a industrial setting is discussed.

5.1 Answering the Main Research Question

Based on the investigations carried out in this thesis and the results gathered throughout the studies, the Main Research Question (1) can be answered.

(1) Can a MR headset serve as a suitable visual system for training helicopter pilots?

Yes, a MR headset can serve as an effective visual system for training helicopter pilots. Consideration should be given to the following aspects:

- **Presence:** MR provides a higher level of presence than conventional simulators utilizing TV screens, but falls short of the presence offered by simulators with fully equipped cockpits using projection systems.
- **Flight Performance:** Performance with MR is comparable to that of a basic mock-up simulator with TV screens, though it lags behind full-flight simulators.
- **Simulation Sickness:** Although MR simulation is more prone to cause simulation sickness, this effect can be mitigated through training and technical adjustments.
- **Gaze Behavior:** Minor differences in gaze behavior between MR and traditional visual systems exist, their occurrence depends on the individual flight task.
- **Comfort:** The weight of MR headsets is comparable to pilots' helmets, but its uneven weight distribution may cause discomfort.

In summary, the MR headset currently cannot fully replicate the presence and comfort offered by a projection dome, but it holds the potential to enhance or complement training conducted in FFSs. Continued advances in hardware, software and regulatory standards will determine the extent of MR adoption in helicopter-pilot training.

This thesis demonstrates that MR simulators are a viable option for helicopter training, but further development is needed to enhance their effectiveness. By optimizing critical factors such as presence, simulator sickness and flight performance, MR simulators can be improved to approach the quality of Full Flight Simulators. Given the rapid evolution of MR technology, it's likely that current limitations will soon be overcome. In addition, MR-based simulation platforms offer greater flexibility due to their compact design, allowing for more frequent training sessions on the ground as more simulators can be accommodated in the space of a single FFS. Moreover, they can be easily transported to different training facilities to train pilots, thereby increasing accessibility and reducing the need for dedicated training facilities. This makes them an attractive alternative to traditional FFS systems.

The results and findings in this thesis are subject to limitations primarily due to the small size of participants in the conducted studies (three pilots). Nevertheless, the combination of various assessment methods, with both subjective and objective metrics, allows for meaningful interpretation of the collected data. Due to the high effort of pilot evaluations similar number of participants are common in the helicopter research community.

The small number of pilots also limits the applicability of statistical evaluation (e.g. correlation analysis between presence and flight performance). Nevertheless, the results of this work reveal areas of interest for future research that may be further investigated in focused analyses with higher numbers of participants.
Overall, the studies highlight useful evaluation approaches, important factors to consider and the specific limitations of MR technology that should be taken into account in more extensive investigations (see the next Section 5.2).

5.2 Knowledge Transfer for Industrial Applications

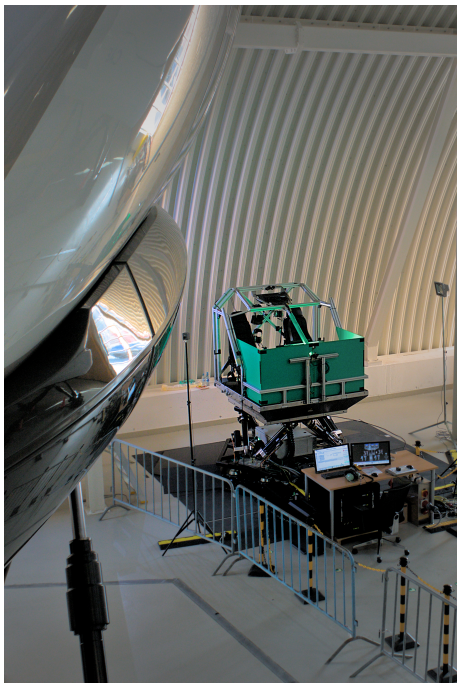
Studies 1 (Section 3.1) and 3 (Section 4.2) conducted for this thesis were an integral element of the research project *AR/XR in Training Simulation* (ARiTra) funded by the German Federal Ministry of Economic Affairs and Climate Action (BMWK) within the German Federal Aeronautical Research Program (LuFo VI-2). The ARiTra project was carried out from 2022 to 2025 with the objective to assess the potential of AR and MR technologies in helicopter pilot training and to identify requirements for qualification.
For this purpose, the industrial project partner REISER Simulation and Training GmbH built a helicopter flight simulator utilizing MR technology. This prototype MR simulator, named ARiTra Simulator, consists of a motion platform with a H145 cockpit mockup with two seats, one being equipped with a CLS for the pilot (see Figure 5.1a). For the vision system, the Varjo XR-3 headset was used. The ARiTra simulator was compared to a qualified FFS within a study. Nine pilots took part in this campaign.

