



OPEN ACCESS

EDITED BY

James Harland,
RMIT University, Australia

REVIEWED BY

Katharina Petri,
Otto von Guericke University Magdeburg,
Germany
Radu Petrușe,
Lucian Blaga University of Sibiu, Romania

*CORRESPONDENCE

Fiona Duruaku,
✉ Fiona.duruaku@ucf.edu

RECEIVED 28 March 2025

ACCEPTED 23 June 2025

PUBLISHED 14 July 2025

CITATION

Duruaku F, Sims V and Jentsch FG (2025)
Individual differences influencing procedural
learning outcomes in virtual reality: a case study
on an exterior preflight inspection.
Front. Virtual Real. 6:1601562.
doi: 10.3389/frvir.2025.1601562

COPYRIGHT

© 2025 Duruaku, Sims and Jentsch. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Individual differences influencing procedural learning outcomes in virtual reality: a case study on an exterior preflight inspection

Fiona Duruaku^{1*}, Valerie Sims² and Florian G. Jentsch¹

¹Team Performance Laboratory, Department of Psychology and Institute for Simulation & Training, University of Central Florida, Orlando, FL, United States, ²Applied Cognition and Technology Group, Department of Psychology, University of Central Florida, Orlando, FL, United States

Introduction: Virtual reality (VR) has been increasingly used across safety-critical industries for training procedures because it allows for practice without real-world risks. Its effectiveness may be further influenced by individual differences. This paper examined technology features, including immersion and interactivity, and individual differences factors, specifically sex, spatial ability and personality traits, that could affect learning in VR, particularly within the context of procedural training. The study aimed to understand how VR functions and to identify who benefits most from its use.

Methods: In the experiment, 79 undergraduate students were trained to conduct an exterior preflight inspection of a passenger aircraft in VR, with varying levels of immersion (desktop PC vs. immersive VR) and interactivity (passive learning vs. active exploration). Participants were randomly assigned to one of four training groups: PC Passive, PC Active, VR Passive, and VR Active. The PC group used a mouse and keyboard, while the VR group used a head-mounted display and hand controllers to interact with the VR environment. Individual differences in sex, spatial ability, and personality traits were also investigated to determine their effects on procedural learning outcomes. Learning outcomes were assessed using two measures: a practical assessment using the desktop PC or immersive VR and a post-knowledge test. Data analyses were conducted using analyses of covariance (ANCOVAs) to examine the individual and combined effects of interactivity and immersion on procedural learning outcomes while controlling for pre-knowledge test scores. Additionally, stepwise multiple regression analyses were employed to evaluate the effects of individual differences on procedural learning.

Results: The results indicated no difference in procedural learning outcomes across the levels of immersion and interactivity. Specific individual differences, including sex, and spatial ability, however, significantly predicted VR procedural learning outcomes.

Discussion: Our findings challenge the assumption that higher immersion and higher interactivity alone, or in combination, always lead to better procedural

learning outcomes. Furthermore, the study emphasizes the importance of considering individual differences when implementing VR in learning environments, as they play a critical role in shaping learning outcomes.

KEYWORDS

virtual reality, procedural training, individual differences, immersion, interactivity, technology-human interaction

1 Introduction

With the advancements in technology, immersive virtual reality (IVR) has become popular in research, education and training. The ability to create immersive and interactive learning environments offers unique opportunities for enhancing skill acquisition and retention. VR, as used here, implies “reality simulated virtually” (Li et al., 2020), and it affords learners in various fields the opportunity to acquire skills in a fully immersed and highly interactive environment.

Learners can progress at their own pace, benefiting from personalization and training adaptability, for example, by using gamification techniques (Maroukias et al., 2024). VR has been recognized as an effective educational tool because it enables users to visualize, explore, manipulate, and interact with objects and environments in a simulated, computer-generated space, which may encourage deeper learning (Parong and Mayer, 2018; Petersen et al., 2022). Safety-critical industries such as aviation (Guthridge and Clinton-Lisell, 2023), medicine (Li et al., 2017) and construction (Akindele et al., 2024) often require that certain skills be mastered before they can be performed efficiently in real-world settings. These industries have benefitted from the realism offered by learning procedures in VR, taking full advantage of the agency and fidelity it affords learners (Scorgie et al., 2024).

Using virtual, immersive, and interactive simulations, trainees can learn in realistic simulated experimental environments and examine various phenomena, thereby advancing their cognitive engagement and understanding (Lin et al., 2024; Wang et al., 2024). Since practice is an important part of procedural training (Ganier et al., 2014; Lorenzis et al., 2023), and VR requires some activity from the learner, it has been found that VR could serve as a valuable resource for learning procedures, offering several advantages and applications in comprehension and retention (Radianti et al., 2020). This has made VR particularly valuable in fields where hands-on experience is crucial, offering a safe and controlled environment for learners to master complex procedures before applying them in actual work settings. VR can help pilots learn to perform risky maneuvers (Hight et al., 2022), doctors can utilize VR for surgical training (Ntakakis et al., 2023), and those in manufacturing can use VR for training on emergency drills (Scorgie et al., 2024). However, studies have reported mixed findings on how the technology features of VR—immersion and interactivity—affect procedural learning outcomes.

Researchers have demonstrated that IVR can effectively enhance motivation and engagement (Makransky et al., 2019), but it can also pose challenges during learning, particularly because learners may become distracted by elements unrelated to the instructional goals, thus increasing extraneous cognitive load. In VR, the immersive visual experience can impose an additional visual burden on learners

(Mayer et al., 2023; Yang et al., 2023). This heightened visual demand is also associated with simulation sickness, which can negatively impact learning procedures.

Other researchers have found no significant differences in learning procedures in immersive three-dimensional versus non-immersive two-dimensional VR environments (Barrett et al., 2022; Urhahne et al., 2011). Yet others have found positive effects of immersive VR in learning procedures (Coban et al., 2022; Hamilton et al., 2021). Furthermore, the high interactivity afforded by VR is believed to enable learners to take on an active role in their learning, allowing them to navigate through the material at their own pace. If the interactions help in the generative processing of the content, it may be more effective than just being in an immersive condition (Jang et al., 2017; Johnson et al., 2022). However, Khorasani et al. (2023) argued that increased interactivity does not necessarily lead to better learning performance, particularly for procedural training tasks. They emphasized the need to consider factors such as the complexity of interactions, the alignment between the technology and the learning task, and the cognitive demands placed on learners.

