EVALUATION OF PILOT FLIGHT PERFORMANCE AND COCKPIT INTERACTION WITH A MIXED REALITY HEADSET IN A HELICOPTER SIMULATOR

Tanja Martini, German Aerospace Center (DLR), Institute of Flight Systems, Braunschweig, Germany

Abstract

The demand for pilots is expected to increase in the future. More efficient and cost-effective training methods are required to meet the growing demand. Currently, training is carried out both in a real helicopter and in simulation. However, simulation-based training with a mixed reality (MR) headset has the potential to complement the traditional methods by offering an affordable, movable training platform, and safe training environment for helicopter pilots. MR technology holds a unique advantage in helicopter simulator training, as it integrates physical controls and instruments, creating a more authentic and engaging experience that surpasses virtual reality alone in terms of realism and immersion. This enables pilots to practice complex procedures that require manual interaction with cockpit elements, such as pressing buttons or adjusting knobs, in a way that virtual reality simulations may not fully replicate. However, mixed reality technology also has its own set of challenges. This study investigates the benefits, drawbacks, and limitations of employing a mixed reality headset as a visual system within a helicopter simulation environment. To achieve this goal, a piloted simulation campaign was conducted at the German Aerospace Center's (DLR) Air Vehicle Simulator (AVES). A MR headset and a projection dome were utilized as the two visual systems under comparison. Pilots were instructed to perform standard flight maneuvers and malfunction procedures. The evaluation focuses on the comparison of the visual systems regarding presence, simulation sickness and flight performance.

1. INTRODUCTION

Forecasts show that the demand for helicopter services such as Helicopter Emergency Medical Services (HEMS), surveillance, police, aerial work, agriculture, and offshore operations will continue to grow (Ref. 1). However, the industry is facing a significant challenge: a shortage of qualified helicopter pilots. As experienced pilots retire, the profession is competing with other industries for talent, and access to flight training facilities is limited. To address this issue, providing efficient and affordable flight training opportunities could be an effective solution. In addition to regular helicopter flight training, simulators have been proven to be an effective, safe, and cost-efficient way to train pilots. This approach aligns with the European Aviation Safety Agency (EASA) Rotorcraft Safety Roadmap (Ref. 2), which emphasizes the development of new types of training devices. The current requirements for helicopter training devices in Europe are described in the certification specification for Flight Simulation Training Devices (FSTD) (Ref. 3) by EASA. However, as technology advances and pilot demands increase, these requirements will need to evolve. Two helicopter training devices have been certified by EASA: the Robinson R22 Beta II as FNPT II (Ref. 4) and the Airbus Helicopters H125 as FTD Level 3 (Ref. 5), both equipped with Virtual Reality (VR) headsets. EASA therefore has published the special conditions for using VR technology-based helicopter cockpit in the FSTD Special Conditions for the use of Head Mounted Display (HMD) combined with a motion platform with reduced envelope (Ref. 6). Moreover, research shows that VR and Mixed Reality (MR) technologies have the potential to increase training capability while decreasing costs (Ref. 7).

According to Milgram et al. (Ref. 8), VR is characterized as a state in which an individual is fully immersed within a synthetic environment, while the physical world is governed by the laws of physics. In contrast, MR represents a continuum of possibilities, where the boundaries between the real and virtual worlds could be blended in various degrees. MR and VR can be presented to an individual via two distinct headset technologies: "see-through" displays and "monitor-based" displays. With "see-through" displays, the user views their surroundings through the display while simultaneously perceiving virtual objects or environments as an overlay. In contrast, monitor-based displays concurrently display both virtual and real-world content, with the physical world captured by a camera.

Figure 1: Screenshot of Piloted Campaign in Air Vehicle Simulator (AVES) with Mixed Reality Setup

Research on VR and MR headsets has already gained significant attention for its potential applications for pilot training. As an example of the potential benefits of MR technology in aviation training, NASA and Systems Technology, Inc. conducted a flight test utilizing an MR headset to simulate an air-to-air refueling maneuver (Ref. 9). The pilot flew 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 by replacing the actual tanker aircraft with a simulated aircraft during the maneuver, thereby preventing potential collisions. Based on the findings from earlier in-flight experiments, further flight tests were carried out with the Fused Reality® Flight System in 2015 (Ref. 10). After a familiarization phase, the pilots were tasked with performing certain maneuvers, such as precision landings. Observations revealed that visual cueing issues arose due to the limited Field of View (FoV) presented by the VR system. This limitation had a direct impact on the pilot's ability to accurately execute tasks, highlighting the importance of optimizing FoV for optimal performance in virtual reality-based flight training and operations. VR is also utilized to create a proof of concept for novel cockpit designs. For instance, it has been employed to evaluate virtual cockpits and instruments for helicopter offshore operations in confined areas (Ref. 11). In addition, three studies have been conducted to assess flight decks during the early stages of design (Ref. 12), focusing on the pilot's interaction within the cockpit. It was concluded that VR is a valuable addition to the design process, however, there are also limitations associated with its use, such as when employing fully virtual buttons. This leads to slower movements by the pilot and an increased time required to complete tasks. Another study compared a VR flight simulator with a physical flight simulator for cockpit familiarization (Ref. 13). The results showed that participants were able to successfully complete check procedures in both simulators; however, some individuals reported symptoms of simulation sickness following use of the VR flight simulator. Furthermore, current VR/MR headsets are limited by their FoV. In the study of (Ref. 14), the effects of FoV restrictions on the head movement of rotorcraft pilots have been investigated. To achieve this, pilots performed various maneuvers while wearing HMD with different FOV settings. The findings shows 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 not visible, whereas a larger FoV allows the pilot to view visual cues in their peripheral vision. Furthermore, 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 (Ref. 15). Several rating scales were used in the study to quantify the pilots' subjective experiences of workload during the simulation-based tasks. Moreover, research has explored the physiological implications of using pass-through video in MR environments (Ref. 16). This study discusses the challenges associated with pass-through video, including distortion, latency, and simulator sickness. Despite these issues, it is predicted that pass-through video will become the standard for MR headsets over the next few years.

