BEYOND TRADITIONAL RADIO: EXPLORING SPATIAL-AUDIO SYSTEMS FOR ENHANCED COMMUNICATION IN MULTITASKING FLIGHT ENVIRONMENTS

S.-M. Stasch[a], F. Ernst[b], J. von Mankowski[a], E. Sachau[b], W. Mack[a], D. Sachau[b]

[a]Universität der Bundeswehr München, Institut für Psychologie, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Deutschland
[b]Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Institut für Mechatronik, Holstenhofweg 85, 22043 Hamburg, Deutschland

Abstract

Pilots often operate within large formations that include both aircraft and drones, making the spatial arrangement of these units critical for successful formation flying and the prevention of safety hazards. Effective communication among formation members, typically transmitted through radio messages, heavily relies on understanding the spatial relationships between individual aircraft. This is particularly crucial in low-visibility situations, where visual contact is limited. Current systems lack spatial cueing capabilities, which can hinder the clarity and effectiveness of communication. Recent advancements in spatial audio technology, specifically 3D audio, offer a potential solution by incorporating spatial cues into radio transmissions. By providing directional information about the transmitting aircraft, spatial audio could significantly improve situational awareness within the formation. Building on prior research into audiovisual integration in cockpit environments, this study investigates the impact of congruent and incongruent spatial cueing, as well as the absence of spatial cueing, on radio communication in low-fidelity flight scenarios. The study aims to determine whether 3D audio systems can enhance communication performance in multitasking flight environments. The findings have implications for the design of cockpit interfaces and radio communication systems.

Keywords

3D audio support, multitasking, radio messages, task prioritization

1. INTRODUCTION

1.1. Spatial Awareness in the Cockpit

Flight operations are inherently complex, requiring pilots to manage multiple tasks simultaneously under often high-pressure conditions. These tasks include monitoring aircraft systems, interpreting instrument data, maintaining communication with air traffic control (ATC) and other aircraft, and responding to dynamic changes in the flight environment. The cognitive demands of multitasking are particularly pronounced in complex flight scenarios, such as formation flying, where precise coordination between aircraft is essential for mission success and safety. Additionally, the integration of unmanned aerial vehicles (UAVs) into modern military flight operations further increases the complexity of that multitasking environment, requiring pilots to manage both manned and unmanned assets within a unified operational framework [1].

Historical and contemporary aviation accidents underscore the critical importance of spatial situational awareness and multitasking capabilities in the cockpit. In the military domain, the June 19, 1962 [2], crash of four Lockheed F-104F "Starfighter" jets from the German Air Force, resulting in the loss of all the pilots, serves as a tragic reminder of the potential consequences of compromised spatial situational awareness. More recently, the April 23, 2024, collision of two Royal Malaysian Navy helicopters [3] during a military parade rehearsal resulted in the loss of 10 lives. In the civil domain, the November 12, 1996, mid-air collision between Saudi Arabian Airlines Flight 763 (Boeing 747) and Kazakhstan Airlines Flight 1907 (Ilyushin Il-76) near Charkhi Dadri, India, remains the deadliest in history, claiming 349 lives [4]. These accidents highlight the vital role of spatial situational awareness for safety in aviation, particularly in time-critical, high-stakes environments.

1.2. Theoretical Framework of Multitasking in Aviation

To better understand the cognitive mechanisms underlying errors in aviation, it is essential to examine multitasking behavior, a key determinant of pilot performance. According to [5], multitasking exists along a continuum ranging from concurrent multitasking, characterized by rapid task switching, to sequential multitasking, where tasks are executed serially with sustained focus before transitioning to another task. Wickens' Multiple Resource Theory from 1980 [6] provides an explanatory framework for multitasking errors in aviation. The theory posits that cognitive resources are limited and that errors arise when multiple tasks compete for the same processing resources. Specifically, information processing is structured along three dimensions (see Fig. 1):

1. *Codes*: This dimension represents the type of information being processed. Verbal codes refer to verbally transmitted information, such as ATC instructions, whereas spatial codes pertain to spatially presented information, such as those displays on flight displays or radar screens.
2. *Modalities*: This dimension distinguishes between visual and auditory processing. The visual modality encompasses information acquired through sight, while the auditory modality pertains to information processed through hearing.
3. *Stages*: This dimension delineates the phases of information processing: encoding, central processing, and response execution. Encoding involves the perception and mental representation of incoming information. Central processing entails cognitive operations such as interpretation and decision-making. Finally, response execution involves the selection and implementation of an appropriate action, which may be manual (e.g., pressing a button) or vocal (e.g., issuing a verbal command).

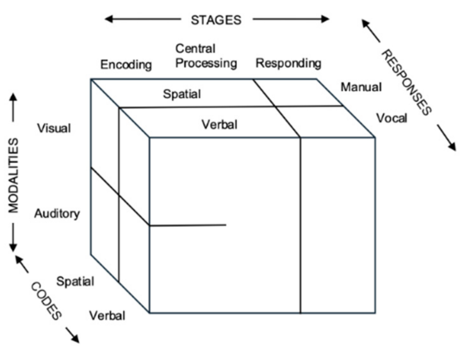

Figure 1: Conceptual framework of the Multiple Resource Theory [6].

According to Wickens [6], errors due to multitasking can emerge at all three stages of information processing, especially when two or more tasks compete for the same limited resource. This might be the case when a pilot attempts to process two verbal information simultaneously. An example of that scenario would be the communication with ATC to receive new approach instructions (verbal code/ auditory modality) and monitoring an auditory engine performance alert (verbal code/ auditory modality). Both tasks create a high degree in overlap, which will very likely result in task interference and degraded task performance. This can manifest itself in a failure to fully understand the approach instruction due to distraction by the engine performance alert or a misjudgement of the severity of the engine alert. Tasks that engage different cognitive resources, however, are less likely to interfere with one another. Instead, an auditory engine alert paired with a visual display can reduce cognitive load on the auditory channel, allowing the pilot to more effectively process ATC instructions.

1.3. Task Prioritization and Cockpit Task Management Theoretical Framework of Multitasking in Aviation

Given these limitations in cognitive resources, task prioritization emerges as one strategy for managing concurrent tasks effectively in the cockpit. Task prioritization is often discussed under the broader concept of cockpit task management (CTM), which involves initiating, monitoring, prioritizing, and terminating tasks based on their relative importance [7]. Research conducted by Chou, Madhavan, and Funk [8] underscores the significance of task prioritization in the cockpit, noting that errors in CTM contributed to nearly half (49%) of incident reports and 23% of accident reports. Among these CTM errors in accident reports, task prioritization errors were the second most frequent (28%). These findings underscore the necessity of deliberate cognitive resource allocation to mitigate task interference and maintain high operational performance.

1.4. Enhancing Cockpit Communication Through Spatial Audio

Modern flight operations rely predominantly on the visual modality, with pilots depending on instrument displays and external visual references for flight control, navigation, and situational awareness [9]. Peng et al. [10] estimate that approximately 80% of the information processed by pilots during flight is visual. In contrast, the auditory modality remains rather underutilized, primarily serving as a medium for radio communication and auditory alerts. However, traditional auditory communication lacks spatial cues, necessitating additional cognitive effort to interpret the relative positions of aircraft and environmental threats.

In this context, three-dimensional (3D) audio technology presents a promising advancement. 3D audio systems simulate sound sources in three-dimensional space, enabling pilots to perceive the direction and distance of auditory cues. Previous research has demonstrated that spatial audio can enhance response times [11, 12] and improve target localization [13, 14, 15]. A comprehensive overview of the benefits of 3D audio in aviation can be found in Niermann [9].