While previous studies have explored how immersion and interactivity impact learning in VR environments, only very few investigations have addressed both immersion and interactivity. One notable study in this context is Johnson et al. (2022), who investigated the effects of immersion (desktop display vs. HMD) and interaction modality (gesture vs. voice) on procedural learning and also studied how the effects were moderated by spatial ability. Their findings suggested that gesture-based VR may be particularly beneficial for learners with lower spatial ability. However, the generalizability of their work was constrained by a focus on a narrow task type, a limited participant sample, and reliance on written recall measures alone. In contrast, our present study integrated both system-level manipulations (immersion $\times$ interactivity) and person-level factors (sex, spatial ability, and personality traits) within a 2×2 factorial design, using a realistic aviation inspection task. Moreover, learning outcomes were assessed using both practical performance and knowledge-based measures, offering a more ecologically valid understanding of procedural training effectiveness in VR.

2 Related work

2.1 VR in procedural learning

Learning procedures involve mastering a sequence of tasks or steps, often with a focus on accuracy and efficiency. It involves acquiring skills and knowledge through the repetition of specific tasks. Training using VR is especially effective for learning procedures, as it supports skill acquisition, retention, and

TABLE 1 Description of exterior pre-flight inspection procedural task in the simulation.

Task	Description	Duration
Task 1: Tutorial	Participants learn simulation controls in both desktop virtual reality (DVR) and immersive virtual reality (IVR). Tricks include teleporting, movement, pointing to objects and other interface interactions	5 min
Task 2: Training	Participants perform an exterior preflight inspection: Learning to identify faults, errors, and damage on a simulated Boeing 737. They are placed in either a passive or active condition.	20 min
Task 3: Assessment	Participants apply learned steps from the training phase to inspect the aircraft, identifying faults, missing parts, or damage. This is done using the mouse in DVR or VR controllers in IVR. Procedural learning was measured by calculating the total number of faults or damages accurately identified on the aircraft	10 min

transfer, and can improve both the effectiveness and efficiency of training (Chang and Hung, 2019; Jongbloed et al., 2024). In the study described here, the procedure to be learned was an exterior preflight inspection (Table 1 provides details of the training task). An exterior preflight inspection is a visual examination of the external components of an aircraft to ensure they are in good condition for safe flight operations.

VR may be particularly effective for learning procedures (Conrad et al., 2024; Jongbloed et al., 2024), with significant benefits in helping learners acquire, retain, and apply new skills (Hamilton et al., 2021). Research shows that VR can improve the effectiveness and efficiency of skill development. However, it may not be suitable for all learners, as individual differences can influence learning outcomes. Understanding how learner characteristics affect outcomes can help tailor the learning content and process to accommodate these differences. Therefore, the study addresses the following research questions:

- RQ1: Are there mean differences in learning procedures using VR associated with immersion and interactivity (technology factors), after adjusting for differences in pre-knowledge test scores?
- RQ2: Do individual differences factors (sex, spatial ability, personality traits) predict procedural learning outcomes in VR?

While existing research has explored certain individual differences such as sex and spatial ability in VR, there remains a gap in our understanding of how these factors specifically influence procedural learning. There has been a disconnect between the technology and the individual characteristics of the user that could impact procedural learning, and this poses a challenge that needs to be addressed. As a result, the goal of this paper is to redress this issue by examining VR effectiveness for training procedures from the individual learner's perspective. The aim is to contribute to the literature on how technology-learner interaction may lead to better outcomes for trainees. The focus was first on isolating the technology factors of VR (immersion and interactivity) by investigating whether immersive virtual reality (using a head-mounted display and controllers) was more effective than desktop virtual reality (using a personal computer, keyboard, and mouse) for learning a procedure. Then, we examined how certain individual characteristics predicted learning outcomes. Understanding these key variables in VR learning can inform both research and practical applications for procedures training. This approach was taken because the effectiveness of VR training may vary depending on individual learner factors and the specific VR implementation. Note: While VR is widely used across various industries, this study focuses on its application for training purposes.

To effectively answer the research questions, the study integrated the technological affordances of VR with individual learner characteristics, drawing on Salzman et al. (1999) technology media model, which emphasized the input, process, and output mechanisms (Figure 1 details the technology media model).