This research compares two visual system configurations used in a helicopter simulator: a monitor-based headset and a dome projection system with both setups being tested and evaluated in a piloted simulation campaign (Figure 1). The evaluation focuses on presence, simulation sickness and flight performance. This investigation builds upon previous work (Ref. 17) where two headsets, a monitor-based and a see-through headset, were compared to each other in a piloted simulation campaign at the Dual Pilot Active Sidestick Demonstrator (2PASD), which is a fixed base simulator with TV screens as visual system at German Aerospace Center (DLR) Braunschweig.

2. SYSTEM OVERVIEW

This section provides a detailed description of the helicopter facility, which served as the platform for the investigation. Furthermore, the experimental setup is described in detail. Additionally, this section provides details about the two visual systems utilized in the study.

2.1. Simulation Facility

The piloted study was conducted in the Air Vehicle Simulator (AVES) (Figure 2) at the DLR in Braunschweig (Ref. 18). AVES provides a fixed-based and a motion simulator with exchangeable cockpits that resemble an

EC135 and Airbus 320. For the campaign the EC135 cockpit on the fixed platform was used. The EC135 Cockpit is a replica of the cockpit of DLR's research helicopter Active Control Technologies / Flying Helicopter Simulator (ACT/FHS). The simulator features a nonlinear flight dynamics model of an EC135 helicopter (Ref. 19) and a flight controller for stabilization of the aircraft.

Figure 2: Air Vehicle Simulator (AVES) © DLR

2.2. Experimental Setup

The standard AVES configuration with the dome projection system serves as the baseline for comparing a conventional full-flight simulator with an MR setup using the Varjo XR-3 headset (Ref. 20) (see Figure 5). The Varjo headset features a monitor-based display with a FoV of $115°$ and $95°$ (horizontally $\times$ vertically). In this study the Varjo XR-3 Focal Edition is used, which provides a focal distance ranging from 30 cm to 80 cm (Table 2). Additionally, the Varjo has a focus area within the rendered image, where the displayed image is rendered in higher resolution. This focus area, located in the lower center of the screen, has a FoV of $27° \times 27°$, providing a clearer and more readable cockpit instrument display. The headset is connected to the computer used for rendering via a cable and VIVE Base Stations (Ref. 21) were used for head tracking. The two head-tracking stations were installed in the cockpit for the MR setup. One is located in the center above the cockpit panel and another at the top left next to the cockpit door. In addition, a computer used for rendering the MR image is integrated into the back of the pilot seats. The headset is connected to the AVES infrastructure and has access to the internet over a dedicated firewall. Furthermore, an additional lighting was installed to illuminate the cockpit interior, thereby enhancing the quality of the filmed image displayed within the MR headset. In the AVES, chroma key was employed to generate an immersive MR experience using the Varjo XR-3 headset. Chroma key is a visual effects technique that allows for the replacement of a background with a different image or video by matching the color hues of the subject to be replaced (Ref. 22). In this case, the dome has been illuminated green by the projection system (Figure 1). That implicates that wherever the the color green is depicted, the Virtual Environment (VE) is shown in the MR headset. Everywhere else the real world captured by the headset's camera is shown (Figure 5a). Additionally, with the use of masking, the cockpit area was defined as mask/cutout, as flight instruments and panels also use green in their layout. With the mask the cockpit area is free of chroma key artifacts during the simulation.

Table 1: Overview of Visual Systems

Technical Specifications	Varjo XR-3	Dome Projection
Visual System	monitor-based	projection
Field of View (h $\times$ v)	$115° \times 90°$	$240° \times 90°$
Refresh Rate	$90\,\mathrm{Hz}$	-
Weight	$1320\,\mathrm{g}$	-

The conventional visual system features nine FL40MKII projectors with spans a FoV of $240° \times 90°$. The resolution is 2.560×1.600 The MR configuration is rendered with Unity. The conventional setup, the rendered image by the projections system, is based on OpenGL. However, the 3D models for the visual cues of the flight maneuvers are the same. This setup enables a comparison between the dome projection system and MR headset as visual system, allowing for a thorough evaluation of their respective capabilities and limitations in a simulated flight environment.