This thesis provided the scientific knowledge base (Section 1.2.1) for the ARiTra project and made following contributions to the MR ARiTra Simulator and the conducted comparative study:

- Reference implementation of MR in 2PASD
- Argument for selection of the monitor-based headset Varjo XR-3
- Application of chroma keying
- Definition of flight tasks for assessment of a MR simulator
- Evaluation of applicability of assessment methods

An important step was creating **reference implementations** for Varjo XR-3 and JVC headsets into DLR's 2PASD, which contributed to the knowledge and expertise for the development of the MR-capable ARiTra simulator (see also [61]). The **decision to utilize the Varjo XR-3 Focal Edition headset** instead of the JVC headset for the ARiTra simulator has been made on the basis of the experiences gathered with the reference implementation in 2PASD in Study 1 (Section 3.2).

The recommendation to use both masking and **chroma keying** for the presentation of a mixed reality environment was taken into account due to outcome of Study 1. This led to a green screen being installed at the ARiTra Simulator (Figure 5.2). The subsequent software adjustments were made accordingly.
Figure 5.1b shows an example of the resulting VE: the pilot's feet are visible instead of being cut off by a static mask, while the VE with view to the ground remains visible through the chin window. This differentiation can be achieved using chroma keying if non-stationary real components are to be integrated into the VE. The use of green screen technology significantly improved the MR experience of the ARiTra Simulator.



(a) ARiTra Simulator
©DLR. All rights reserved.



(b) MR View in ARiTra Simulator Cockpit
©REISER Simulation and Training GmbH.
All rights reserved.

Figure 5.1: ARiTra Simulator

The **flight tasks for assessment of a MR simulator** developed in this thesis (Section 4.1) were used in the comparative study of the ARiTra project.
After a short familiarization flight, the pilots flew the two tasks Ground Work and Aerial Work on both the ARiTra simulator and the baseline FFS. The original Ground Work task proposed in Section 4.2 was slightly adapted due to feedback gathered in the pre-study validation of the ARiTra simulator: After flying both maneuvers, an operational pilot stated it would be beneficial to begin the Ground Work Task with the traffic circuit instead of ending with it. This would rearrange the task elements in increasing effort, allowing the pilots to conduct the comparatively easy traffic circuit task when they are more comfortable and less susceptible to simulation sickness. Afterwards they would tackle the more challenging maneuvers close to the ground, which are more likely to induce fatigue and simulation sickness symptoms to the pilots. The ARiTra comparative study was conducted utilizing the rearranged procedure. The updated flight task is illustrated in Figure 5.3.

The **assessment methods** proposed in Section 4.4 were used for evaluation of the ARiTra simulator. Each task was evaluated by the pilots using standardized questionnaires including PQ and SSQ. In addition, eye tracking and flight state data were recorded during this study.

The results of this theses positively impacted the qualification process of a MR helicopter simulator.

REISER Simulation and Training GmbH, in collaboration with Varjo, has announced the pre-qualification of the REISER H145 D3 Mixed Reality Helicopter Simulator by the German Federal Aviation Office (LBA) [94].

The results of Study 1 (Section 3.2) led to a technological change in 2025, where the JVC headsets were replaced with Varjo headsets in the REISER H145 D3 Mixed Reality Helicopter Simulator. The LBA was presented with the research results of Study 1 [61] demonstrating the benefits of the Varjo headset. The results thus directly assisted with the qualification process [95].



Figure 5.2: ARiTra Simulator Cockpit with Green Screen ©DLR. All rights reserved.

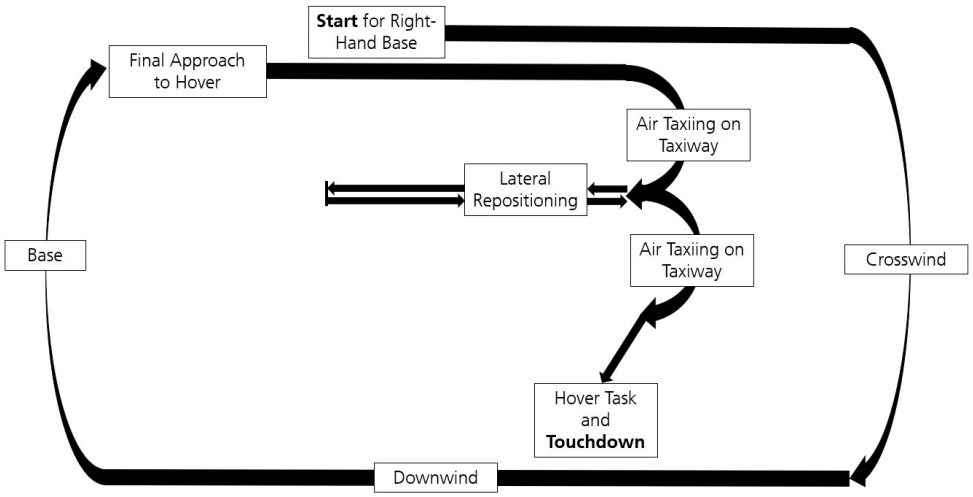


Figure 5.3: Updated Ground Work Task ©DLR. CC BY-NC-ND

6 Conclusion and Outlook

This chapter concludes this thesis on application of MR technology in helicopter training. First, a brief summary of the research objectives and results gathered from three studies is presented. Finally, an outlook on future research is given.