2.2 Technology factors in VR learning

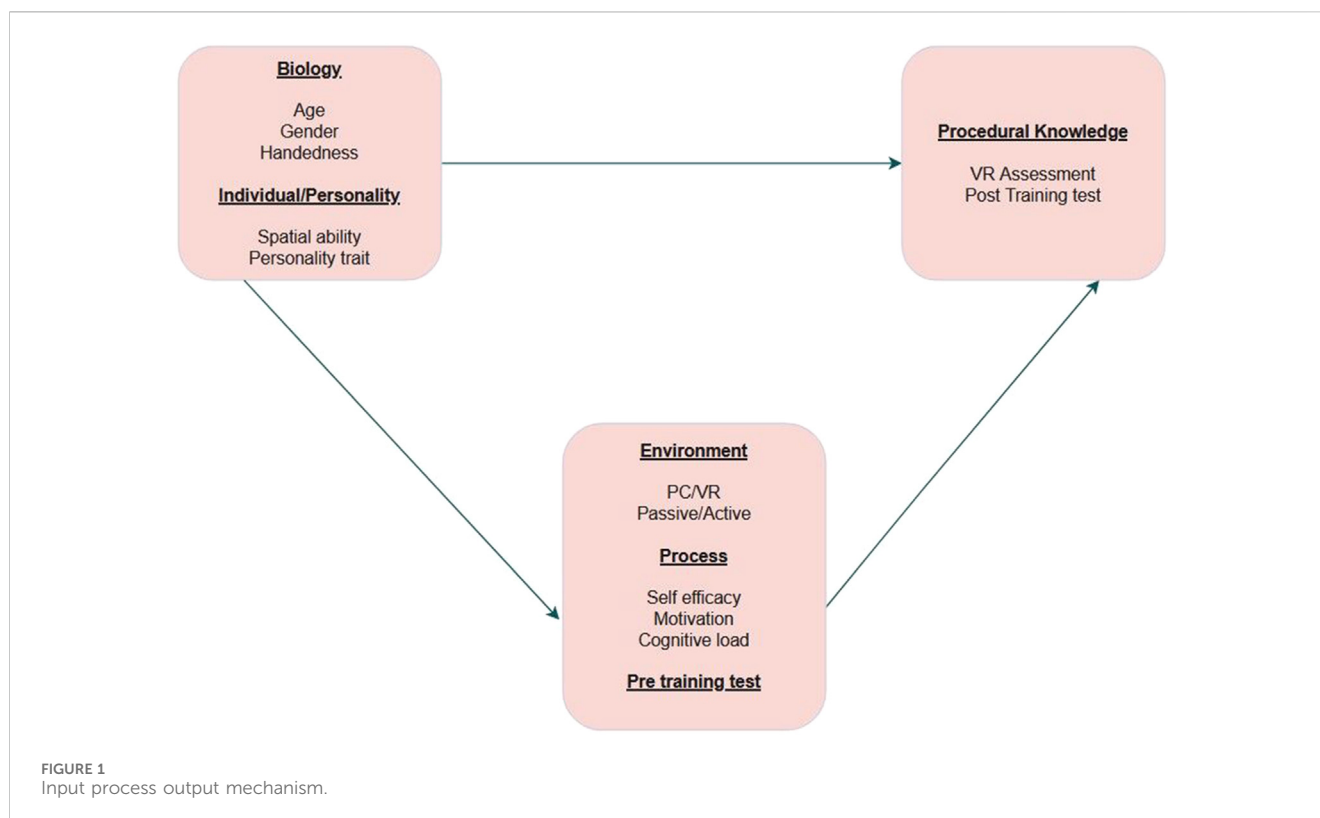
In VR learning environments, the level of immersion influences learning outcomes (Conrad et al., 2024). Two common display formats that differ in immersion, how the virtual content is presented, and the level of sensory engagement are desktop virtual reality and immersive virtual reality.

2.2.1 Desktop virtual reality (DVR) and immersive virtual reality (IVR)

VR systems can operate on relatively inexpensive platforms, such as desktops. This type of VR is referred to as desktop virtual reality (DVR), allowing users to interact within VR using a monitor, keyboard and mouse (Lee et al., 2010; Liu et al., 2024). Research has shown positive learning outcomes in DVR when compared to traditional classroom training (Lee et al., 2010; Makransky and Petersen, 2019). DVR also limits the inherent issue of simulation sickness associated with using head mounted displays (HMDs) (Stanney et al., 2020). DVR has been known to provide a low sense of presence for learners. However, the sense of presence in a virtual environment is not an inherent characteristic of the environment itself, rather, it is a psychological state induced by representational fidelity and user interactivity (Lee et al., 2010). The combination of these elements can create a compelling sense of presence, even without the use of specialized immersive hardware like the HMD.

IVR, on the other hand, operates using HMDs and hand controllers to provide a more immersive experience that blocks out the physical world during learning. This technology fosters a strong sense of presence and immersion in VR (Slater, 2009), allowing users to feel as though they were present in the simulated space. However, IVR has drawbacks, including the need for costly specialized hardware and the potential to cause motion sickness or discomfort for some users.

Some researchers argue that learning procedures in IVR is more effective than using a desktop PC. While some studies suggest that IVR enhances the recall of abstract concepts and improves procedural knowledge acquisition compared to non-immersive technologies, other studies have found no significant advantages. As pointed out



earlier, one reason for this is that learners tend to pay attention to the semantic details and not to the main learning content, however, DVR eliminates the extraneous cognitive load caused by the distractive elements of IVR. Another reason is that many of these learners lack prior IVR experience since technology and task familiarity also impact learning outcomes. A meta-analysis conducted by [Coban et al. \(2022\)](#) and [Wu et al. \(2020\)](#) revealed that IVR has a small advantage over non-immersive learning environments. It is essential to recognize here that even minor effects on learning can have considerable practical implications in real-world scenarios. IVR has demonstrated particular strengths in areas that require visual and spatial perception, making it especially effective for tasks such as safety and medical training ([Conrad et al., 2024](#); [Hamilton et al., 2021](#)). IVR also shows promise in the development of psychomotor and practical skills, potentially offering advantages over non-immersive technologies like desktop computers.

Two key technological factors that set VR apart from other training tools are immersion and interactivity ([Johnson-Glenberg, 2018](#); [Makransky and Petersen, 2021](#)). These elements play crucial roles in shaping the learning experience in VR. As we pointed out above, very few studies have investigated both immersion and interactivity within the same research (e.g., [Johnson et al., 2022](#)), and our research here extends this work by systematically manipulating both in a 2×2 design and also investigating the effects of three individual-differences variables.

2.2.2 Immersion and interactivity

Immersion can be described both objectively and subjectively. From the objective standpoint, it is viewed as the capacity of computer displays (such as the hardware and software) to create

a comprehensive and vivid illusion of reality that surrounds the user's senses ([Slater et al., 2022](#); [Slater and Wilbur, 1997](#)). System immersion is determined by features such as visuals, audio, haptics, tracking, and the alignment of multiple sensory inputs ([Slater and Wilbur, 1997](#)). It is also measured using metrics like stereoscopy, image quality, and the field of view or field of regard ([Bowman and McMahan, 2007](#); [Cummings and Bailenson, 2016](#)). Subjectively, immersion is described as the psychological state where users feel their senses are isolated from the real world. It refers to the degree to which learners feel present in the virtual environment, achieved by HMDs that provide a 360-degree view and isolate them from the physical world. It encompasses factors such as the degree of physical reality exclusion, the range of sensory inputs, the display resolution and accuracy ([Cummings and Bailenson, 2016](#); [Slater and Wilbur, 1997](#)). The attractive aspect of immersive virtual reality (IVR) for learning is that high immersion leads to a greater sense of presence ([Makransky and Petersen, 2021](#); [Petersen et al., 2022](#)), which in turn, is thought to boost the learner's motivation and engagement in the learning content. With this understanding, we suggest that there will be a main effect of immersion on procedural learning outcomes meaning that the level of immersion significantly affects how well individuals learn procedural tasks, on average, across all groups or conditions.