3. PILOTED SIMULATION CAMPAIGN

This section presents an overview of the simulation campaign setup, including the flight task definition, the assessment methods utilized, and the experiences of the three participating pilots.

3.1. Task Definition

Initially, a familiarization flight was performed to familiarize the pilots with the simulation environment and allow them to become accustomed to wearing the MR headset. Thereafter, three different flight tasks (see Table 2) were flown. Each flight task was first flown with the MR configuration and afterwards with conventional helicopter simulation setup with dome projection.

Within the first flight task, the pilots are asked to plan the flight with the GARMIN GNS430. The GARMIN is located in the front area of the center console (Figure 3 (①)). The aim is to assess how well the pilot can interact with a standardized system and also evaluate the readability of the display. After completing the flight plan, they were asked to depart from the airport

Table 2: 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 Mission-Task-Elements (MTE)
MR	Mixed Reality (monitor-based headset)	Slalom Mission-Task-Elements (MTE)
CON	Conventional simulator with projection dome	Malfunction Procedure
CON	Conventional simulator with projection dome	Hover Mission-Task-Elements (MTE)
CON	Conventional simulator with projection dome	Slalom Mission-Task-Elements (MTE)

and proceed according to their flight plan. During the flight the generator disconnected (GEN DISCON) warning indication (Figure 3 (②)) appeared and they had to act accordingly to the procedure.

If the GEN DISCON appears on the Caution and Advisory Display (CAD) it is because of an electrical failure. The following procedure is given in the Flight Manual:

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

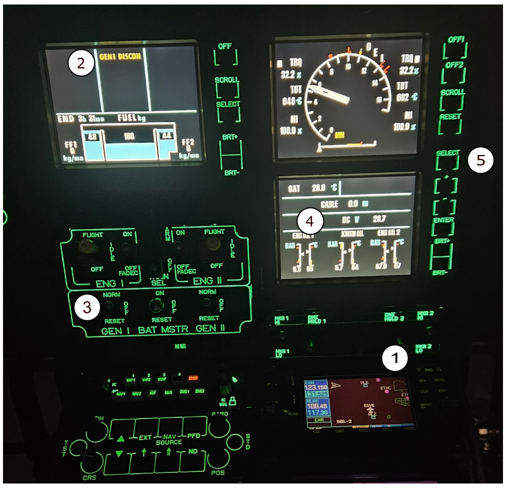

Figure 3: Cockpit Instruments

Therefore, the pilot's interaction consists of switching off the switch of the affected generator (see Figure 3 (③)) and additionally pressing the buttons of the Vehicle and Engine Multifunction Display (VEMD) on the right-hand side to change the displayed GEN field options DC [V], GEN [AMPS] and BAT [AMPS] in order to be able to monitor the parameters (see Figure 3 (④ and ⑤)).

In this scenario, the battery discharged. Once the pilots assessed the situation, they were asked to decide how they wanted to proceed. For example, 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 the cruise flight to the destination, depending on the pilot's decision.

The next two tasks are based on the ADS-33 (Ref. 23), where the pilots flew the Mission-Task-Elements (MTE) Hover and Slalom. The hover maneuver was chosen to assess the effects of the visual system with its limited FoV on the pilot's peripheral vision. The hover maneuver starts with an acceleration to 6 to 10 knots with a relative heading of $45°$ to the south to the hover point. As the pilot is in a stable hover, the pilots were asked to hold the over position for 30 seconds by maintaining heading and altitude less than 20 ft (see also Table 3). Furthermore, the slalom maneuver was selected to investigate if the rapid movement causes simulation sickness. Within the slalom maneuver, the pilot starts at 60 knots. Then the pilot has to maneuver through the slalom parcour with smooth turn at a 500 ft interval (see also Table 4). As support, visual cues for each maneuver have been integrated in a MTE parkour scenario (Figure 4). Each MTE task have been flown three times.

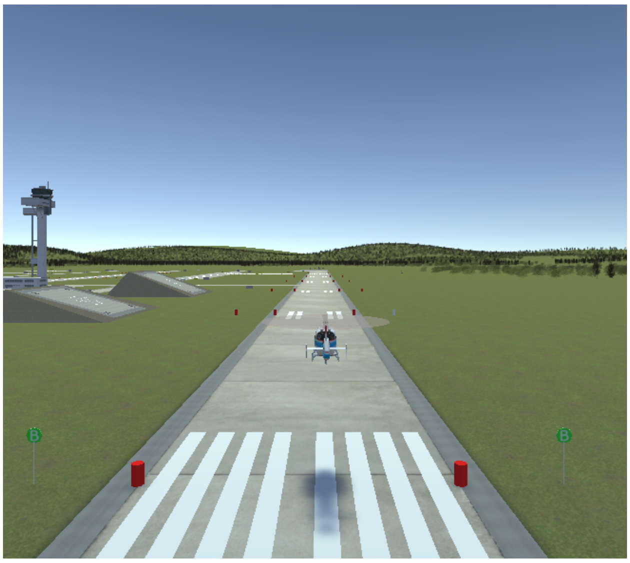

Figure 4: Slalom Parkour within Unity

(a) AVES Conventional Setup **(b)** AVES Mixed Reality Setup

Figure 5: Comparison of two Visual Systems

Table 3: MTE Hover Performance Parameters based on (Ref. 23).