6.1 Conclusion

The primary goal of this dissertation was to investigate the effectiveness of MR headsets as a visual system for helicopter pilot training, addressing whether MR can improve or complement training experiences compared to traditional simulators. The thesis was structured into four key objectives:

- **Comparison of MR and AR headsets** (monitor-based vs. see-through)
- **Assessment of flight performance, presence and well-being** (including simulation sickness and gaze behavior)
- **Evaluation of assessment methods** to determine their applicability in MR-based training
- **Proposal of a standardized flight task** for future comparisons between MR and conventional visual systems

Summary of Study Results

To answer the research questions, this dissertation conducted three studies using two distinct helicopter simulators: the 2PASD and the AVES. A broad evaluation approach was adopted, combining flight performance evaluation, presence and sickness questionnaires and eye tracking data.

Study 1: Comparison of MR Configurations in 2PASD

The first study focused on comparing different MR configurations, specifically, monitor-based (Varjo XR-3) and see-through (JVC) headsets, to determine their suitability for helicopter training.

Key findings include:

- **Presence:** The MR configuration with the Varjo headset received the highest presence ratings, indicating that pilots perceived it as more immersive than the JVC AR configuration and the CON with TV screen setup.
- **Flight Performance:** No significant differences were observed in pilot performance between all configuration (VR, MR with the Varjo headset, AR with the JVC headset and CON with TV screens), indicating that all of them could be viable for basic training tasks.
- **Simulation Sickness:** The pilots exhibited no sickness symptoms with the MR configuration, except for one pilot experiencing slight symptoms caused by an unintentional rapid maneuver, which lead to fast movements in the headset display.
- **Technical Challenges:**
  - **Varjo monitor-based headset:** Blurred images during fast head movements, chroma-keying artifacts and reflections from illuminated surfaces.
  - **JVC see-through headset:** Reflections, brightness issues, insufficient opacity (real-world shimmering through the virtual environment) and a small FOV, which reduced immersion

Because of its superiority in presence, negligible sickness symptoms and comparable flight performance, the Varjo MR configuration was selected for further investigation. Some of its technical drawbacks can be mitigated with improved lighting and the use of cut-outs (masks) in addition to chroma keying.

Study 2: Investigation of MR in a Fully Equipped Cockpit

The second study involved integrating the Varjo headset into the fully equipped cockpit of the AVES, which allows for full cockpit interaction (e.g., switch engagement, instrument manipulation).

Key optimizations and findings include:

- **Hybrid MR Approach:** To mitigate chroma keying artifacts and reflections, cutouts (masks) were implemented, blocking virtual elements from overlaying real-world objects in the cockpit area while preserving the external view.
- **Cockpit Interaction:** Unlike the 2PASD, the AVES provides full cockpit engagement and haptic feedback. Malfunction procedures that require knob turning and use of the navigation system were feasible with the Varjo MR configuration.
- **Usage Challenges:** FOV limitations restrict pilot perception of co-pilot actions and peripheral airspace, increasing the need for head movements.

- **Technical Challenges:** The closed-cockpit environment constrains placement of head-tracking devices and illumination sources, requiring careful lighting calibration to avoid visual distortions.

Study 3: Proposing a Standard Maneuver for Evaluation of New Visual Systems for Helicopter Simulation

The third study assessed the applicability of evaluation methods (e.g., presence questionnaires, flight performance metrics, eye tracking) and proposed standardized maneuvers for future MR simulator assessments.

Key insights include:

- **Standardized Assessment Methods:** Subjective Questionnaires (PQ, SSQ) measured the pilots' perceived presence and simulation sickness symptoms.
- **Flight Performance Evaluation:** Provides an objective assessment of the pilots' flight performance.
- **Proposed Flight Task:** Two specific maneuvers based on the annual proficiency check for pilots were developed to compare MR systems against conventional simulators. The combined flight task's properties allow for a more comprehensive assessment of simulation setups.
- **Combined Subjective and Objective Metrics:** A positive correlation between presence ratings and flight performance was identified allowing for mutual validation of the individual metrics.
- **Eye Tracking Analysis:** Pilots focused more on external visual cues (terrain, obstacles) when in MR configuration compared to using a conventional visual system, indicating that MR setups influence pilots' gaze behavior.

The MR headset has potential to enhance or complement training of a FFS, but cannot fully replicate the presence and comfort offered by a projection dome. However, ongoing developments in hardware, software and regulatory frameworks suggest that MR could become a valuable technology for enhancing helicopter pilot training in simulation.

6.2 Outlook

The use of MR technology in pilot training has shown promising results in this thesis, particularly in single-pilot operation. However, to fully leverage the potential of MR technology, it is essential to explore its integration into Multi-Crew-Cooperation (MCC) environments. Studies like [96] have shown the effectiveness of communication, teamwork and specialized training for crew members in simulation. To further enhance the training capabilities of MR simulators for helicopter pilot training, MCC with MR headsets should be investigated.