However, the same immersive quality that makes IVR attractive can also present itself as an obstacle that impedes effective learning. High immersion can divert the learner's attention from the core content, resulting in extraneous cognitive load - where learners focus on the non-essential elements and decrease essential processing which is necessary for understanding procedures and processes. The risk is that learners may become distracted by aesthetics causing

them to lose focus on the main lesson objective. Also, due to the limited capacity of the human working memory (Baddeley, 1992; Chandler and Sweller, 1991; Mayer, 2022), when a learner focuses on irrelevant content (extraneous information), there will be insufficient mental capacity for essential processing which is the cognitive processing required for meaningful procedural learning outcomes. Thus, we hypothesize that:

Hypothesis 1a: There will be a significant effect of immersion on procedural learning outcomes suggesting that those in high-immersive condition will demonstrate better procedural learning outcomes compared to those in low-immersive conditions.

Interactivity refers to the degree of engagement and reciprocal communication between learners and the VR environment. It relates to the user's ability to manipulate objects, typically using hand-held controllers or gesture-based commands. By allowing learners to manipulate virtual elements, make decisions, and receive instant feedback, interactivity contributes to more dynamic and participatory learning experiences. This active engagement may be associated with deeper understanding and improved knowledge retention. The interactive elements in VR may support generative processing by aiding in the selection, organization, and integration of information. These interactive elements may be more effective than immersion alone in supporting learning outcomes (Jang et al., 2017; Johnson et al., 2022). It is sometimes easy to assume that interactivity or active VR automatically means active learning, while the more traditional or passive VR means passive learning. However, this view can be misleading. The key point here is that true active learning is not determined by outward behavior but rather by the level of cognitive engagement (Mayer, 2021; Mayer, 2022). According to the cognitive theory of multimedia learning (CTML, Mayer, 2022; Mayer et al., 2023), it is the depth of mental processing specifically essential processing (understanding the steps and processes during procedural learning) and generative processing (integrating new information with existing information), that define active learning. This means that a seemingly passive activity like training in a desktop environment or using an iPad can be considered active learning if it stimulates cognitive engagement. On the other hand, an outwardly interactive experience like a poorly designed immersive learning environment might result in superficial engagement but not meaningful learning. With this understanding, we hypothesized that:

Hypothesis 1b: There will be a significant effect of interactivity on procedural learning outcomes suggesting that those in the high interactive (active) condition will demonstrate better procedural learning outcomes compared to those in the low-interactive (passive) condition.

Research on the combined effects of immersion and interactivity on VR learning has yielded varied outcomes. Some studies have found that higher levels of immersion lead to increased feelings of presence and agency which leads to improved declarative learning outcomes and embodied learning (Petersen et al., 2022). Others, like Johnson et al. (2022), found no significant interaction between immersion and interactivity when learning procedures in VR. Buttussi and Chittaro (2017) reported that using an HMD resulted in a greater sense of presence and engagement and not safety knowledge when compared to learning the same content on a desktop PC. Cummings and Bailenson (2016) also conducted a

meta-analysis that revealed a strong relationship between immersion and participant-reported presence. Various theories like the CTML (Mayer, 2024; Parong and Mayer, 2018) have considered how the technological capabilities of VR and the human cognitive processing capacity influence learning. The fundamental principles of CTML have been applied to understand how immersion and interactivity may drive meaningful learning in VR. The first principle is dual channels, developed from the dual coding theory (Paivio, 1990) which means that humans process verbal and visual information through different pathways, hence, when the learning content is presented in multiple modes, learners benefit more as long as the channels are complementary. The second principle of limited capacity recognizes that our ability to process information is constrained by the finite capacity of our working memory. This principle is particularly important to manage cognitive load. The third principle is active processing which emphasizes that meaningful learning is not a passive reception of information, but rather an active cognitive process. As learners interact in the VR environment, it increases generative processing which may be essential for enhancing their understanding and recall of procedural tasks. Integrating the above, we suggest that there will be an interaction effect between immersion and interactivity meaning that the effect of immersion on procedural learning outcomes may differ depending on whether interactivity is low or high. Thus, we hypothesized that:

Hypothesis 1c: There will be an interaction effect between immersion and interactivity such that those in the high immersion (VR) group and high interactivity (active) group will have disproportionately positive procedural learning outcomes.

The effectiveness of immersion and interactivity in VR learning may be influenced by individual differences among learners. Spatial ability, for example, plays an important role in how individuals interact with and benefit from 3D virtual environments. Sex is also associated with differences in spatial cognition, technology acceptance and the occurrence of simulation sickness, which could influence VR learning outcomes. Additionally, personality traits such as openness to experience, conscientiousness and agreeableness may affect how individuals engage with and learn from VR.

2.3 Individual differences factors

For this paper, we adopted the viewpoint that the effectiveness of immersive technology is not solely dependent on the media used, but rather on the alignment between the learning environment and the individual characteristics of learners. Individual differences play an important role in learning procedures and can account for learning outcomes in VR. In this study, we refer to individual differences as the differences in sex, spatial ability, and personality traits among learners engaging with VR for procedural training.

2.3.1 Sex

Research has shown that sex can influence spatial cognition and technology acceptance, which may, in turn, affect VR learning outcomes. Generally, males have been found to outperform

females on certain spatial tasks, such as mental rotation (Sneider, 2015; Wei, 2016), for example, because males have more interest and experience with spatial tasks, such as interactive spatial video games. Sex has also been identified as a significant factor influencing experiences and outcomes in VR, particularly concerning the phenomenon known as simulator sickness.