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 ±X ft of a point on the ground:	3 ft	6 ft
Maintain altitude within ±X ft:	2 ft	4 ft
Maintain heading within ±X deg:	5 deg	10 deg

Table 4: MTE Slalom Performance Parameters based on (Ref. 23).

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

3.2. Assessment Methods

For the purpose of the evaluation, various assessment techniques have been employed. On one hand, standardized subjective assessment methods have been utilized, including the Presence Questionnaire (PQ) (Ref. 24) and the Simulation Sickness Questionnaire (SSQ) (Ref. 25). On the other hand, for an objective assessment of flight performance, flight data has been recorded, allowing for a comparison of each pilot's performance across different configurations.

In the existing literature, it appears that the sense of presence emerges as a crucial factor in the evaluation of virtual environments (Ref. 26). Nevertheless, there exists a disparity in definitions surrounding the concept of presence, accompanied by various questionnaires designed to assess it. The most commonly employed questionnaire for this purpose is the Presence Questionnaire devised by Wittmer and Singer. According to their definition, presence is described as "the subjective experience of being in one place or environment, even when one is physically situated in another" (Ref. 24). The PQ consists of 32 questions, each question having been assigned to at least one of the following factors that influence presence: Control Factors (CF), Sensory Factors (SF), Distraction Factors (DF), and Realism Factors (RF). The PQ utilizes a seven-point scale. Another widely utilized questionnaire is the questionnaire developed by Slater Usaoh and Steed (SUS) (Ref. 27). This rating scale consists of six questions designed to assess presence in relation to real-life experiences. The PQ by Wittmer and Singer, with its extensive catalogue of questions, takes into account a large number of factors that relate to both virtual and real environments and was therefore selected from these two questionnaires for this study.

Simulation sickness, a phenomenon characterized by a range of physical symptoms, can occur when individuals utilize VR and MR headsets. The human sense of orientation relies on the integration of information from various sensory organs, including visual, vestibular, and proprioceptive systems. These sensory inputs are processed to determine one's current position and movement in space. However, a mismatch between visual and vestibular information can lead to a sensory conflict, which may persist for an extended period, ultimately culminating in the development of simulation sickness. In order to assess simulation sickness, various types of questionnaires have been employed. The most widely utilized questionnaire is the Simulation Sickness Questionnaire ((Ref. 28) and (Ref. 29)). This standardized assessment method enables a rapid evaluation of 16 physical symptoms that may occur during a simulation study. Participants can rate their symptoms on a scale of none, light, moderate and severe.

In addition to subjective evaluation methods, the pilots' flight performance during respective maneuvers is also assessed. As the tasks for hovering and slalom flying are clearly defined, it is possible to compare flown maneuvers with different visual systems. This makes it possible to evaluate the impact of different visual systems on pilot performance.

3.3. Pilot Experience

Overall, three helicopter pilots participated in the campaign (Table 5). All were familiar with see-through and monitor-based headsets.

Table 5: Pilot Experience

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

4. RESULTS

In this section, the results of the piloted study are presented. Initially, the subjective ratings provided by the pilots for the PQ and the SSQ are shown. Subsequently, an analysis is conducted on the flight performance achieved by each pilot during hovering and slalom maneuvers with the two configurations.

Results of the Questionnaire

The analysis begins with an overview of the PQ results for the Malfunction Procedure. This will be followed by a review of all PQ outcomes. Subsequently, the SSQ Scores for the Malfunction Procedure is analyzed, after which all SSQ results are presented and discussed.

The visualization of the PQ Score follows the grading scale presented in (Ref. 30), which enables the evaluation of system usability in addition to the PQ score. A grading score below $50\,\%$ of the PQ score is considered unacceptable, while scores above $50\,\%$ are classified in increments of $10\,\%$ as unsatisfactory, marginal, satisfactory, very good, and excellent. In this study, the audio-related questions of the PQ were excluded, as no audio was used during the piloted simulation campaign.