Integrating spatial audio into cockpit communication systems may alleviate visual workload and herby facilitate more efficient multitasking. Findings from Brock et al. [16] further illustrate that. The authors showed that that by using spatialized auditory displays, the response times for tactical decision tasks were significantly reduced without sacrificing accuracy in a dual-task simulated multiscreen watchstation environment. This effect was attributed to the ability of spatialized auditory displays to effectively direct attention without requiring additional visual checks. This mechanism allows operators to process auditory information about multiple tasks simultaneously, leading to improved response times and reduced cognitive load when managing tasks across the multi-screen environment.

Integrating 3D audio into a cockpit communication system is particularly relevant in formation flying, where pilots must continuously monitor the relative positions of multiple aircraft. However, given the inherent susceptibility of technical systems to errors—whether due to malfunctions or deliberate adversarial interference in military contexts—it is also crucial to investigate the potential consequences of erroneous, hereafter also referred to as spatially incongruent, 3D radio messages.

The underlying rationale is that incongruent information increases the need for conflict monitoring to resolve discrepancies, thereby elevating cognitive load and attentional demands [17]. As more cognitive resources are devoted to managing this conflict, fewer remain available for other concurrent tasks, potentially resulting in task interference and performance degradation.

Therefore, a systematic evaluation of the risks associated with incongruent information is essential to assess the operational viability of 3D audio technology, but also to understand its broader impact on task performance.

1.5. Research Question

While the benefits of 3D audio in cockpit environments have been well-established in prior research, the effects of spatial audio in high-demand multitasking scenarios, such as formation flight with varying task prioritization, remain underexplored.

Task prioritization has been shown to exert a substantial influence on task performance in earlier studies [18, 19]. Due to the limited scope of this paper, this aspect will not be discussed in detail here. Readers interested in a more comprehensive examination of the prioritization effects are referred to [20], which specifically addresses the corresponding findings of that study.

Therefore, this study seeks to address the following research question: What are the effects of correctly presented (congruent) and incorrectly presented (incongruent) 3D radio messages on multitasking performance, compared to non-spatial radio messages, under varying task prioritization conditions in a simulated formation flight?

1.5.1. Hypotheses

The following hypotheses were derived from our research question:

1. *H1*: Congruent 3D audio messages will enhance performance across all flight tasks and improve subjective support compared to both incongruent 3D audio messages and non-spatial radio messages (no3D).

This result is expected because with the correct presentation of 3D audio messages, a redistribution of cognitive resources from the visually overloaded channel to the auditory channel can take place. That would not be the case with incongruent or no 3D radio messages, since the relative position in formation flight would need to be visually verified.

2. *H2*: Incongruent 3D audio messages will impair performance across all flight tasks and decrease subjective support compared to congruent 3D audio messages and non-spatial radio messages.

This hypothesis is attributed to the necessity for visual reverification of the relative position in formation flight, which arises from the incorrect spatial representation of radio messages. This additional load on the visual channel may exacerbate task interference and degrade overall task performance.

3. *H3*: Non-spatial radio messages will result in better performance across all flight tasks and higher subjective support ratings than incongruent 3D audio messages but lower performance and support ratings than congruent 3D audio messages.

This result is expected because, although pilots must still visually determine relative positions in formation flight, non-spatial radio messages do not introduce the additional cognitive burden of detecting and correcting for incorrect spatial positioning.

2. METHODS

2.1. Current study

To address this research question, the current study examines multitasking behavior in a standardized test environment using a modified version of the Multi-Attribute Task Battery (openMATB) [21], originally developed by Comstock & Arnegard [22]. The MATB traditionally comprises four flight-relevant tasks a tracking task, a system monitoring task, a communication task, and a resource management task. In the modified setup, participants engage in a simulated formation flight, during which other formation members communicate with them via radio. These radio transmissions are spatially modulated using a 3D audio system [23, 24] to reflect the relative direction of the transmitting aircraft relative to the participant's aircraft within the formation. Specifically, radio messages are either directed to the participant's own aircraft or to other aircraft within the formation. These transmissions contain radio frequencies of other formation members, which participants must compare to tuned frequencies displayed in the user interface (UI). Additionally, some radio calls include "evade" commands, requiring participants to immediately steer away from the transmitting aircraft by moving the tracking task cursor in the opposite direction. Details on the single MATB tasks are discussed in the following section.

2.2. The OpenMATB

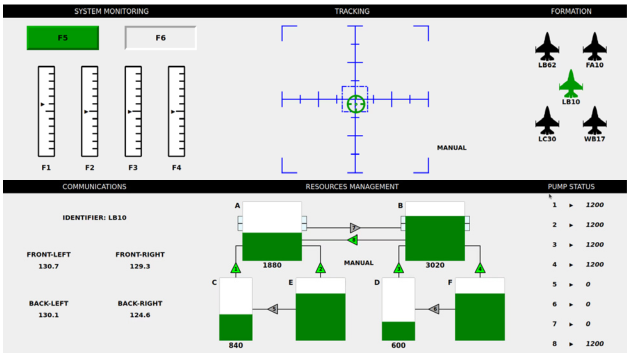

Figure 2: Modified OpenMATB Software [21] depicting the new FORMATION (top right) and COMMUNICATION (bottom left) plugins.

The original OpenMATB [21] provides six views, displaying various information and tasks that utilize similar cognitive resources (auditory and visual) as pilots do when operating an aircraft in a formation flight. This setup offers an ecologically valid framework for studying multitasking performance by allowing flexible adjustments of task complexity. OpenMATB, an open-source software, enables straightforward modifications to the original interface for simulating our formation flight use case (see Fig. 2).. As in the original OpenMATB, the UI is displayed on a screen, and participants provide input using a keyboard and a joystick for the tracking task (see Fig. 3).

In the SYSTEM MONITORING task, participants must reset the scales and lights to a predefined state when deviations exceed a preset threshold. This is accomplished by pressing the function keys (F1–F6) on the keyboard. In the RESOURCE MANAGEMENT task, the objective is to maintain the levels of the main tanks (A, B) at a specified target by toggling pumps on and off using the numeric keys (1–8). The PUMP STATUS indicates the pumping capacity of each pump. In the TRACKING task, participants must keep the green cursor centered using a joystick.

Compared to OpenMATB, the former SCHEDULER view is replaced by the FORMATION view, which displays the call signs and positions of other aircraft in the formation. Additionally, a new auditory task is introduced. Participants must evade an aircraft upon hearing a command structured as 'Lima Bravo One Zero (LB10), evade immediately.' This is followed by the call sign of the aircraft to avoid. To evade, the participant pulls the joystick in the opposite direction of the specified aircraft until the circle reaches the corner of the display, then quickly returns it to the center. During this task, the TRACKING task is not evaluated. Not all evasion commands apply to the participant; for instance, if the command begins with a call sign other than LB10, it should be ignored. The call signs change with each trial but remain consistent within a single trial.

Additionally, the communication plugin is redesigned. Previously, participants had to set the frequencies for each radio. Now, radio frequencies are displayed spatially to align with the aircraft formation, linking each radio to a specific aircraft and corresponding joystick button. The joystick buttons are arranged in the same spatial formation as the radios, so the top-left button (5) is linked to the top-left radio (FRONT-LEFT) and the top-left aircraft (LB62). In this task, participants must verify whether a stated frequency is correct. They hear an audio command such as "Lima Bravo One Zero, check frequency" followed by a call sign and frequency. Participants must first determine whether the command is directed at them, locate the call sign in the formation display, and then find the corresponding radio in the communication section. If the called-out frequency matches the displayed radio frequency, participants confirm by pressing the corresponding joystick button once (confirm-trial). If it does not match, they reject it with a double-click (reject-trial). Similar to the evasion task, radio calls that start with a call sign other than the own (here: LB10) should be ignored.

Both audio tasks must be completed within 15 seconds of the audio onset. Tasks were presented according to the hypotheses (see Section 1.5) in three conditions: no spatial modulation (no3D), congruent spatial modulation (congruence), or incongruent spatial modulation (incongruence). In the congruent condition, sounds are spatially modulated with the 3D audio system to originate from the position corresponding to the second call sign. In the incongruent condition, the audio source is also spatially modulated but originates from a different direction than the second call sign. Without spatial modulation, the unmodulated mono signal is played on both headphone channels.