Research has indicated that women are generally more susceptible to simulator sickness than men, in part due to the design of HMDs (i.e., sub-optimal inter-pupillary distance, IPD; Munafò et al., 2017), but also due to lower chronic adaptation (Grassini and Laumann, 2020; MacArthur et al., 2021). Thus, we hypothesized that:

Hypothesis 2: Sex will significantly predict procedural learning outcomes, with males demonstrating higher procedural learning outcomes compared to females because of more experience and interest in technology particularly video games and differences in visuo spatial working memory.

2.3.2 Spatial ability

As briefly mentioned with respect to the difference between males and females above, sex-based differences in spatial ability have been identified as a critical factor in how individuals interact with and learn from 3D virtual environments. It refers to the cognitive capacity to mentally manipulate spatial information and it plays a vital role in how users interact with and navigate virtual environments. However, even after accounting for sex differences, spatial ability is an important predictor of performance in spatial tasks. According to the ability as compensator hypothesis, low spatial ability learners benefit more from VR environments because it helps them build their mental representation whereas high spatial ability learners can create mental visualizations by themselves (Huk, 2006; Mayer 2001). In IVR, spatial ability is crucial for tasks like navigating complex virtual worlds, understanding spatial relationships, and manipulating objects within the 3D space. Individual differences in spatial ability can significantly impact performance and learning outcomes in VR (Johnson et al., 2022). In the context of procedural learning, spatial ability may be relevant for tasks that involve understanding spatial relationships between objects or components. Many VR environments are not designed specifically to train spatial navigation skills, but their spatial layout and navigational demands can significantly affect procedural learning outcomes, as efficient navigation is often necessary for engaging with the content. With this understanding, we hypothesized that:

Hypothesis 3: Spatial ability will significantly predict procedural learning outcomes, with higher spatial ability associated with better performance on procedural tasks because high spatial ability enables individuals to process spatial information more efficiently, reducing cognitive load and freeing up mental resources for task performance in VR.

2.3.3 Personality traits

Personality traits include the enduring patterns of thoughts, feelings, and behaviors that define an individual (Thorp et al., 2023). In VR, traits like openness to experience, curiosity, risk-taking, and

preference for autonomy can influence engagement, exploration, and learning outcomes.

Few studies have investigated the impact of individual personality traits on VR learning, and none related to the effect of personality on learning procedures in VR. Among those that have been conducted, only a limited number have identified significant relationships. For instance, research examining the relationship between surgeons' technical performance and personality traits in VR settings found no significant correlation between these traits and technical performance outcomes (Rosenthal et al., 2013). Other studies have indicated that certain personality traits may correlate with better academic performance. For example, Thorp et al. (2023) found that high agreeableness and low conscientiousness significantly predict knowledge transfer in VR. Komarraju et al. (2011) also found positive relationships between conscientiousness and agreeableness on academic performance. In the current study, personality trait was measured using the five-factor model which conceptualizes personality through five dimensions: extraversion, agreeableness, conscientiousness, emotional stability, and openness. We hypothesized that:

Hypothesis 4: Personality traits will significantly predict procedural learning outcomes, with high conscientiousness, high agreeableness and high openness positively associated with better procedural learning outcomes.

2.4 Research hypotheses

In summary, we stated the following hypotheses for this study on the basis of our theoretical framework and review of the literature:

H1a: There will be significant differences in procedural learning outcomes based on the level of immersion. This means that participants in high-immersion conditions (VR) will demonstrate better procedural learning outcomes compared to those in low-immersion conditions (PC) regardless of the level of interactivity.

H1b: There will be significant differences in procedural learning outcomes based on the level of interactivity. This suggests that participants in high-interactivity conditions (active) will demonstrate better procedural learning outcomes compared to those in low-interactivity conditions (passive) regardless of the level of immersion.

H1c: There will be an interaction effect between immersion and interactivity on procedural learning outcomes. This means that the effect of immersion on procedural learning outcomes will depend on the level of interactivity. Specifically, participants in the high-immersion and high-interactivity condition (VR active group) will demonstrate the best procedural learning outcomes compared to all other conditions (high-immersion/low-interactivity, low-immersion/high-interactivity, and low-immersion/low-interactivity).

H2: Sex will significantly predict procedural learning outcomes, with males demonstrating higher procedural learning outcomes

compared to females because of more experience and interest in technology particularly video games and differences in visuo spatial working memory.

H3: Spatial ability will significantly predict procedural learning outcomes, with higher spatial ability associated with better performance on procedural tasks because high spatial ability enables individuals to process spatial information more efficiently, reducing cognitive load and freeing up mental resources for task performance in VR.

H4: Personality traits will significantly predict procedural learning outcomes, with high conscientiousness, high agreeableness and high openness positively associated with better procedural learning outcomes. This is because learning procedures in VR often involve exploratory, adaptive, and trial-and-error methods and people with high conscientiousness exhibit a strong preference for structure, orderliness, and goal-directed behavior. Openness to experience is linked to curiosity, imagination, and willingness to explore new concepts.

Individuals high in openness may be more engaged and receptive to visual-spatial cues, and problem-solving tasks, leading to better procedural learning outcomes.

Note: We measured procedural learning outcomes using a practical assessment in the simulation using desktop PC or immersive VR and post-knowledge test scores. The reason for this was to avoid the “getting good at the game effect” (Parong and Mayer, 2018) where learners can perform procedures virtually but not transfer to the real world, which is what is considered procedural knowledge.

3 Methods

3.1 Power analysis

We conducted a sample size estimation using G*Power software to find the optimal number of participants (Faul et al., 2007). Earlier studies have found bigger effect sizes, so we chose to adopt a similar approach by choosing a large effect size of $d = 0.40$, which is supported by existing literature (Mayer, 2017). By setting the significance level at 0.05 and the desired power at 0.80, we determined that a total sample size of 64 participants would be necessary.

3.2 Design

We utilized both 2 (Immersion: PC vs. VR) $\times$ 2 (Interactivity: Passive vs. Active) between-subjects analysis of covariance (ANCOVA) and stepwise multiple regression analyses. For the ANCOVAs, the independent variables were immersion and interactivity, with pre-knowledge test scores as the covariate. The dependent variable was procedural learning outcomes measured using a practical assessment and post-knowledge test scores. The scoring range for the procedural learning outcomes was theoretically 0 to 15 points. Sex, spatial ability and personality traits were the predictor variables in the Multiple Regression/Correlation Analyses.