Figure 6 presents the total score for the PQ evaluation performed by all three pilots for the malfunction procedure. All pilots rated the MR configuration at least one point lower than the conventional configuration, indicating a perceived lower presence with the MR configuration. However, they were able to perform the malfunction

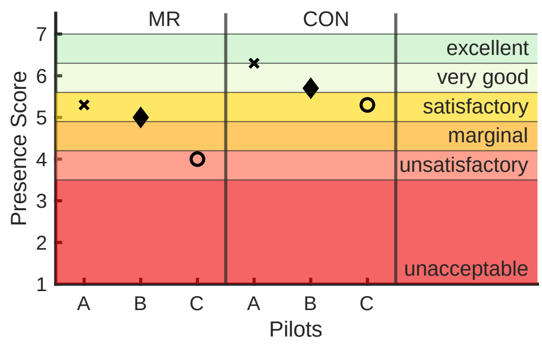

Figure 6: PQ Overall Score for Malfunction Procedure

procedure. With both configurations the pilots were able to operate the toggle switch for switching off generator 1 without encountering any issue. It was also possible to toggle and read the displays for BAT, GEN and DC with both configurations. Visual navigation to the destination airport also was possible. The pilots were able to determine their position using a channel and a highway as reference points. However, the pilots stated that they preferred the conventional configuration with the dome projection for several reasons. Initially, they reported that it was easier to use the GARMIN for route planning to the destination airport with the MR headset while still on the ground than in flight, as the displays were barely readable. Without the MR headset, however, the pilots had no problems reading the display of the GARMIN. They used it as part of their normal navigation procedures. Pilot A and Pilot C mentioned that without wearing the MR headset they felt safer to use the system and still maintain a stable cruise flight. Pilot C added that with the peripheral vision of the conventional configuration, it is easier to assess and be aware of the flight situation than with the limited FoV of the headset. Another reason why they favored the conventional configuration is that the image in the MR headset occasionally flickered or lagged. This can be attributed to the fact that head tracking setup was not optimal in the compact cockpit. In addition, there were some unintentional reflections due to the cockpit windows and additional lighting, which led to unwanted artifacts in the MR headset. Nonetheless, Pilot A and B rated the presence with the MR with a 5 and 5.3 which is according to the grading already satisfactory. Considering that there is still room for improvement through software and hardware optimization, this indicates potential for the usability of MR headset for future helicopter pilot training.

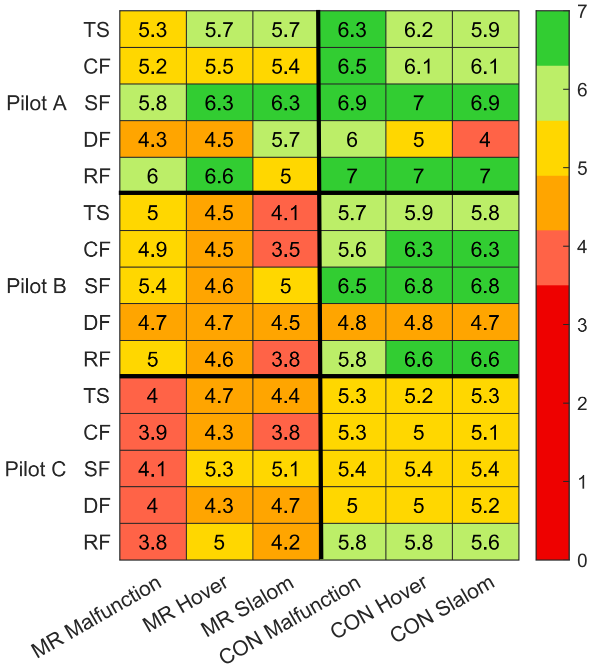

Figure 7: PQ Score Heatmap for all Configurations

These results are also reflected in Figure 7, which provides an overview of all pilots PQ ratings. A clear trend emerges, indicating that not only for the Malfunction Procedure but also in general, presence is rated lower when using the MR headset. The pilots reported that with the MR headset, flight instruments were displayed less sharp and blurred even more when they moved their heads. Additionally, due to the limited FoV, pilots had to focus more to identify values on the flight instruments or visual cues outside. Even though, they had a person sitting on the copilot seat, they were not able to detect their movement. Normally, the pilots see the movement of the crew member in his peripheral vision, however, with the MR headset they needed to move their head more to scan their environment. For example, Pilot B's rating for the slalom maneuver shows a score of 3.8 for both Realism Factor and Control Factor (Figure 7), indicating that the MR headset degraded presence during this maneuver. Especially in a dynamic maneuver, a fast scan of the flight instruments is necessary. Additionally, the pilot emphasized that latency-free rendering is essential to perform the maneuver within limits, as any lagging would require corrective control inputs.