Future investigations should focus on exploring the effective integration of MR technology into MCC environments. It is essential to address the challenges associated with MR technology, including the difficulty of face-to-face communication and particularly the restricted FOV, which can hinder awareness as well as interactions between pilots and cockpit elements (Figure 6.1).

Furthermore, the integrated eye tracking of the MR headset can be used for providing direct feedback to the flight instructor on how a student pilot applies gaze patterns, which could improve flight training effectiveness. This could provide a more comprehensive understanding of the student pilot's visual attention and awareness during training.

In conclusion, the potential of MR technology for pilot training is significant, but it also comes with challenges that need to be further addressed. By investigating the effective integration of MR technology into MCC environments, significant progress can be made in improving not only single pilot training but also MCC.



Figure 6.1: MCC Application with MR ©DLR. All rights reserved.

Appendix A

Standardized Questionnaires

A.1 Presence Questionnaire [65]

Items		Factors	Subscale
1.	How much were you able to control events?	CF	INV/C
2.	How responsive was the environment to actions that you initiated (or performed)?	CF	INV/C
3.	How natural did your interactions with the environment seem?	CF	NATRL
4.	How completely were all of your senses engaged?	SF	
5.	How much did the visual aspects of the environment involve you?	SF	INV/C
6.	How much did the auditory aspects of the environment involve you?	SF	NATRL
7.	How natural was the mechanism which controlled movement through the environment?	CF	NATRL
8.	How aware were you of events occurring in the real world around you?	DF	
9.	How aware were you of your display and control devices?	DF	
10.	How compelling was your sense of objects moving through space?	SF	INV/C
11.	How inconsistent or disconnected was the information coming from your various senses?	RF	
12.	How much did your experiences in the virtual environment seem consistent with your real-world experiences?	RF,CF	NATRL
13.	Were you able to anticipate what would happen next in response to the actions that you performed?	CF	INV/C
14.	How completely were you able to actively survey or search the environment using vision?	RF,CF,SF	INV/C
15.	How well could you identify sounds?	RF,SF	AUD
16.	How well could you localize sounds?	RF,SF	AUD
17.	How well could you actively survey or search the virtual environment using touch?	RF,SF	HAPTC

18.	How compelling was your sense of moving around inside the virtual environment?	SF	INV/C
19.	How closely were you able to examine objects?	SF	RESOL
20.	How well could you examine objects from multiple viewpoints?	SF	RESOL
21.	How well could you move or manipulate objects in the virtual environment?	CF	HAPTC
22.	To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session?	RF	
23.	How involved were you in the virtual environment experience?		INV/C
24.	How distracting was the control mechanism?	DF	
25.	How much delay did you experience between your actions and expected outcomes?	CF	INV/C
26.	How quickly did you adjust to the virtual environment experience?	CF	INV/C
27.	How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	CF	INV/C
28.	How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?	DF	IFQUAL
29.	How much did the control devices interfere with the performance of assigned tasks or with other activities?	DF,CF	IFQUAL
30.	How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?	DF	
31.	Did you learn new techniques that enabled you to improve your performance?	CF	
32.	Were you involved in the experimental task to the extent that you lost trackof time?		INV/C

Major Factor Category: CF = Control Factors, SF = Sensory Factors, DF = Distraction Factors, RF = Realism Factors. Subscales: INV/C = Involvement/Control, NAT = Natural, AUD = Auditory, HAPTC = Haptic, RES = Resolution, IFQUAL = Interface Quality.

A.2 Slater, Usoh, and Steed (SUS) Questionnaire (A Presence Questionnaire) [67]

1. Please Rate your sense of being in the virtual environment, on the following scale from 1 to 7, where 7 presents your normal experience of being in a place.

   I had a sense of "being there" in the virtual environment:

Not at all.						Very much.

2. To what extent were there times during the experience when the virtual environment was the reality for you?

   There were times during the experience when the virtual environment was the reality for me. . .

At no time.						Almost all the time.

3. When you think back about your experience, do you think of the virtual environment more as images that you saw, or more as somewhere that you visited?

   The virtual environment seems to me to be more like. . .

Images I saw.						Somewhere that I visited.

4. During the time of the experience, which was strongest on the whole, your sense of being in the office space, or of being elsewhere?

   I had a stronger sense of. . .

Being elsewhere.						Being in the virtual environment

5. Consider your memory of being in the virtual environment. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By "structure of the memory," consider things like the extent to which you have a visual memory of the virtual environment, whether that memory is in color, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements.

   I think of the virtual environment as a place in a way similar to other places that I've been today. . .

Not at all.						Very much so.

6. During the time of the experience, did you often think to yourself that you were actually in the virtual environment?

   During the experience I often thought that I was really standing in the virtual environment. . .

Not very often.						Very much so.

A.3 Simulation Sickness Questionnaire

Table A.2: Simulation Sickness Questionnaire and the calculation of nausea (N), oculomotor disturbance (O) and disorientation (D) based on [71] and for calculation clarification [92].