2.3. Experimental Set-Up

Participants were seated in a soundproof cabin in front of a 2560*1140-pixel monitor, which displays the modified openMATB multitasking environment (see

Fig. 3). The 3D audio radio messages are presented to the participant through Beyerdynamic DT 770 M headphones. Participants use a Logitech Extreme 3D Joystick to operate the tracking task and make radio inputs for the communication task. A keyboard is used to make input for the system monitoring and the resource management task.

Figure 3: The experimental set-up (top image). The participant performs the tracking task and inputs radio responses for the communication task using a joystick with the right hand (bottom image), while the left hand is used to operate the system monitoring and resource management tasks.

2.4. 3D Audio

The perception of spatial sound is based on interaural time differences (ITD), interaural level differences (ILD), and spectral cues that vary with the direction of arrival [25]. These auditory localization cues are captured by Head-Related Transfer Functions (HRTFs), which describe how sound waves interact with the listener's head, ears, and torso before reaching the eardrums. By filtering a mono audio signal through HRTFs corresponding to a specific direction, spatialized sound is generated, enabling listeners to perceive sound sources in 3D space when using headphones.

In this study, 3D audio was generated using the system developed by Ernst & Sachau [23, 24], which is based on the ARI HRTF Database [26]. This database provides HRTFs with a 2.5-degree resolution in azimuth, allowing fine-grained spatial audio representation. A single, non-individualized HRTF set was used for all participants. All sound sources were presented exclusively on the horizontal plane (0° elevation). During an initial localization training, the azimuth of the sound sources was randomized to familiarize participants with the system. In the main experiment, however, auditory cues originated from only four distinct directions corresponding to the positions of the formation members at 45°, 135°, 225°, and 315°.

To enhance spatial perception, a camera-based head tracker was integrated to enable dynamic rendering of the 3D audio. The Neuralnet tracker within the open-source software OpenTrack [27] was utilized to capture head position data. The tracking latency was below 35 ms. This data was sent directly to a dSPACE MicroLabBox for real-time convolution of the audio signal with the corresponding HRTFs. This setup allows participants to improve sound localization through natural head movements, a method shown to reduce localization errors [23, 28].

2.5. 3D audio localization training

Before the main experiment, participants undergo a training phase to familiarize themselves with the spatial characteristics of the 3D audio system. In this phase, they listen to voice announcements originating from randomized spatial locations and indicate the perceived direction on a screen. Immediate visual feedback is provided displaying the correct position of the voice announcement. The deviation between the estimated and actual direction is calculated as a moving average. Training consists of a minimum of 10 trials and continues until participants achieve an acceptable level of localization accuracy. This ensures sufficient familiarity with spatial sound cues before proceeding to the main study.

2.6. Experimental Procedure

The experiment begins with 3D audio localization training, followed by task-specific training where participants practice each modified MATB task individually (see Fig. 4). During this phase, only the relevant task is displayed, with all other task modules hidden.

Once participants have completed individual task training, they proceed to combined task training, where all MATB tasks are performed simultaneously under the congruent spatial audio condition. Thus, participants completed all subtasks concurrently after having previously practiced each one independently. In this phase, the communication task included only correctly presented radio calls to facilitate participants' learning of the spatial auditory cues. Training continues for up to six rounds but can be terminated early if participants reach an 80% accuracy rate in the auditory task.

The main experiment consists of six blocks, which were presented in a randomized order. Each block is defined by two experimental factors: audio modulation (no spatial modulation, congruent spatial modulation, and incongruent spatial modulation) and task prioritization (priority vs. no priority).

- *Priority condition:* Participants are instructed to prioritize auditory tasks. This condition is visually reinforced by displaying a cockpit scene in bad weather with an instructional message emphasizing the need to avoid collisions.
- *No priority condition:* Participants are instructed to distribute their attention equally across all tasks. A cockpit scene in good weather is displayed, accompanied by a message emphasizing balanced task performance.

Combining the two factors (priority and spatial modulation) yielded six experimental conditions: priority–no spatial modulation, priority–congruent spatial modulation, priority–incongruent spatial modulation, no priority–no spatial modulation, no priority–congruent spatial modulation, and no priority–incongruent spatial modulation. Each condition consists of five trials, during which all MATB subtasks were performed simultaneously according to the respective experimental condition. After each trial, participants receive a feedback score indicating how well they adhered to task prioritization instructions, based on the principles described by Stasch & Mack [29]. At the end of each block, participants complete an intermediate questionnaire assessing the perceived difficulty of the tasks (see Section 2.6). Upon finishing the experimental session, they complete a demographic questionnaire.

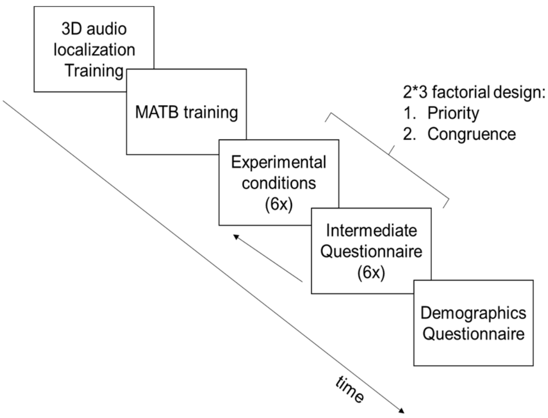

Figure 4: Experimental sequence. The six experimental conditions encompass each combination of the two factors: priority (priority of the communication task, equal focus on all tasks) and congruence (congruent audio presentation, incongruent 3D audio representation, no 3D audio).

2.7. Questionnaires

After each block, participants are asked the following two questions: "How difficult did you find the communication task?" and "How helpful did you find the spatial representation of the radio messages?". An answer can be selected on a 7-point scale (very easy to very difficult and not at all helpful to very helpful). Thus, the subjective task difficulty and perceived support was recorded. Also, everyone can provide their subjective feedback in an open-ended question. At the conclusion of the experiment, we procure demographic data, including information regarding the participants' experience of flying.

2.8. Data Analysis

2.8.1. Data Processing

Performance measures from the modified openMATB were averaged across trials for each condition and participant. Communication performance was further divided into radio performance and evade performance:

- Radio performance: Further divided into confirm-trials (when participants had to confirm the radio frequency) and reject-trials (when they had to reject it). For both, the hit proportion (correct responses) and error proportion (incorrect responses) were calculated.
- Evade performance: Measured as the hit proportion, which represents the number of correctly executed evasive maneuvers relative to the total number of evade commands.

For instances where no evasion was required in the tracking task, Root-Mean-Square Error (RMSE) was computed to quantify deviations from the center.

In the system monitoring task, performance was assessed using:

- Hit rate: The number of correctly detected changes in lights or scales.
- False alarm rate: Instances where participants incorrectly indicated a change when no input was required.

For the resource management task, the deviation from the optimal fuel level was computed.

2.8.2. Statistical Analysis

Linear Mixed Models (LMMs) were used to estimate fixed effects, group-specific variability, and to address inherent data dependencies, using the lme4 package [30] in RStudio. Participants were modeled as a random effect to account for individual differences.

Linear mixed models were chosen as the suitable statistical method because they account for both within- and between-subject variability and can appropriately handle repeated-measures data. Random slopes were not included, as we expected participants to exhibit a similar response pattern. The statistical effects of task prioritization, 3D-audio congruence (hereafter referred to as congruence), and supporter status (see Section 3.3) were examined using a stepwise model comparison approach. In cases where Model 3 provided the best fit, post-hoc comparisons were conducted using the emmeans-package in RStudio.