The diagram in (Figure 8) shows the results of the questionnaire on simulation sickness for the malfunction procedure. To facilitate comparative analysis with the corresponding PQ diagram, the Simulation Sickness Score axis was inverted and truncated at 35 to ensure optimal visual alignment. Furthermore, a color coding scheme was incorporated to enable rapid visualization and comparison of the results with those obtained through the PQ assessment. All pilots had no symptoms while flying the conventional helicopter simulator configuration, meanwhile Pilot B had significant symptoms wearing the MR headset. Pilot B experienced symptoms which are classified as "bad" based on (Ref. 31). The SSQ scale ranges from 0 to 235, making this classification of (Ref. 31) highly stringent. Pilot B rated five symptoms with "slight" which already gives a SSQ of 30. The ratings for "slight fatigue" and "slight general discomfort" were attributed to mild neck pain experienced while wearing the MR headset. It is worth noting that the initial calibration process for the SSQ was conducted with military pilots, who may not be representative of the general population, as outlined in (Ref. 32). Figure 9 presents

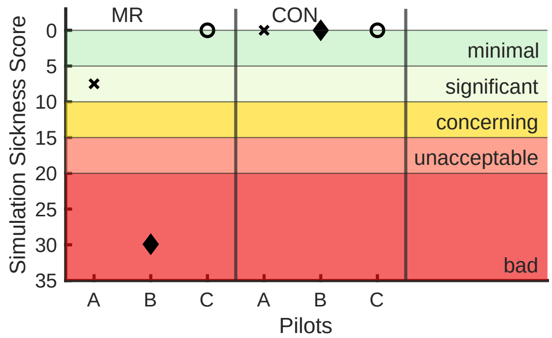

Figure 8: SSQ Overall Score for Malfunction Procedure Task

the SSQ results for diverse maneuvers and configurations. All pilots reported no symptoms of simulation sickness when flying with the conventional setup. With this results it should take into account that all pilots are very familiar with the AVES simulator, having already undergone familiarization with this simulation environment. The further analysis shows that the MR configuration requires further refinement, as all pilots reported symptoms of simulation sickness. In addition, these symptoms may be mitigated by optimizing the hardware and software setup. For example, 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. Despite some limitations in this experimental setup, there is room for improvement in the MR configuration, which could lead to even better results with a well-designed setup.

Results of the Flight Performance

The analysis commences by examining pilot B's hover flight performance as an exemplar. This is followed by a evaluation of the overall performance of all pilots during the hover maneuver. Then, the slalom flight performance is analyzed, with pilot B's results serving as a representative example. Finally, the flight performances of all pilots for the slalom maneuver are presented and discussed.