SSQ Symptoms	N	O	D
General discomfort	1	1	
Fatigue		1	
Headache		1	
Eyestrain		1	
Difficulty focusing		1	1
Increased salvation	1		
Sweating	1		
Nausea	1		1
Difficulty concentrating	1	1	
Fullness of head			1
Blurred vision		1	1
Dizzy (eye open)			1
Dizzy (eye closed)			1
Vertigo			1
Stomach awareness	1		
Burping	1		
Total	[1]	[2]	[3]

$$N = [1] \cdot 9.54 \tag{A.1}$$

$$O = [2] \cdot 7.58 \tag{A.2}$$

$$D = [3] \cdot 13.92 \tag{A.3}$$

$$TS = ([1] + [2] + [3]) \cdot 3.74 \tag{A.4}$$

Bibliography

[1] L. Allen. Evolution of flight simulation. *American Institute of Aeronautics and Astronautics*, AIAA-93-3545-CP, August 1993.

[2] Lufthansa-Bildarchiv CHD761-3-09. LINK Trainer.

[3] Max Baarspul. A review of flight simulation techniques. *Progress in Aerospace Sciences*, 27(1):1–120, January 1990.

[4] D. Solar and C. Audard. Rotorcraft safety roadmap. Technical report, EASA, December 2018.

[5] Glen Ross and Andrew Gilbey. Extended reality (xR) flight simulators as an adjunct to traditional flight training methods: a scoping review. *CEAS Aeronautical Journal*, 14(4):799–815, October 2023.

[6] Anon. Certification Specifications for Helicopter Flight Simulation Training Devices. EASA Annex, EASA, June 2012.

[7] European Union Aviation Safety Agency (EASA). Easy Access Rules for Helicopter Flight Simulation Training Devices (CS-FSTD(H)) (Initial issue). Technical Report Initial issue, July 2020.

[8] Anon. 14 CFR Part 60 - Flight simulation training device initial and continuing qualification and use. Technical report, Federal Aviation Administration (FAA), 2026.

[9] Anon. FAA Approval of Aviation Training Devices and Their Use for Training and Experience. Advisory Circular 61-136B, U.S. Department of Transportation, Federal Aviation Administration, December 2018.

[10] Anon. FSTD Special Condition fur the use of Head Mounted Displays (HMD) combined with a motion platform with reduced envelope. EASA Information Paper, EASA, March 2023.

[11] Michael Gordon Whitaker. Integration of powered-lift: Pilot certification and operations; miscellaneous amendments related to rotorcraft and airplanes. Technical report, Federal Aviation Administration (FAA), Department of Transportation, October 2024.

[12] Anon. Flight Simulation Training Device (FSTD) Certification Guidance Document. Technical Report Version 1.0, ICAO NACC Regional Office, May 2024.

[13] Anon. Easy acess rules for flight crew licencing (Part-FCL). EASA eRules: aviation rules for the 21st century, Federal Aviation Administration (FAA), August 2020.

[14] Loft Dynamics. EASA's new ruling drops FFS requirement for helicopter training and checking, prioritizes Loft Dynamics VR FSTD. https://www.loftdynamics.com/easas-new-ruling-drops-ffs-requirement-for-helicopter-training-and-checking-prioritizes-loft-dynamics-vr-fstd/. Accessed 2026-01-05.

[15] Anon. 14 CFR part 61: Certification: Pilots, flight instructors, and ground Instructors). Special federal aviation regulation NO. 100–2, Federal Aviation Administration (FAA), April 1997.

[16] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. In *Photonics for Industrial Applications*, pages 282–292, 1995.

[17] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems,*, Vol E77-D(No.12), December 1994.

[18] Johannes M. Ernst, Tim Laudien, and Sven Schmerwitz. Implementation of a mixed-reality flight simulator: blending real and virtual with a video-see-through head-mounted display. In *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications V*, page 23, Orlando, United States, June 2023. SPIE.

[19] Andrew Yeh Ching Nee and Soh Khim Ong, editors. *Springer Handbook of Augmented Reality*. Springer Handbooks. Springer International Publishing, Cham, 2023.

[20] Johannes Maria Ernst. *Creating a Virtual Helicopter Cockpit with an Immersive Head-Mounted Display*. DLR Forschungsbericht, DLR-FB-2023-21, TU Braunschweig, 2023.

[21] David M. Johnson. Introduction to and review of simulator sickness research. Technical report, American Psychological Association, 2005.

[22] Anon. NATO - Guidelines for mitigation cybersickness. STO TECHNICAL REPORT AC/323(HFM-MSG-323)TP/1039, NATO, October 2021.

[23] M. Park, S. Serefoglou, L. Schmidt, K. Radermacher, C. Schlick, and H. Luczak. Hand-eye coordination using a video see-through augmented reality system. *The Ergonomics Open Journal*, 1(1):46–53, October 2008.