The models were structured as follows:

Model 0:

(1) $$y_{ij} = \beta_0 + u_{0j} + \epsilon_{ij}$$

Model 1:

(2) $$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + \epsilon_{ij}$$

Model 2:

(3) $$y_{ij} = \beta_0 + \beta_2 z_{ij} + u_{0j} + \epsilon_{ij}$$

Model 3:

(4) $$y_{ij} = \beta_0 + \beta_3 w_j + u_{0j} + \epsilon_{ij}$$

Model 4:

(5) $$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_{ij} + u_{0j} + \epsilon_{ij}$$

Model 5:

(6) $$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_{ij} + \beta_4 (x_{ij} z_{ij}) + u_{0j} + \epsilon_{ij}$$

Hereby:

y_{ij} = Outcome variable for each model at observation *i* within subject *j*

β_0 = Intercept (baseline outcome when predictors are zero)

x_{ij} = Task prioritization for observation *i* within subject *j*

z_{ij} = Congruence for observation *i* within subject *j*

w_j = Supporter status within subject *j*

$x_{ij} z_{ij}$ = Interaction term (task prioritization x congruence)

$\beta_1, \beta_2, \beta_3, \beta_4$ = Regression coefficients representing the effects of each predictor

u_{0j} = Random effect (random intercept) for subject *j*, accounting for individual differences

ϵ_{ij} = Residual error term for observation *i* within subject *j*

Specifically, the *intercept* (β_0) refers to the predicted value of an outcome when all factors (such as task prioritization or congruence) are set to zero. It acts as a baseline from which the influence of the factors of interest is measured. The *regression coefficients* (β_1, β_2, β_3, β_4) represent the strength and direction of the relationship between each factor and the respective outcome. For instance, β_1 is the regression coefficient for task prioritization and shows how much the outcome changes when task prioritization increases by one unit. The random effect (u_{0j}) accounts for individual differences between participants. Each participant may have their own baseline value (a random intercept), which means that the model allows each person to start from a slightly different point. The residual error term (ϵ_{ij}) represents the variability in the outcome that cannot be explained by the factors included in the model. It accounts for other unknown influences on the outcome that are not captured by the predictors (task prioritization, congruence, supporter status, etc.). Essentially, it measures the "noise" or unexplained variation in the data. Model fit was evaluated using the *Akaike Information Criterion (AIC)* and *Bayesian Information Criterion (BIC)*, both of which assess model adequacy while penalizing excessive complexity. Lower AIC/BIC values indicate a better model fit [31].

In the following section, M refers to the mean and SD to the standard deviation. Priority will be abbreviated with p, no priority with np. The congruence condition will be abbreviated with c, the incongruent condition with ic, and the no 3D condition with no3D. Additionally, the confidence interval (*CI*) for a fixed effect is provided for a fixed effect and represents the range of values within which the true population parameter (e.g., the true effect of task prioritization on subjective support for 3D audio) is likely to fall with 95% confidence, based on the data from the sample. The t-value (*t*) tests whether a fixed effect significantly contributes to explaining the outcome, after accounting for random effects (i.e., individual participant differences). Finally, the *p*-value (*p*) tests the null hypothesis that the fixed effect has no impact on the outcome variable. It represents the probability of obtaining a result as extreme as (or more extreme than) the observed data, assuming the null hypothesis is true. If the p-value is smaller than 0.05, it suggests that the predictor significantly contributes to the model (i.e., it has a real, non-zero effect on the outcome).

3. RESULTS

This section presents the findings of the study. The responses to the open-ended questions at the end of the experiment revealed a mixed evaluation of the 3D audio system. One subset of participants rated the 3D audio as supportive, very supportive, or highly supportive, while the other subset did not share this sentiment (see Section 3.3). To account for this variability, the sample was divided into two groups after the experiment: a 3D-audio supporter group, henceforth referred to as the supporters indicated with index s, and a 3D-audio non-supporter group, referred to as the non-supporters with index ns.

3.1. Demographics

A total of 34 participants were included in the study, with an average age of 33.76 years (*SD* = 9.43). The sample consisted of 10 females and 24 males. Three participants had prior flight experience, while the majority were members or students of the University of the Bundeswehr.

3.2. 3D audio localization training

During the 3D audio localization training, participants exhibited a mean deviation of 34.13° (*SD* = 37.96°). Front-back confusions occurred in 12.6% of trials, where sounds intended to come from the front were perceived as coming from the back, and vice versa. Such confusions negatively impacted the accuracy of localization. However, these errors can be corrected by mirroring the perceived direction, which resulted in a reduction of the mean deviation to 24.90° (*SD* = 30.57°). When considering only the last 10 trials of each participant, after applying corrections, the mean deviation was 22.58° (*SD* = 25.67°).

Although the error rates are higher than those reported in previous studies utilizing the 3D audio system [20], they align with the findings from localization trials conducted without head tracking [19]. This suggests that participants may not have fully utilized the head tracking feature, which in prior research significantly reduced localization errors. Notably, there were significant individual differences, with some participants demonstrating an average deviation as low as 3.4°. Nevertheless, given that the radio message directions in the main experiment were set at 45°, 135°, 225°, and 315°, it is assumed that participants were still able to distinguish these directions adequately.

3.3. Subjective Support of 3D Audio

Model 3 (*AIC* = 639.82, *BIC* = 666.29), including the priority and congruence predictors as well as their interaction, best explained the subjective support of the 3D audio in comparison to the null model (*AIC* = 655.82, *BIC* = 665.74). Inclusion of the priority condition as the only main effect (*AIC* = 657.73, *BIC* = 670.97) or the congruence as the only main effect (*AIC* = 648.48, *BIC* = 665.02) did not result in a better model fit than including both predictors as well as their interaction in model 3. See Fig. 5 for an illustration of that effect.

Within that model, the effect of the priority condition is statistically significant and positive (β = 0.58, 95% *CI* [0.13, 1.02], *t*(194) = 2.54, *p* = 0.012), as well as the effect of the congruent condition (β = 0.99, 95% *CI* [0.54, 1.43], *t*(194) = 4.37, *p* < .001). The no 3D condition had no effect on the subjective support of the 3D audio (β = 0.13, 95% *CI* [-0.31, 0.58], *t*(194) = 0.59, *p* = 0.558). The interaction between the priority condition and the congruent condition is statistically significant and negative (β = -1.22, 95% *CI* [-1.85, -0.60], *t*(194) = -3.84, *p* < .001), while the interaction between the priority condition and the no3D condition is statistically non-significant and negative (β = -0.61, 95% *CI* [-1.23, 0.02], *t*(194) = -1.90, *p* = 0.059). Post-hoc comparisons further illustrate this interaction effect (see file on *Pairwise comparison subjective Support* in the repository (https://osf.io/d2vgk/).

Generally, 13 participants (38%) rated the 3D audio as rather supportive, supportive, or very supportive in the congruent condition. These individuals will be referred to as the 3D audio supporter group to account for potential individual differences in the sample.

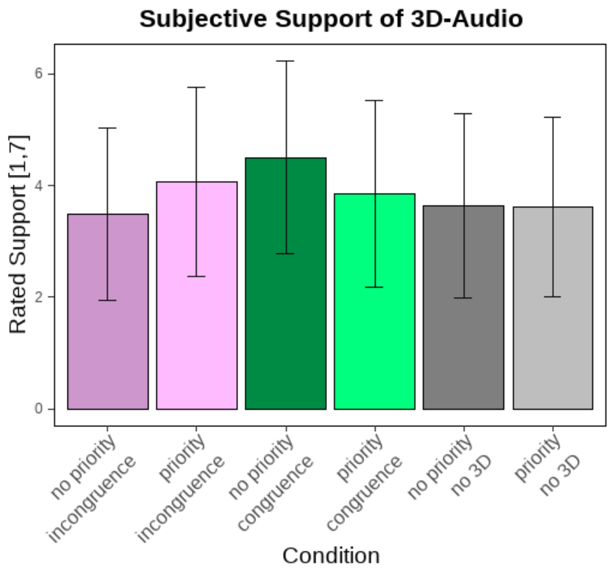

Figure 5: Rated support on a 7-point Likert scale. 1 = not at all supportive, 7 = very supportive. The bar chart illustrates the subjective rating of 3D audio support across the six experimental conditions. Error bars refer to the standard deviation of the rated support within each experimental condition.