TABLE 2 Participant distribution by condition.

Condition	Immersion	Interactivity	<i>n</i>
PC - Passive	Low	Low	18
PC - Active	Low	High	19
VR - Passive	High	Low	22
VR - Active	High	High	20

Note. *N* = 79.

We employed a stratified random assignment method, ensuring that students were randomly assigned to one of four groups (PC passive, PC active, VR passive, VR active), each varying in levels of immersion and interactivity, and stratified for participant sex.

3.3 Participants

A total of 83 undergraduate students recruited from a large public university in the Southeastern United States signed up to participate in the study. However, four participants were excluded due to technological issues, simulation sickness or lack of interest in completion. This resulted in a final sample size of 79 students. All participants received course credits for research participation. The sample included 69 participants between the ages of 18–20 and 10 between the ages of 21–30. There were 43 females and 36 males. Forty-two participants were randomly assigned to the VR group and thirty-seven were assigned to the PC group. None of the participants had prior experience with aviation or with using desktop virtual reality (DVR) or immersive virtual reality (IVR) for learning instructional content. Table 2 has details of participants and experiment conditions.

3.4 Apparatus

3.4.1 Simulation testbed: flightcrew procedures experimental training (FlightPET)

The simulation testbed for the experiment was a virtual reality simulation of an exterior preflight inspection for a generic commercial aircraft developed at the Institute for Simulation & Training of the University of Central Florida (Sonnenfeld et al., 2023). This system was developed using the Unity 3D gaming engine (Figures 2, 3 has images of the simulation testbed). The simulation allowed participants to interact in VR under two conditions: a 2D environment using a desktop PC, keyboard and mouse, and a 3D environment using an HTC Vive HMD and controllers (Figure 4 provides an image of the HMD).

The training included a radial tool that enabled students to select various instructional components for each aircraft section (Figure 5 has an image of the instructional components), such as text, audio, point-to-object interactions, state comparisons, animations, and a question-and-answer feature. During the practical assessment, participants utilized a laser pointer—controlled via a mouse in DVR or hand controllers in IVR—to identify and select flaws in aircraft components, following the guidance provided in the training.

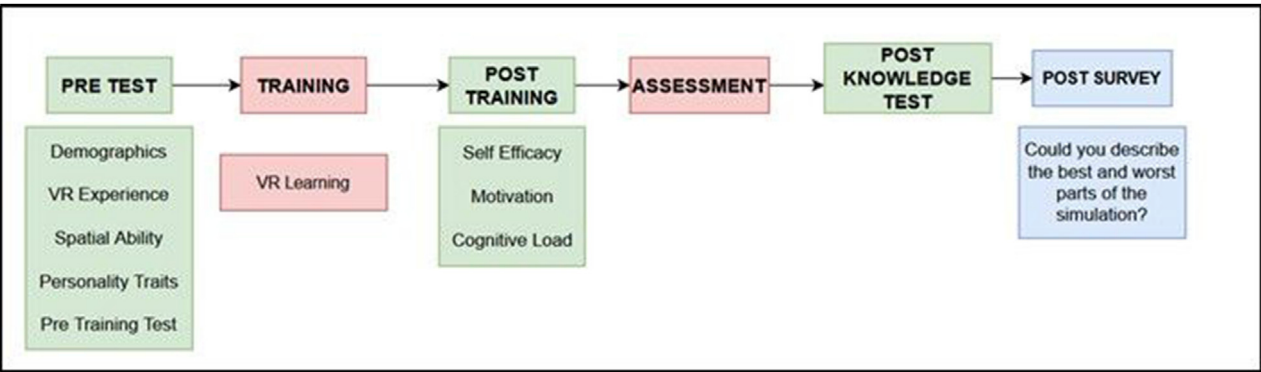


FIGURE 2
Experimental research design.

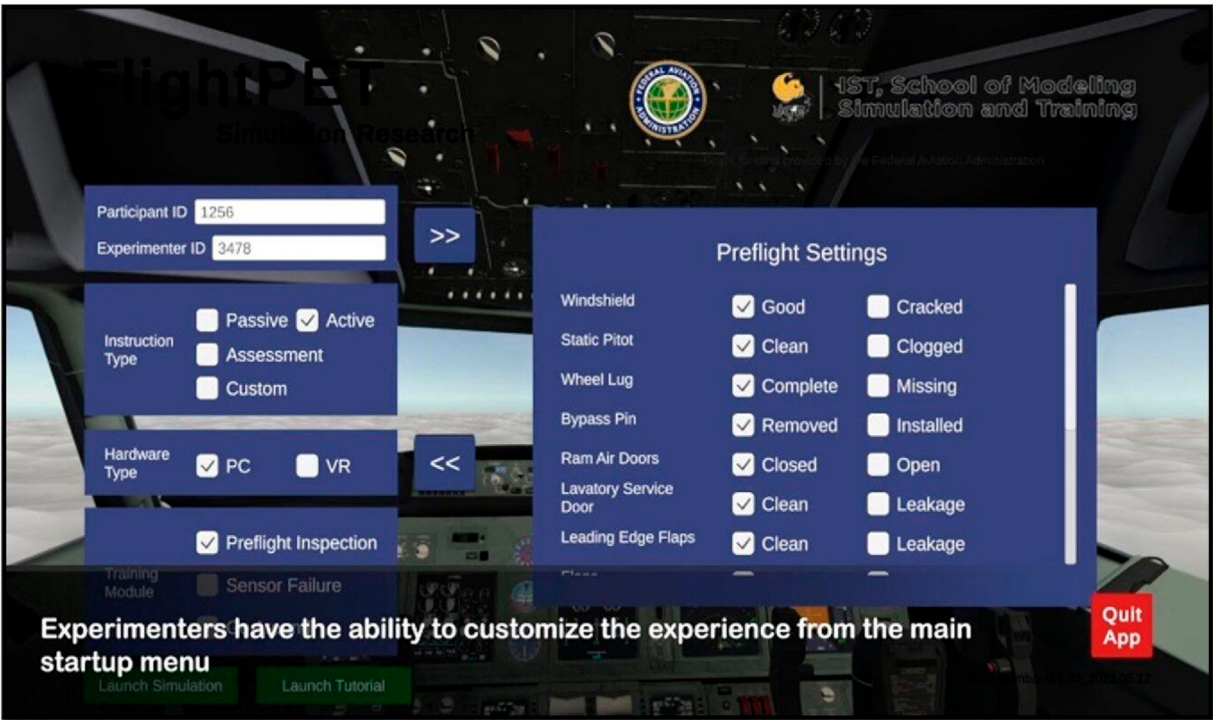


FIGURE 3
The flightcrew procedures experimental training (FlightPET) simulation testbed.