[24] Jonathan W. Kelly. Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics. *IEEE Transactions on Visualization and Computer Graphics*, 29(12):4978–4989, December 2023.

[25] Joong Ho Lee and Ji-Hyung Park. Visuomotor adaptation to excessive visual displacement in video see-through HMDs. *Virtual Reality*, 24(2):211–221, June 2020.

[26] Yan Yan, Ke Chen, Yu Xie, Yiming Song, and Yonghong Liu. The effects of weight on comfort of virtual reality devices. In *Advances in Ergonomics in Design*, volume 777, pages 239–248. Springer International Publishing, Cham, 2019. Series Title: Advances in Intelligent Systems and Computing.

[27] Thibault Louis, Jocelyne Troccaz, Amélie Rochet-Capellan, and François Bérard. Is it real? Measuring the effect of resolution, latency, frame rate and jitter on the presence of virtual entities. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pages 5–16, Daejeon Republic of Korea, November 2019. ACM.

[28] David Urban and Ryan Pritchard. Beyond Illusions: Navigating VR Fidelity in Undergraduate Pilot Training - A 3-Year Data Analysis. In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, volume Paper No. 24378, 2024.

[29] Ylli Dalladaku, Jacob Kelley, Brycen Lacey, James Mitchiner, Braden Welsh, and Matthew Beigh. Assessing the effectiveness of virtual reality in the training of army aviators. In *Proceedings of the 2020 Annual General Donald R. Keith Memorial Capstone Conference*, West Point, New York, USA, April 2020.

[30] Matthias Oberhauser, Daniel Dreyer, Reinhard Braunstingl, and Ioana Koglbauer. What's real about virtual reality flight simulation?: Comparing the fidelity of a virtual reality with a conventional flight simulation environment. *Aviation Psychology and Applied Human Factors*, 8(1):22–34, March 2018.

[31] Talon M Pope. *A cost-benefit analysis of pilot training next*. Thesis, Air Force Institute of Technology Air Force Institute of Technology, March 2019.

[32] Travis H Sheets and Matthew P Elmore. Abstract to Action: Targeted Learning System Theory Applied to Adaptive Flight Training. *Air command and staff college*, April 2018.

[33] Agata Lawrynczyk. *Exploring virtual reality flight training as a viable alternative to traditional simulator flight training*. Thesis, Carleton University, Ottawa, Ontario, 2018.

[34] Stefan Auer, Jens Gerken, Harald Reiterer, and Hans-Christian Jetter. Comparison between virtual reality and physical flight simulators for cockpit familiarization. In *Proceedings of Mensch und Computer 2021*, pages 378–392, Ingolstadt Germany, September 2021. ACM.

[35] Edward Bachelder, David H. Klyde, Noah Brickman, Sofia Apreleva, and Bruce Cogan. Fused reality for enhanced flight test capabilities. In *AIAA Atmospheric Flight Mechanics (AFM) Conference*. American Institute of Aeronautics and Astronautics, 2013.

[36] David H. Klyde, Justin R. Gray, and George Park. A mixed reality simulation tool for in-flight evaluations. In *AIAA Scitech 2020 Forum*, Orlando, FL, January 2020. American Institute of Aeronautics and Astronautics.

[37] Eugene F. Kasper, Loran A. Haworth, Zoltan P. Szoboszlay, Robert D. King, and Zsolt L. Halmos. Effects of in-flight field-of-view restriction on rotorcraft pilot head movement. In *SPIE*, volume Vol. 3058, pages 34–45, Orlando, FL, June 1997.

[38] Perry Comeau, Alanna Wall, and Andrew Law. Comparison of pilot workload between integrated reality in-flight simulation and flight test of helicopter landings on a frigate. In *Vertical Flight Society's 80th Annual Forum & Technology Display*, Montreal, Quebec, Canada, May 2024.

[39] Mark D. White and Gareth D. Padfield. Flight simulation in academia - Progress with HELIFLIGHT at the University of Liverpool.

[40] Sunjoo Edvani, Etienne va, Sander Bettendorf, Sunjoo Edvani, Etienne va, and Sander Bettendorf. SIMONA - A reconfigurable and versatile research facility. In *Modeling and Simulation Technologies Conference*, New Orleans,LA,U.S.A., August 1997. American Institute of Aeronautics and Astronautics.

[41] Boris Englebert and Guido Tillema. Objective evaluating short-stroke motion systems in helicopter flight simulation, January 2026.

[42] Gabriele Luzzani, Michael T Morcos, and Umberto Saetti. Multimodal cueing in precision hovering tasks: Predicting pilot cognitive workload via physiological measurements. In *51st European Rotorcraft Forum*, Venice, Italy, September 2025.

[43] First-in-Victoria spinning simulator to turn heads. https://www.rmit.edu.au/news/all-news/2024/feb/nova-simulator. Accessed 2026-04-25.

[44] Holger Duda, Sunjoo K. Advani, and Mario Potter. Design of the DLR AVES research flight simulator. In *AIAA Modeling and Simulation Technologies (MST) Conference*, Boston, MA, August 2013. American Institute of Aeronautics and Astronautics.