3.4. Subjective Task Difficulty

Model 1 (*AIC* = 553.70, *BIC* = 566.94), including priority as a predictor (M_p = 3.40, SD_p = 1.29; M_{np} = 3.72, SD_{np} = 1.25), is best explaining the subjective task difficulty of the communication task (see Fig. 6) in comparison to the null model (*AIC* = 560.75, *BIC* = 570.68). Within model 1, the effect of the priority condition is statistically significant and negative (β = -0.33, 95% *CI* [-0.54, -0.11], *t*(198) = -3.04, *p* = 0.003). Including congruence as predictor (M_c = 3.57, SD_c = 1.35; M_{ic} = 3.5, SD_{ic} = 1.18; M_{no3D} = 3.6, SD_{no3D} = 1.31; *AIC* = 563.93, *BIC* = 580.47), the supporter status (M_s = 3.36, SD_s = 1.04; M_{ns} = 3.68, SD_{ns} = 1.39, *AIC* = 561.93, *BIC* = 575.16) or the interaction (*AIC* = 559.45, *BIC* = 585.92) did not improve model fit in comparison to the null model.

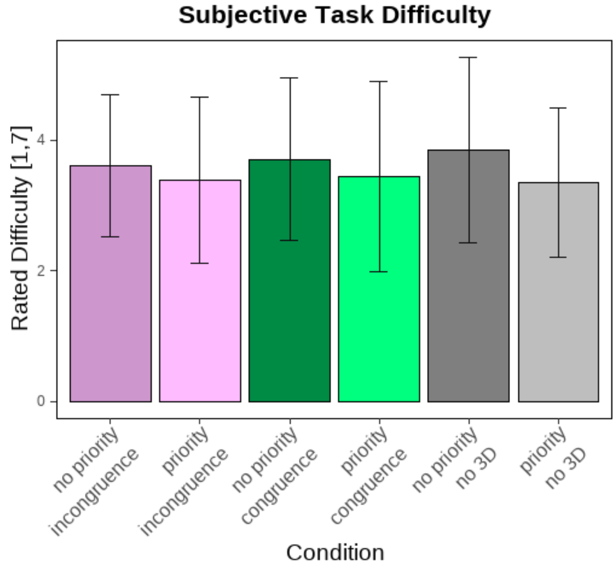

Figure 6: Rated communication task difficulty on a 7-point Likert scale. 1= very difficult, 7 = very easy.

3.5. Communication Task Performance

3.5.1. Confirmation of correct radio calls - Hits

Model comparison indicates that neither the model including priority as a predictor (M_p = 0.91, SD_p = 0.15; M_{np} = 0.92, SD_{np} = 0.13; *AIC* = -231.19, *BIC* = -217.95), nor congruence (M_c = 0.93, SD_c = 0.15; M_{ic} = 0.89, SD_{ic} = 0.16; M_{no3D} = 0.92, SD_{no3D} = 0.11; *AIC* = -232.78, *BIC* = -216.23), nor the model including the interaction of both factors (*AIC* = -220.42, *BIC* = -174.11) explains the hit proportion of confirming correct radio calls better than the null model (*AIC* = -233.09, *BIC* = -223.16).

Exploratory analysis of the descriptive statistics of model 1, including congruence as a predictor, indicates that 3D audio has the tendency to improve the hit rate of correct frequencies in the congruent condition. However, this effect is statistically non-significant on a $p < 0.05$ significance level in that model (β = 0.04, 95% *CI* [-2.45e^{-03}, 0.08], *t*(197) = 1.86, *p* = 0.065). Similarly, the effect of the no 3D condition is statistically non-significant (β = 0.03, 95% *CI* [-0.01, 0.07], *t*(197) = 1.36, *p* = 0.176) in that model. Respecting the supporter group status as a predictor did not lead to any model improvement (*AIC* = -220.42, *BIC* = -174.11; see Fig. 7 for an illustration of the effect) in comparison to the null model.

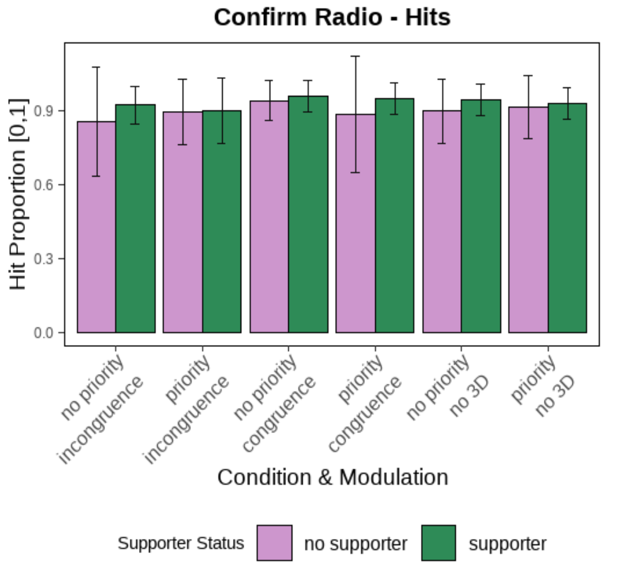

Figure 7: Hit proportion of confirming correct radio calls.

3.5.2. Confirmation of correct radio calls - Errors

The wrong proportion when confirming correct radio calls is best explained by the null model (*AIC* = -364.98, *BIC* = -355.05), indicating that neither priority (M_p = 0.06, SD_p = 0.01; M_{np} = 0.05, SD_{np} = 0.01; *AIC* = -363.76, *BIC* = -350.53) nor congruence (M_c = 0.05, SD_c = 0.01; M_{ic} = 0.07, SD_{ic} = 0.1; M_{no3D} = 0.05, SD_{no3D} = 0.01; *AIC* = -364.85, *BIC* = -348.31) had a significant influence on the wrong proportion on a *p* =0.05 level. However, exploratory analysis revealed that the effect of the congruent condition is statistically significant at a *p* = 0.1 level and negative (β = -0.03, 95% *CI* [-0.06, 1.99e-03], *t*(197) = -1.85, *p* = 0.066). The effect of the no 3D condition is within that model statistically non-significant and negative (β = -0.02, 95% *CI* [-0.06, 7.14e-03], *t*(197) = -1.52, *p* = 0.130). Including the supporter status as a predictor in the model comparison did not yield any model improvement (*AIC* = -365.07, *BIC* = -351.84; see Fig. 8 for an illustration of that effect). Also, the inclusion of the interaction between priority and congruence did not explain the data better than the null model (*AIC* = -359.88, *BIC* = -333.41).

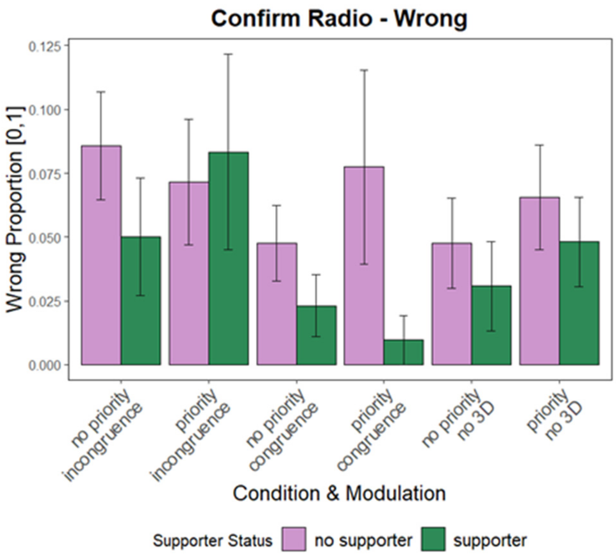

Figure 8: Wrong proportion of confirming correct radio calls.