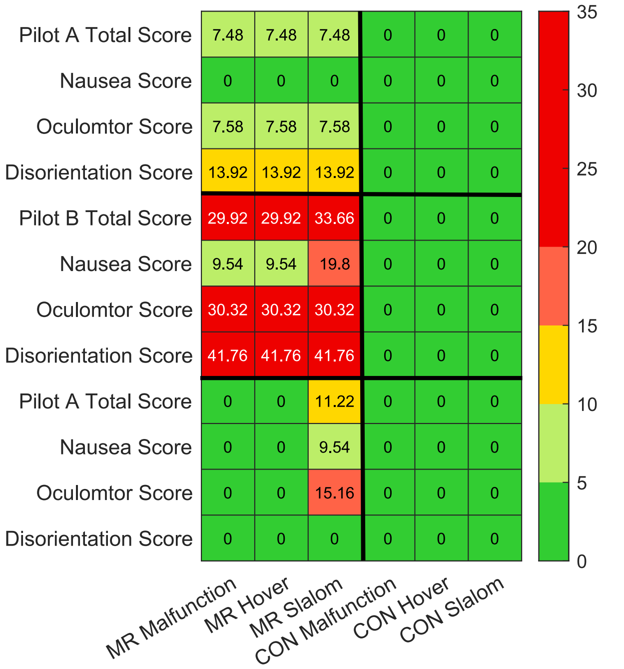

Figure 9: SSQ Score Heatmap for all Configurations

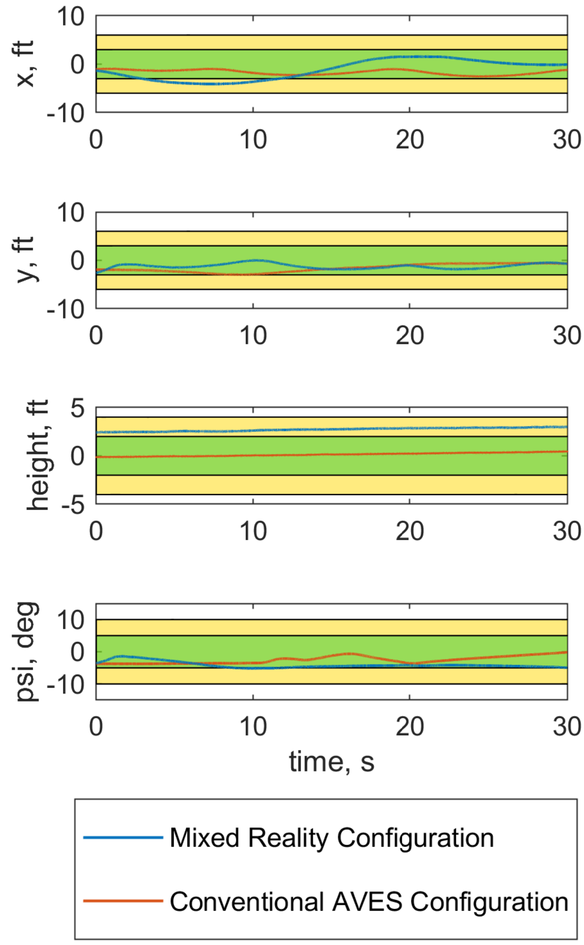

Figure 10: MTE Hover Performance of Pilot B

In Figure 10 an example of the flight performance for hover maneuver of Pilot B is shown. In the two upper diagrams the x- and y position is presented for the duration of the 30 seconds of hover. The green color within the diagram indicates that the pilot's performance is within the desired limits of the hover MTE. The yellow area represents adequate limits. For the MR configuration the pilot stays for all parameter (x- and y position, height and heading (psi)) within the the adequate limits. As for the conventional configuration, the pilot meets desired performance for all parameters. The findings from the questionnaires are consistent with the flight performance data. A correlation seems to exists between the pilots' evaluations of presence with the MR configuration and their subsequent hover performances. When pilots rated their presence with the conventional configuration higher than with the MR configuration, they demonstrated improved flight skills during the hover. As depicted in Table 6, it can be observed that each pilot demonstrated improved flight performance when utilizing the conventional configuration. If all performance parameters meet the desired limits, the notation "desired" is used in the table.

Table 6: Overview of the Flight Performance for Hover Maneuver

Hover	Pilot A	Pilot B	Pilot C
MR	not adequate	adequate	not adequate
CON	adequate	desired	adequate

Similarly, if all performance parameters meet the adequate criteria, the notation "adequate" is used. It should also be noted that an effective tracking and positioning of the camera image within the virtual environment is particularly important for this maneuver. A slight shift in the image can already make a significant difference between "desired", "adequate" or "not adequate" regarding the heading performance limits. In Pilot C's videos and with the pilot's comment, it was evident that his view was slightly offset, which falsified his heading and ultimately led to an assessment of "not adequate" for this maneuver, as this deviation was not taken into account in the evaluation.

Figure 11 shows the flight performance of the slalom maneuver of Pilot B. The upper diagram shows the flight path from above. All pilots, including Pilot B, had difficulty navigating through the final gate with the MR configuration. They reported that flying with the conventional setup was easier because of a wide FoV. This allowed pilots to quickly scan their surroundings or recognize other cues in their peripheral vision. Pilot B maintains ground speed (see middle diagram of 11) between 42 and 70 kts and is with that within the adequate limits of the MTE Parameters. With both configurations Pilot B reached desired performance parameters regarding the altitude. The overview Table 7 reveals that all

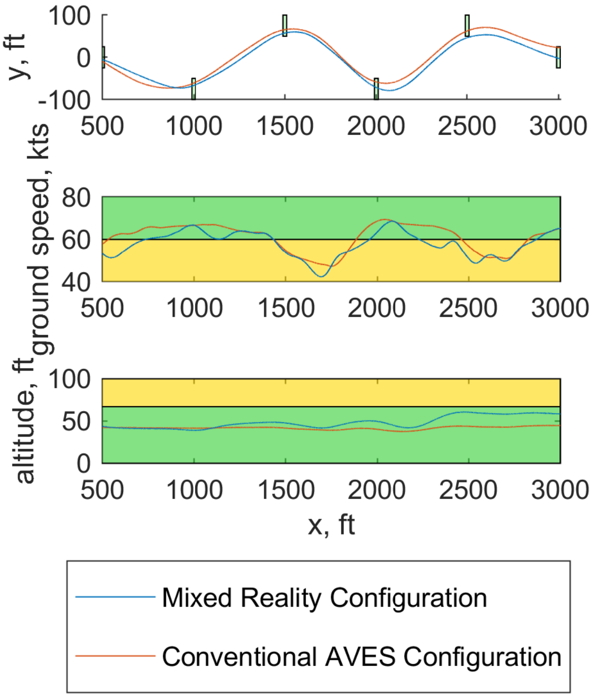

Figure 11: MTE Slalom Performance of Pilot B

pilots achieved adequate performance in the slalom maneuver with the conventional setup, with most capable of meeting altitude and position parameters. However, they faced challenges maintaining speeds above 60 kts throughout the entire 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.

Table 7: Overview of the Flight Performance for Hover Maneuver

Slalom	Pilot A	Pilot B	Pilot C
MR	not adequate	not adequate	not adequate
CON	adequate	adequate	adequate

5. DISCUSSION

This section begins with an examination of the results from the SSQ and PQ, followed by an analysis of the flight performance data. The relationship between Presence and flight performance will then be explored. Finally, the study's limitations and potential avenues for future research will be discussed.

Presence and Simulation Sickness

The pilots' ratings for presence within the simulation environment generally decreased when using the MR configuration compared to the conventional configuration. Although the MR configuration was rated worse than the conventional configuration in terms of presence, two pilots rated it as "satisfactory". This result indicates that MR technology could be a promising tool for training helicopter pilots. However, this result also shows that the experimental MR setup needs to be improved in order to achieve a level of presence comparable to a full flight simulator.