[45] Jürgen Kaletka, Hermann Kurscheid, and Ulrich Butter. FHS, the new research helicopter: Ready for service. *Aerospace Science and Technology*, 9(5):456–467, July 2005.

[46] Jürgen Gotschlich and Michael Jones. Online trimming of flight dynamics models using the 2simulate realtime simulation framework. In *Proceedings of AIAA Scitech 2018 Forum*. American Institute of Aeronautics and Astronautics, January 2018.

[47] D. M. Pitt and D. A. Peters. Theoretical prediction of dynamic-inflow derivatives. In *Sixth European Rotorcraft and Powered Lift Aircraft Forum*, Paper No. 47, Bristol, England, September 1980.

[48] The Khronos Group Inc. OpenGL: The industry's foundation for high performance graphics. https://www.opengl.org/. Accessed 2026-01-05.

[49] OpenSceneGraph. https://openscenegraph.github.io/openscenegraph.io/. Accessed 2026-01-05.

[50] Alexej Dikarew, Mario Müllhäuser, and Tanja Martini. Dual Pilot Active Sidestick Demonstrator 2PASD: Flexible rapid prototyping platform for research flight simulation. In *50th European Rotorcraft Forum*, Marseille, France, September 2024.

[51] Varjo. Varjo XR-3 - the first true mixed reality headset. https://varjo.com/products/varjo-xr-3/. Accessed 2024-03-20.

[52] JVC. Head mounted display HMD-VS1W. https://de.jvc.com/pro/headset/HMD-VS1W/. Accessed 2024-03-20.

[53] Jeremy N. Bailenson, Brian Beams, James Brown, Cyan DeVeaux, Eugy Han, Anna C. M. Queiroz, Rabindra Ratan, Monique Santoso, Tara Srirangarajan, Yujie Tao, and Portia Wang. Seeing the world through digital prisms: Psychological implications of passthrough video usage in mixed reality. *Technology, Mind, and Behavior*, 5(2):1–16, June 2024.

[54] Jürgen Gotschlich, Torsten Gerlach, and Umut Durak. 2Simulate: A distributed real-time simulation framework. *ASIM Mitteilung 149*, February 2014.

[55] QNX Software Systems. QNX Neutrino TROS Overview. https://www.qnx.com/products/intl/neutrino_rtos/. Accessed 2026-01-05.

[56] Open Robotics. ROS 2 Documentation: Foxy. https://docs.ros.org/en/foxy/index.html. Accessed 2026-01-05.

[57] VIVE. SteamVR basisstation 2.0. https://www.vive.com/de/accessory/base-station2/. Accessed 2024-03-20.

[58] Anthony Steed. A simple method for estimating the latency of interactive, real-time graphics simulations. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 123–129, Bordeaux France, October 2008. ACM.

[59] Ding He, Fuhu Liu, Dave Pape, Greg Dawe, and Dan Sandin. Video-Based Measurement of System Latency. *IPT 2000, International Immersive Projection Technology Workshop*, 2000.

[60] Mathworks. Matlab: Math. Graphics. Programming. https://www.mathworks.com/products/matlab.html. Accessed 2026-01-05.

[61] Tanja Martini. Evaluation of virtual and mixed reality technologies in helicopter simulation. In *Vertical Flight Society's 80th Annual Forum & Technology Display*, Montreal, Quebec, Canada, May 2024.

[62] Tanja Dikarew-Martini. Comparative analysis: See-through vs. monitor-based headsets in mixed reality helicopter simulation. In *American Helicopter Society Journal*. American Helicopter Society, 2026.

[63] Anon. ADS-33, Aeronautical Design Standard,Performance Specification, Handling Qualities Requirements for Military Rotorcraft. Technical report, US Army, March 2000.

[64] Guilherme Gonçalves, Hugo Coelho, Pedro Monteiro, Miguel Melo, and Maximino Bessa. Systematic review of comparative studies of the impact of realism in immersive virtual experiences. *ACM Computing Surveys*, 55(6):1–36, July 2023.

[65] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 1998.

[66] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, Glasgow Scotland Uk, May 2019. ACM.

[67] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9(5):497–503, October 2000.

[68] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, June 2001.

[69] Miguel Melo, Guilherme Gonçalves, josé Vasconcelos-Raposo, and Maximino Bessa. How much presence is enough? Qualitative scales for interpreting the Igroup presence questionnaire score. *IEEE Access*, 11:24675–24685, 2023.

[70] Eunhee Chang, Hyun Taek Kim, and Byounghyun Yoo. Virtual reality sickness: A review of causes and measurements. *International Journal of Human–Computer Interaction*, 36(17):1658–1682, October 2020.

[71] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, July 1993.

[72] Jelte E Bos, Scott N MacKinnon, and Anthony Patterson. Motion sickness symptoms in a ship motion simulator: Effects of inside, outside, and no view. *Aviation, Space, and Environmental Medicine*, Vol. 76, No.12, 2005.