3.5.3. Rejection of wrong radio calls - Hits

The correct rejection of wrong radio calls was best explained by the null model (*AIC* = -150.62, *BIC* = -140.69), showing that neither including priority (M_p = 0.85, SD_p = 0.25; M_{np} = 0.87, SD_{np} = 0.26; *AIC* = -149.56, *BIC* = -136.33) nor congruence (M_c = 0.87, SD_c = 0.25; M_{ic} = 0.85, SD_{ic} =0.27; M_{no3D} = 0.87, SD_{no3D} = 0.25; *AIC* = -147.60, *BIC* = -131.06) explained the data better. Including the supporter status did also not yield any model improvement (M_s = 0.92, SD_s = 0.1; M_{ns} = 0.82, SD_{ns} = 0.31, *AIC* = -149.76, *BIC* = -136.53), similar to the interaction between congruence and priority (*AIC* = -143.53, *BIC* = -117.07). For conciseness, the figures for the remaining categories have been omitted. The same analytical approach applies, and the corresponding statistical results are described in the text. Interested readers are invited to consult the repository for further details (https://osf.io/d2vgk/).

3.5.4. Rejection of wrong radio calls - Errors

The errors in rejecting wrong radio calls was best explained by the model including task priority (M_p = 0.06, SD_p = 0.1; M_{np} = 0.04, SD_{np} = 0.08; *AIC* = -374.39, *BIC* = -361.16) as a factor in comparison to the null model (*AIC*=-371.64, *BIC* = -361.72). Within that model, the effect of the priority condition is statistically significant and positive (β = 0.03, 95% *CI* [2.80e-03, 0.05], $t(198)$ = 2.19, p = 0.030). Including congruence (M_c = 0.05, SD_c = 0.1; M_{ic} = 0.06, SD_{ic} =0.1; M_{no3D} = 0.05, SD_{no3D} = 0.08; *AIC* = -367.99, *BIC* = -351.44) or the supporter status (M_s = 0.05, SD_s = 0.07; M_{ns} = 0.05, SD_{ns} = 0.11, *AIC* = -369.70, *BIC* = -356.47) did not yield any model improvement, similar to the interaction between congruence and priority (*AIC* = -367.65, *BIC* = -341.18).

3.6. Tracking Task Performance

3.6.1. Evade

The number of correctly performed evade commands was best explained by the null model (*AIC* = -239.66; *BIC* = -229.73), neither including priority (M_p = 0.92, SD_p = 0.19; M_{np} = 0.94, SD_n = 0.18; *AIC* = -238.25, *BIC* = -225.02;) or congruence (M_c = 0.93, SD_c = 0.2; M_{ic} = 0.96, SD_{ic} = 0.11; M_{no3D} = 0.91, SD_{no3D} = 0.22; *AIC* = -237.71, *BIC* = -221.17) as a factor. The inclusion of the supporter group status did not lead to a model improvement (M_s = 0.92, SD_s = 0.23; M_{ns} = 0.94, SD_{ns} = 0.15; *AIC* = -238.01, *BIC* = -224.78). Also, the model including the interaction between congruence and priority (*AIC* = -234.80, *BIC* = -208.33) did explain the data better than the null model.

3.6.2. RMSE

The RMSE was best explained by the null model (*AIC* = -1034, *BIC*= -1024.08), neither including priority (M_p = 0.08, SD_p = 0.03; M_{np} = 0.08, SD_{np} = 0.03; *AIC* = -1033.1, *BIC* = -1019.79) or congruence (M_c = 0.08, SD_c = 0.03; M_{ic} = 0.08, SD_{ic} =0.03; M_{no3D} = 0.08, SD_{no3D} = 0.03; *AIC* = -1030.6; *BIC* = -1014.05) as a factor. The inclusion of the supporter group status did not lead to a model improvement (M_s = 0.09, SD_s = 0.03; M_{ns} = 0.08, SD_{ns} = 0.03; *AIC* = -1035.9, *BIC* = -1022.64). Including the interaction between congruence and priority did also not improve the model fit (*AIC* = -1025.8, *BIC* = -999.22).

3.7. System Monitoring Task Performance

3.7.1. Hits

The number of hits in the system monitoring task was best explained by model 1 including priority (M_p = 4.98, SD_p = 1.07; M_{np} = 5.20, SD_{np} = 1.02; *AIC* = 499.20, *BIC* = 512.35) as a predictor compared to the null model (*AIC* = 502.17, *BIC* = 512.03). Within that model, the effect of the priority condition is statistically significant and negative (β = -0.22, 95% *CI* [-0.42, -0.03], $t(194)$ = -2.24, p = 0.026). The inclusion of the congruence condition (M_c = 5.1, SD_c = 1.07; M_{ic} = 5.14, SD_{ic} =1.02; M_{no3D} = 5.03, SD_{no3D} = 1.07; *AIC* = 505.32, *BIC* = 521.76) did not yield any model improvement. Including neither the interaction between congruence and priority (*AIC* = 506.30, *BIC* = 532.61) nor the supporter group status (M_s = 4.84, SD_s = 1.31; M_{ns} = 5.23, SD_{ns} = 0.83; *AIC* = 502.34, *BIC* = 515.50) explained the data better than the null model.

3.7.2. False Alarm

The number of false alarms on the system monitoring task was best explained by the null model (*AIC* = 561.66, *BIC* = 571.53), neither including priority (M_p = 0.75, SD_p = 1.11; M_{np} = 0.68, SD_{np} = 1.17; *AIC* = 563.30, *BIC* = 576.46), congruence (M_c = 0.76, SD_c = 1.09; M_{ic} = 0.63, SD_{ic} = 1.27; M_{no3D} = 0.76, SD_{no3D} = 1.06; *AIC* = 564.68, *BIC* = 581.12), the supporter status (M_s = 0.94, SD_s = 1.41; M_{ns} = 0.59, S_{ns} = 0.93; *AIC* = 562.28, *BIC* = 575.44), or the interaction between priority and congruence (*AIC* = 567.36, *BIC* = 593.67) as a predictor.

3.8. Resource Management Task Performance

3.8.1. Deviation from optimal fuel level

The average deviation from the optimal fuel of the resource management task is best explained by model 1 including the priority as a predictor (M_p = 537, SD_p = 179; M_{np} = 319, SD_{np} = 83.2; *AIC* = 2455.2, *BIC* = 2468.4) compared to the null model (*AIC* = 2599.6, *BIC* = 2609.5). Within that model, the effect of the priority condition is statistically significant and positive (β = 217.82, 95% *CI* [189.75, 245.89], $t(194) = 15.31$, $p < .001$). Including congruence (M_c = 447, SD_c = 190; M_{ic} = 411, SD_{ic} = 164; M_{no3D} = 426, SD_{no3D} = 177; *AIC* = 2601.8, *BIC* = 2618.3), the supporter status (M_s = 445, SD_s = 190; M_{ns} = 419, SD_{ns} = 169; *AIC* = 2601.2, *BIC* = 2614.3), or the interaction between priority and congruence (*AIC* = 2458, *BIC* = 2484.3) as a predictor did not explain the data better than the null model.

3.9. Open-ended Question

At the end of the experiment, participants were required to answer the following open-ended questions: Question 1: “*How did the spatial-acoustic sound presentation help you with the flight tasks?*” and Question 2: “*How did the spatial-acoustic sound presentation disturb you when performing the flight tasks?*”.