3.5 Materials

3.5.1 Pre-test

3.5.1.1 Demographic survey

Participants provided general demographic information. They were asked about their age (in age groups of 18–21 and 21–30), sex, handedness (right-handed, left-handed, and ambidextrous), and any vision correction methods (none, contact lenses, glasses). Additional questions pertained to aviation experience, such as employment as a pilot and holding various pilot certificates.

Participants were asked about their prior experience with VR procedure trainers, including whether they had used such trainers before and a description of their experience. They rated their skill levels in using various VR interfaces, such as a mouse and keyboard, touchscreen devices, handheld controllers, and motion-tracking controllers, using a 5-point self-assessment scale with the following levels: Novice (no prior experience), proficient (basic familiarity with occasional use), competent (comfortable with independent use), advanced (frequent use with a high level of efficiency), and expert (extensive experience, including the ability



FIGURE 4
Htc Vive 3 headset.

to troubleshoot and instruct others). Additionally, participants reported the frequency of experiencing symptoms like nausea, headache, dizziness, fatigue, and eye strain while using these devices.

3.5.1.2 Paper folding test

Spatial ability was measured using the Paper Folding Test (PFT; Ekstrom and Harman, 1976). The PFT consists of two sets of ten items that participants completed within 3 min for each section. Each item presented images illustrating a piece of paper being folded with a hole punched through it. Participants were required to select one of five images that accurately depicted how the paper would appear when unfolded. Scoring involved counting the number of correct answers. The paper folding task was chosen because it requires both mental rotation and maintenance of mental alterations in working memory.

3.5.1.3 Ten-item personality inventory (TIPI)

The ten-item personality inventory (TIPI; Gosling et al., 2003) was used to measure personality traits. Participants rated their

agreement with ten statements on a 7-point scale ranging from 1 (strongly) to 7 (strongly agree). The TIPI assesses the big five personality domains: extraversion, agreeableness, conscientiousness, emotional stability, and openness.

3.5.1.4 Practical assessment task

In the practical assessment task, participants performed an exterior preflight inspection based on the procedures learned during training. Using either the desktop PC or immersive VR setup, they walked around the aircraft to identify any flaws, faults, or missing or damaged components. Each participant was given 10 min to complete the task. During the simulation, participants remained seated to ensure standardization and safety. In the VR condition, interaction was facilitated through hand-held controllers and mouse with a visible laser pointer. No full-body avatar or locomotion was implemented. The environment was visually realistic but lacked ambient audio, environmental changes, or haptic feedback.

3.5.1.5 Pre and post-knowledge questions

Participants completed pre and post-knowledge questions on exterior preflight inspection procedures (Nguyen et al., 2023). Pre and post-knowledge questions were similar in content and structure. The content validity of both tests was determined by expert judgment. The evaluation included a variety of question formats, including three multiple-choice, three fill-in-the-blank, and three scenario-based questions. It covered key aspects of the training on preflight inspections, such as its purpose, the recommended sequence for inspecting an aircraft's exterior, and specific components within each inspection criteria. The tests measured learning outcomes cognitively based on the number of questions answered correctly. The scoring range for the pre and post-knowledge test was theoretically 0 to 15 points (see Supplementary Appendix A for a complete set of questions).

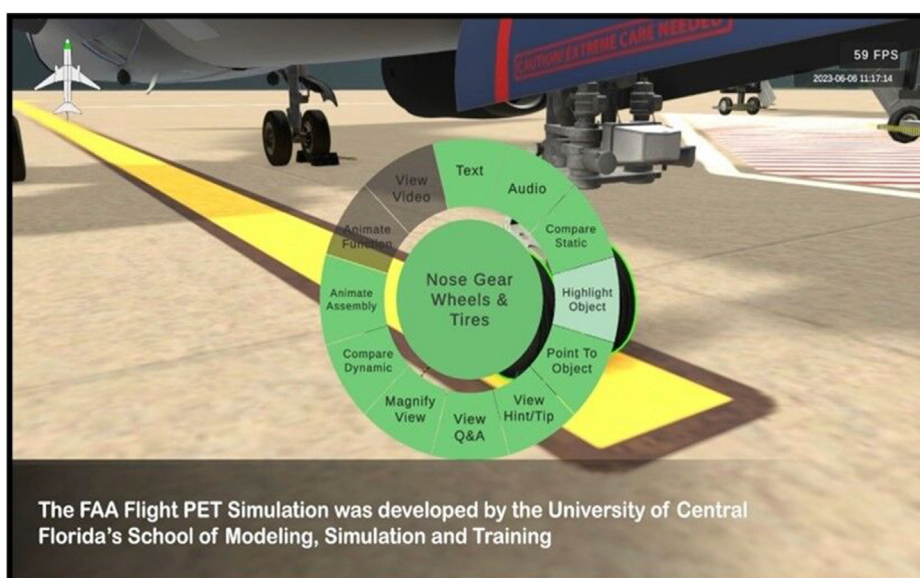


FIGURE 5
Radial tool for learning component selection.