[73] Peter J Gianaros, Eric R Muth, J Toby Mordkoff, Max E Levine, and Robert M Stern. A Questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviat Space Environ Med.*, 72(2), 2010.

[74] Kay M. Stanney, Robert S. Kennedy, and Julie M. Drexler. Cybersickness is not simulator sickness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 41(2):1138–1142, October 1997.

[75] U.S. Department of Defense. Handling qualities for military rotorcraft. Detail Specification MIL-DTL-32742, technical report, US Army, March 2023.

[76] David L Key, Chris L Blanken, Moffett Field, Roger H Hoh, David G Mitchell, and Bimal L Aponso. Background information and user's Guide (BIUG) for handling qualities requirements for military rotorcraft. Special report RDMR-AD-16-01, December 2015.

[77] Gareth D Padfield. *Helicopter flight dynamics*, volume Second Edition of *AIAA Education Series*. Blacksburg, Virginia, 2007.

[78] EGAST Leaflet Collision-Avoidance. Technical report, European Aviation Safety Agency (EASA), January 2010.

[79] Lin Cheng, Yuan-Chen Shen, Qiang He, and Min-Jie Zhang. Spying with a pilot's eye: Using eye tracking to investigate pilots' attention allocation and workload during helicopter autorotative gliding. *Heliyon*, 10(16):e35872, August 2024.

[80] Joseph Sullivan, Ji Hyun Yang, Michael Day, and Quinn Kennedy. Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine*, 82(9):871–878, September 2011.

[81] Christopher E. Kirby, Quinn Kennedy, and Ji Hyun Yang. Helicopter pilot scan techniques during low-altitude high-speed flight. *Aviation, Space, and Environmental Medicine*, 85(7):740–744, July 2014. Publisher: Aerospace Medical Association.

[82] Maxi Robinski and Michael Stein. Tracking visual scanning techniques in training simulation for helicopter landing. *Journal of Eye Movement Research*, 6(2), July 2013.

[83] Daniel H. Greiwe and Maik Friedrich. Gaze Movements of Helicopter Pilots during Real and Simulated Take-Off and Landing Maneuvers. *Aerospace MDPI*, 11(6):429, May 2024.

[84] Antony William Joseph and Ramaswamy Murugesh. Potential eye tracking metrics and indicators to measure cognitive load in human-computer interaction research. *Journal of scientific research*, 64(01):168–175, 2020.

[85] Tanja Dikarew-Martini and Alexej Dikarew-Martini. Introduction of focus distance distribution method for eye tracking evaluation in helicopter simulation. *Preprint*, 2026.

[86] T.E. Hutchinson, K.P. White, W.N. Martin, K.C. Reichert, and L.A. Frey. Human-computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1527–1534, December 1989.

[87] Anon. Tobii Pro Glasses 3 User Manual v1.10. Technical report, Tobii AB (pupl), July 2022.

[88] V. Onkhar, D. Dodou, and J. C. F. de Winter. Evaluating the Tobii Pro Glasses 2 and 3 in static and dynamic conditions. *Behavior Research Methods*, 56(5):4221–4238, August 2023.

[89] Varjo. Varjo Eye Tracking for Developer. https://developer.varjo.com/docs/get-started/eye-tracking-with-varjo-headset. Accessed 2025-04-11.

[90] Tanja Martini. Evaluation of pilot flight performance and cockpit interaction with a mixed reality headset in a helicopter simulator. In *50th European Rotorcraft Forum*, Marseille, France, September 2024.

[91] EUROCOPTER DEUTSCHLAND GmbH. EC135 T2+, Approved Rotorcraft Flight Manual. Technical report, February 2006.

[92] Pauline Bimberg, Tim Weissker, and Alexander Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 464–467, Atlanta, GA, USA, March 2020. IEEE.

[93] Tanja Dikarew-Martini and Malte-Jörn Maibach. AR/XR in der Trainingssimulation (ARiTra) Abschlussbericht DLR. Technical Report DLR-IB-FT-BS-2025-175, German Aerospace Center (DLR), 2025.

[94] Reiser Simulation and Training GmbH. REISER Unveils XR H145 D3 Simulator Mixed Reality Meets EASA-Compliant MCC Training. https://www.reiser-st.com/reiser-unveils-xr-h145-d3-simulator-mixed-reality-meets-easa-compliant-mcc-training/. Accessed 2026-01-05.

[95] Andrea Anderson. AR/XR in der Trainingssimulation (ARiTra) Abschlussbericht RST. Technical Report ARITRA_20Q2129A_ZB_003, Reiser Simulation and Training GmbH, 2025.

[96] Thomas Luebcke, Hendrik Wilhelm, Oda Schliebusch-Jacob, Tim Kreter, Tanja Martini, and Torsten Gerlach. Inter-professional simulator training for reliable ship-helicopter interactions. *Proceedings of the International ISCRAM Conference*, May 2024.