For the first question, participants' answers were categorized into generally supportive or unsupported answers. Among the generally supportive answers, six out of the 34 respondents surveyed indicated that the 3D audio presentation was primarily helpful for orientation. Three participants mentioned that the 3D audio presentation was overall helpful/supportive. The answer, that left/right differences were recognizable and helpful, but not front/back differences, was given by three participants. One participant indicated that it was helpful in the congruent condition. Among the unsupportive answers, ten participants indicated that the spatial modulation of the radio messages was of little to no help/support. Four participants indicated that they focused only on the content, but not the direction of the radio messages. One participant only became aware of the 3D audio through the questions of the intermediate questionnaire, and one participant had no trust in the 3D audio presentation due to the incongruent condition within the experiment. Regarding the second question, all participants agreed that the 3D audio was not disturbing when performing the flight tasks.

4. DISCUSSION

4.1. Discussion of results

The present study aimed to investigate whether 3D modulated radio calls (levels: congruent, incongruent, no 3D) influence multitasking performance under different task prioritization conditions (levels: priority on communication task, no preassigned priority) in a simulated formation flight using the modified open-source version of the MATB, the openMATB. The goal was to determine whether the positive effects of 3D radio messages also affect task performance under demanding multitasking conditions.

The subjective ratings of support revealed individual differences in the perception of 3D radio messages, with approximately one-third of participants rating the messages as at least somewhat supportive, while the remainder did not report any perceived support. This result was unexpected, as previous research has generally demonstrated a positive impact of 3D audio messages on various outcomes. To control for individual differences, this factor was included as a control variable in the subsequent analysis, labeled as *supporter status*. However, that control variable did not significantly influence multitasking performance or subjective task difficulty in the communication task.

Instead, the perception of support was best explained by the interaction between the congruence of the 3D audio modulation and the prioritization of the communication task. Specifically, 3D radio messages were rated as more supportive when participants prioritized the communication task compared to when no task prioritization was assigned. This finding suggests that task engagement enhances the perception of spatial audio cues as helpful.

Furthermore, the perception of support was best explained by the interaction between the congruence of the 3D audio modulation and the prioritization of the communication task. Specifically, 3D radio messages were rated as more supportive when participants prioritized the communication task compared to when no task prioritization was assigned. This result partly confirms *H1*. However, no difference was found between the no 3D condition and the incongruent condition, which was contrary to *H3*. This suggests that participants were not misled by the false spatial cues in the incongruent condition but likely focused on the content of the radio messages, consciously or unconsciously ignoring the incorrect spatial cue.

The negative interaction between task priority and congruence indicates that the positive effect of congruently modulated 3D radio messages was less pronounced when participants prioritized the communication task. In other words, 3D radio messages were perceived as most supportive when presented spatially correct and when task prioritization was not explicitly assigned. This suggests that spatially accurate 3D radio messages are perceived under subliminal conditions, but only when the task is not the primary focus. When participants focused on the task, the effect was reduced, likely due to limited auditory resources available to process that spatial information [6]. Possibly, congruently modulated 3D radio messages might provide their maximum benefit when cognitive demands on the auditory channel are moderate and the task is not explicitly prioritized. Alternatively, participants may adopt a task-specific strategy under task prioritization that emphasizes content over spatial presentation, consciously or unconsciously ignoring spatial information, which diminishes the perceived or actual supportiveness of congruent 3D audio.

However, do the positive effects of correctly modulated 3D audio messages, as observed subjectively, also translate to measurable performance improvements? Neither the hit nor error rate in confirming correct radio calls or rejecting incorrect radio calls was positively influenced by congruently presented radio calls or negatively influenced by incorrectly presented radio calls. This lack of effect on task performance contradicts previous findings [e.g. 16]. It suggests that participants concentrated more on the content of the radio messages rather than utilizing the spatial cues. The increased auditory load [6] could explain this: when participants focus on processing the content, they may not fully process the spatial cues, consistent with the assumption that cognitive resources are limited. This raises the question for future research of how much spatial information was cognitively processed, especially since participants did report a positive effect of 3D radio messages in the post-task questionnaires.

Interestingly, task prioritization alone accounted for the number of errors in rejecting incorrect radio messages, with participants making more errors when prioritizing the communication task compared to when no such priority was assigned. This result is unexpected, as prioritizing the task should have provided participants with more cognitive resources to avoid these errors. An internal shift in decision criteria [28] may explain this: participants, focusing on the communication task, could adjust their thresholds for rejecting radio messages. This shift in response bias might make participants more likely to incorrectly reject a message if its spatial cues conflict with expectations, even when the message is not incorrect. Such adjustments could lead to errors in interpreting or responding to radio messages, ultimately reducing performance on the communication task.

Could the performance in other flight tasks have benefited from the 3D radio messages, even if the communication task did not? This was also not the case. Instead, task prioritization seemed to influence the performance of other tasks. While no effects of priority or congruence were detected on the performance of the tracking task (evade, RMSE), prioritizing the communication task led to poorer performance in confirming correct radio calls and greater deviation from the optimal fuel level in the resource management task. This effect is consistent with previous research by Stasch and Mack [18, 19], who showed that prioritizing one task (e.g., tracking) can reduce performance on other tasks (e.g., system monitoring or resource management). Fewer cognitive resources may have been available to process other tasks, leading to increased task interference and decreased performance. In summary, while the positive effects of 3D radio messages were evident subjectively, they did not influence objective performance measures. However, no negative effect of incorrectly presented radio messages was observed, which raises questions about how much participants actually relied on the information provided by the spatial cues. Task priority, however, emerged as a significant factor in multitasking research and should be considered in future studies.

4.2. Limitations

The study presents several limitations that may help explain the absence of a positive effect of 3D audio on performance. First, the use of the 3D audio system in this study builds upon previous research demonstrating improved sound localization [23, 24]. However, prior studies focused exclusively on localization tasks, which were conducted outside of the multitasking context examined in the present study. This singular focus, coupled with extended training, likely enabled participants to adapt more effectively to the spatial audio system. In contrast, the current study introduced a multitasking scenario that required participants to manage multiple tasks simultaneously. This likely increased cognitive load, which may have reduced the amount of auditory resources available for processing spatial auditory cues, thus diminishing the benefits of the 3D audio system observed in earlier research.

Furthermore, the inclusion of incongruent audio modulations could have impacted participants' trust in the spatial audio system. Over the course of the experiment, participants may have perceived the spatial information as unreliable, leading them to focus more on the content of the radio messages rather than utilizing the spatial cues. Another significant limitation lies in the task demands of the modified communication task involving simple radio messages. While this task aimed to engage the spatial effects of the 3D audio system, it did not fully exploit the potential of spatial cues in a more complex task setting. The demand of comparing the announced frequency to the displayed frequency required minimal spatial reasoning, which likely led to the underutilization of the spatial cues provided by the 3D audio system. As a result, participants may have prioritized the content of the radio messages over their spatial presentation, reducing the system's effectiveness in this study.

The demands of the collision avoidance (evade) task, which was initially included to introduce a more spatially demanding task, also presented limitations. Although the task required participants to respond based on spatial positioning within the simulated formation, the flight environment used in the MATB was highly controlled and only involved flight-specific tasks. Additionally, the MATB was presented on a normal-sized screen, which required minimal head movement. However, head movements play an essential role in perceiving 3D audio cues [34], and this limitation restricts the generalizability of the current findings to real-world cockpit environments.

4.3. Future research

Future studies should address these limitations by exploring the potential of 3D audio systems in higher-fidelity flight environments with varying workload levels. Additionally, a between-subjects design with a larger sample size would help mitigate the confounding factor of mistrust in the 3D audio system, which may have been influenced by the incongruent condition. This approach could alleviate doubts about the system's reliability and encourage participants to more consistently rely on spatial cues.

A more immersive and realistic experimental setup, such as a virtual reality (VR) flight simulation or high-fidelity flight simulator, could further enhance the ecological validity of future research. VR environments would enable the integration of more realistic flight scenarios that simulate genuine spatial and temporal task demands, such as avoiding mid-air collisions or maintaining formation alignment. Incorporating these scenarios would naturally highlight the potential benefits of spatial audio in enhancing task performance.