The results of the SSQ indicate that the MR configuration in this experimental setup is more prone to triggering simulator sickness than the conventional setup. To address this issue, removing factors such as flickering and latency could help to improve the overall experience.

Flight Performance

The overall flight performance data shows that the conventional configuration is superior to the MR configuration in terms of achieving the desired flight performance. However, the pilots still achieved adequate flight performance with the MR configuration. These results imply that the visual disturbances associated with the MR configuration, such as flickering, blurred vision and delayed images, may have contributed to poorer performance for the pilots. To mitigate these issues, integrating cockpit instruments and navigation displays directly into the MR headset could improve the overall experience and potentially enhance flight performance.

Correlation between Presence and Flight Performance

There is a significant correlation between the pilots' assessment of their presence in the simulation environment and their subsequent hover and slalom performance. When pilots rated their presence higher, they demonstrated improved flight skills during the maneuvers. This finding suggests that the sense of presence can positively impact overall performance in immersive environments.

Limitations and Future Directions

Despite some limitations in this experimental setup, such as sub-optimal lighting conditions and limited tracking station placement, the results suggest that optimizing the hardware and software setup could lead to even better results with the MR configuration. Future studies should aim to address these limitations and explore further ways to improve the MR experience.

6. CONCLUSIONS

This study presents an evaluation of MR helicopter simulation compared to simulator with a conventional dome projection system. The results are based on subjective evaluations using Wittmer and Singer's PQ scale and the SSQ scale developed by Kennedy et al., as well as flight performance assessments. The findings indicate that:

1. In general, pilots preferred the conventional simulator configuration with a dome projection over the MR configuration using a MR headset.
2. Despite this preference, MR offers potential for helicopter training.
3. Presence is strongly correlated with flight performance, suggesting a direct impact.
4. Navigation displays may not be compatible with a MR headset; streaming directly to the headset could be a more effective approach.

In conclusion, the MR headset in this experimental setup does not yet approach the level of presence offered by dome projection systems in helicopter simulator. However, it is a promising, cost-effective and immersive training platform for helicopter pilot training.

Author contact:
Tanja Martini tanja.martini@dlr.de

7. REFERENCES

1. ICAO. ICAO addresses shortage of skilled aviation professionals. https://www.icao.int/newsroom/pages/icao-addresses-shortage-of-skilled-aviation-professionals.aspx. Accessed 7 March 24.
2. D. Solar and C. Audard. Rotorcraft safety roadmap. Technical report, EASA, December 2018.
3. EASA. Certification Specifications for Helicopter Flight Simulation Training Devices. EASA Annex, EASA, June 2012.
4. EASA. EASA approves the first Virtual Reality (VR) based Flight Simulation Training Device. https://www.easa.europa.eu/en/newsroom-and-events/press-releases/easa-approves-first-virtual-reality-vr-based-flight-simulation. Accessed 2024-07-22.
5. Loft Dynamics. About the Airbus H125 FSTD. https://www.loftdynamics.com/flight-training-solutions/products/h125-simulator/. Accessed 2024-07-22.
6. EASA. 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.
7. 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.
8. Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. pages 282–292, Boston, MA, December 1995.
9. 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*, Boston, MA, August 2013. American Institute of Aeronautics and Astronautics.
10. 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.
11. Johannes Maria Ernst and Lars Ebrecht. Virtual cockpit instruments on head-worn displays for helicopter offshore operations in confined areas. In *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, pages 1–10, San Antonio, TX, USA, October 2020. IEEE.
12. Matthias Oberhauser and Daniel Dreyer. A virtual reality flight simulator for human factors engineering. *Cognition, Technology & Work*, 19(2-3):263–277, September 2017.
13. Stefan Auer, Jens Gerken, Harald Reiterer, and Hans-Christian Jetter. Comparison Between Virtual Reality and Physical Flight Simulators for Cockpit Familiarization. In *Mensch und Computer 2021*, pages 378–392, Ingolstadt Germany, September 2021. ACM.
14. 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. pages 34–45, Orlando, FL, June 1997.
15. 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.

16. 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.

17. 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.

18. 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.

19. Alexander Štrbac, Tanja Martini, Daniel H. Greiwe, Frauke Hoffmann, and Michael Jones. Analysis of rotorcraft wind turbine wake encounters using piloted simulation. *CEAS Aeronautical Journal*, 12(2):273–290, April 2021.

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

21. VIVE. SteamVR Basisstation 2.0. https://www.vive.com/de/accessory/base-station2/. Accessed 2024-03-20.

22. S. Shimoda, M. Hayashi, and Y. Kanatsugu. New chroma-key imagining technique with Hi-Vision background. *IEEE Transactions on Broadcasting*, 35(4):357–361, December 1989.

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

24. Bob G. Witmer and Michael J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, June 1998.

25. 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.

26. 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.

27. 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.

28. David M. Johnson. Introduction to and Review of Simulator Sickness Research. Technical report, American Psychological Association, 2005.

29. 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.

30. 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.

31. 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.

32. 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.