5. CONCLUSION

The study suggests that 3D audio offers subjective benefits but does not translate into measurable performance improvements in a controlled multitasking flight environment. However, some participants, despite receiving extensive training on 3D audio localization, reported difficulties in correctly perceiving the spatial accuracy of the 3D radio messages, as revealed in open-ended responses at the conclusion of the experiment. This may have been due to the limited head movements afforded by the experimental setup. Therefore, future research conducted in a realistic cockpit environment, allowing for more natural head movements with an expert sample, is essential to better understand the potential of 3D audio in cognitively demanding multitasking environments.

6. REFERENCES

[1] Donath, D., Schulte, A.: Behavior Based Task and High Workload Determination of Pilots Guiding Multiple UAVs. Procedia Manufacturing. 3, 990-997. https://doi.org/10.1016/j.promfg.2015.07.156 (2015)

[2] Skiba, T.: Sie trainierten für die Premiere und stürzten in den Tod. (2022). Bundeswehr. Retrieved February 18, 2025, from: https://www.bundeswehr.de/de/organisation/luftwaffe/aktuelles/sie-trainierten-fuer-die-premiere-und-stuerzten-in-den-tod-5449260

[3] CNN: Malaysian navy helicopters collide in mid-air, killing 10 crew. (2024). Retrieved February 18, 2025 from: https://edition.cnn.com/2024/04/22/asia/malaysia-lumut-navy-helicopter-crash-intl-hnk/index.html

[4] Lahoti, R. C.: Report of Court of Inquiry on Mid-Air Collision Between Saudi Arabian Boeing 747 and Kazakhstan IL-76 on 12th November, 1996 Near Delhi – India (Charkhi-Dadri, Haryana), 1997

[5] Salvucci, D. D., Taatgen, N.A.: The multitasking mind. University Press, Oxford (2011)

[6] Wickens, C.D.: The structure of attentional resources. In: Nickerson, R. (ed.) Attention and performance, 8, pp. 239-257. Lawrence Erlbaum, Hillsdale (1980)

[7] Funk, K.: Cockpit task management: Preliminary definitions, normative theory, error taxonomy, and design recommendations. The International Journal of Aviation Psychology. 1(4), 271-285 (1991)

[8] Chou, C. C., Madhavan, D., Funk, K.: Studies of cockpit task management errors. The International Journal of Aviation Psychology. 6(4), 307-320 (1996)

[9] Niermann, C.: Can spatial audio support pilots? 3D-Audio for future pilot-assistance systems. Digital Avionics Systems Conference (DASC) (2015). https://doi.org/10.1109/DASC.2015.7311401

[10] Peng, X., Jiang, H., Yang, J., Shi, R., Feng, J., Liang, Y.: Effects of Temporal Characteristics on Pilots Perceiving Audiovisual Warning Signals Under Different Perceptual Loads. Front. Psychol. 13 (2022). https://doi.org/10.3389/fpsyg.2022.808150

[11] Haas, E.C.: Can 3-D auditory warnings enhance helicopter cockpit safety? Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 42(15), 1117-1121 (1998). https://doi.org/10.1177/154193129804201513

[12] Rabl, A.N., Solbeck, C.: Experimental evaluation of a multimodal spatial display in an aircraft cockpit context. 33rd Conference of the European Association for Aviation Psychology (2018)

[13] Oving, A.B., Veltman, J.A., Bronkhorst, A.W.: Effectiveness of 3–D Audio for Warnings in the Cockpit. IN: Int. J. Aviat. Psychol. 14(3), 257-276 (2004). https://doi.org/10.1207/s15327108ijap1403_3, ISSN: 1050-8414

[14] Begault, D.R., Pittmann, M.T.: Three-dimensional audio versus head-down Traffic Alert and Collision Avoidance System displays. The International Journal of Aviation Psychology. 6(1), 79-93 (1996)

[15] Milam, L., Akins, E., Williams, H., Simpson, B., Jones, H.G.: Spatial audio cueing aids pilot navigation during simulated helicopter flight in degraded visual environments. CA SAGE. 46, 1634-1638 (2019)

[16] Brock, D., Stroup, J.L., Ballas, J.A.: Effects of 3D auditory display on dual task performance in a simulated multiscreen watchstation environment. In Proceedings of the human factors and ergonomics society annual meeting. 46, pp. 1570-1573 (2002)

[17] Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D.: Conflict monitoring and cognitive control. Psychological Review. 108(3), 624–652 (2001). https://doi.org/10.1037/0033-295X.108.3.624

[18] Stasch, S., Mack, W.: Why the stability-flexibility-dilemma should be taken into consideration when studying pilots multitasking behaviour. AHFE International, USA (2023). http://doi.org/10.54941/ahfe1003846

[19] Stasch, S., Mack, W., Hilla, Y.: Always in (partner) action? Working in teams may improve simulated flight performance – but only in the apt cognitive control state. Preprint (2024). https://doi.org/10.31234/osf.io/2kajt

[20] Stasch, S.-M., von Mankowski, J., Mack, W.: Extending gamification for multitasking strategies – a study on auditory task prioritization. In: de Waard, D., Onnasch, L., Liebherr, M., Dettmann, A., Siebert, F., Karrer-Gauß, K., Winkler, S., Toffetti, A., de Angelis, M. (eds.) Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting. (April 2025). https://www.hfes-europe.org

[21] Cegarra, J., Valéry, B., Avril, E, Calmettes, C., Navarro, J.: OpenMATB: A Multi-Attribute Task Battery promoting task customization, software extensibility and experiment replicability. Behav Res 52, 1980–1990 (2020). https://doi.org/10.3758/s13428-020-01364-w

[22] Comstock Jr, J.R., Arnegard, R.J.: The multi-attribute task battery for human operator workload and strategic behavior research. Hampton, Virginia (1992)

[23] Ernst, F., Sachau, D.: Räumliches Audiosystem für Headsets mit aktiver Schallreduktion für Hubschrauberpiloten. Fortschritte der Akustik - DAGA (2023)

[24] Ernst, F., Sachau, D.: Influence of background noise on spatial audio localization for helicopter pilot headsets. Proc. Mtgs. Acoust (2023). https://doi.org/10.1207/s15327108ijap1403_3,10.1121/2.0001904

[25] Blauert, J.: Räumliches Hören. Hirzel Verlag, Stuttgart (1974)

[26] Majdak, P., Goupell, M.J., Laback, B.: 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. Attention, Perception, & Psychophysics. 72, 454–469 (2010)

[27] Opentrack 2024, Version 2023.3.0, software available at https://github.com/opentrack/opentrack

[28] Begault, D.R., Wenzel, E.M., Anderson, M.R.: Direct comparison of the impact of head tracking, reverberation and individualized head-related transfer functions on the spatial perception of a virtual speech source. Journal of the Audio Engineering Society 49 (2001) 904–916.

[29] Stasch, S., Mack, W.: A new experimental method to investigate multitasking strategies in flight environments via the use of gamification. Proceedings of the Human Factors and Ergonomics Society Europe Chapter. Annual Conference (2023).

[30] Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 67(1), 1–48 (2015)

[31] Müller, S., Scealy, J.L.,Welsh, A.H.: Model Selection in Linear Mixed Models. Statistical Science, 28(2), 135–167 (2013)

[32] Green, D. M., Swets, J. A.: Signal detection

theory and psychophysics. John Wiley (1966)
[33] Higgins, N.C., Pupo, D.A., Ozmeral, E.J., Eddins, D.A.: Head movement and its relation to hearing. Frontiers in Psychology (2023). https://doi.org/10.3389/fpsyg.2023.1183303

Corresponding author

sophie-marie.stasch@unibw.de