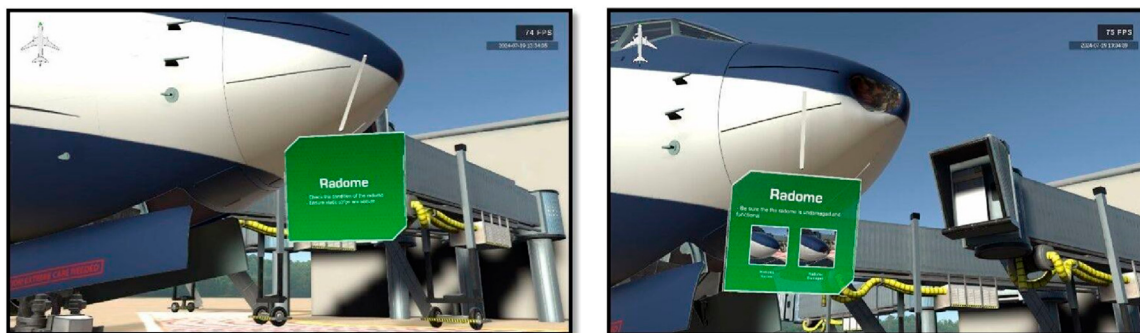


FIGURE 6
Screenshots of the virtual learning environment: FlightPET. Note. Training scenario: The left picture (A) is a good version of the aircraft radome with instructions for the participants to check the radome and ensure that the static strips are secure. The right picture (B) is a full-blown image of a damaged version of the radome. The instructional picture shows a picture of a good and damaged radome, with instructions for learners to ensure that the radome is undamaged and functional.

3.5.2 Post-training questionnaire

After exposure to the simulation training in desktop PC or immersive VR, data was obtained on participants' perceived efficacy, motivation, and cognitive load. Participants responded to a series of standardized questionnaires designed to measure their confidence in learning, their motivation levels, and the mental effort they applied during the training phase.

3.5.2.1 Self-efficacy

Self-efficacy was measured using items adapted from Meyer et al. (2019) using self-ratings on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Participants responded to statements such as "I am confident I have the ability to learn the material taught about exterior preflight inspection" and "I believe that if I exert enough effort, I will be successful on the assessment about exterior preflight inspection." Responses were used to assess participants' confidence in their ability to understand and perform tasks related to the exterior preflight inspection.

3.5.2.2 Motivation

Motivation was assessed using items adapted from Lee et al. (2010) and Makransky and Petersen (2019) using self-ratings on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Participants rated their agreement with statements like "The VR application could enhance my learning interest" and "The realism of the VR application motivated me to learn." These items were designed to gauge how effectively the VR training environment stimulated participants' interest in engaging with the material.

3.5.2.3 Cognitive load

Cognitive load was measured using items adapted from Andersen and Makransky (2021) using self-ratings on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Participants indicated the amount of mental effort they spent learning about preflight inspection. Statements included "The interaction technique used in the simulation was very unclear" and "It was difficult to find the relevant learning information in the virtual environment." These measures helped

evaluate the mental effort required to learn and perform tasks within the VR training environment.

3.5.3 Debrief questions

The debrief questions obtained participant's qualitative feedback on the training they received. It included questions asking participants to describe the best and worst parts of the training experience and to suggest improvements. Participants were asked to evaluate the training's effectiveness, providing insights into how well the training met their learning needs and expectations. This feedback was used to assess the overall satisfaction and perceived value of the VR training.

3.6 Procedure

The study was approved by the institutional review board at the university. We used an experimental design for this study (Figure 6 provides an overview of the research procedure). Upon arrival at the research laboratory, participants were guided through the initial steps of the study. First, they reviewed the informed consent document, which detailed the study's purpose, procedures, potential risks, and their rights as research participants. After ensuring they understood the information provided, they gave their verbal consent to participate in the research. Once consent was obtained, participants proceeded to complete a demographics and individual differences survey, which took approximately 30 min. This survey was designed to gather data on various factors that could influence learning outcomes, including demographics (such as age, sex and handedness), prior VR experience, spatial ability, personality traits, and existing knowledge about exterior pre-flight inspection procedures using pre-knowledge test questions. The insights gained from this survey were essential for understanding the baseline characteristics of each participant and for analyzing how these factors might interact with the instructional interventions.

After completing the survey on demographics, spatial ability and personality traits, participants were randomly assigned to one of four experimental conditions: PC Passive, PC Active, VR Passive, or VR Active and stratified by sex. Each participant then engaged in a 5-minute simulation tutorial using a desktop PC or immersive VR

TABLE 3 Demographic characteristics of participants.

Constructs	<i>n</i>	%	<i>M</i>	<i>SD</i>
Age			0.13	0.34
18–21	69	87.3		
21–30	10	12.7		
Sex			0.46	0.501
Female	43	54.4		
Male	36	45.6		
Handedness			0.09	0.29
Right-handed	72	91.1		
Left-handed	7	8.9		
Immersion			0.53	0.502
PC	37	46.8		
VR	42	53.2		
Interactivity			0.49	0.503
Passive	40	50.6		
Active	39	49.4		
Personality Traits				
Extraversion			5.04	1.98
Agreeableness			5.19	2.2
Conscientiousness			4.74	2.13
Emotional Stability			5.77	2.02
Openness			4.59	1.801
Affective Factors				
Self-Efficacy			15.97	4.28
Motivation			13.28	4.48
Cognitive Load			26.29	7.65
Test Scores				
Pre-training Test			6.73	2.06
VR Assessment			7.13	2.29
Post training Test			7.51	1.83

Note. N = 79.

designed to familiarize them with the simulation controls. The tutorial was essential for ensuring that all participants had a basic understanding of how to navigate within the VR learning environment.

After the tutorial, participants completed a 20-minute simulation training session using the desktop PC or immersive VR that provided detailed procedural instructions for conducting an exterior pre-flight inspection. An exterior preflight inspection is a crucial safety procedure conducted by pilots before each flight, involving a walk-around of the aircraft to ensure its airworthiness. The inspection involves the pilot walking around the aircraft in a clockwise manner to examine critical components like the wings, landing gears, and engines for damage, leaks and obstructions. The training was structured to equip participants with both theoretical knowledge and practical skills necessary for performing the task effectively.

After the simulation training, participants filled out a questionnaire that assessed various psychological factors such as self-efficacy (their confidence in performing tasks), motivation (their enthusiasm for learning), and cognitive load (the mental effort required during training). The questionnaire provided valuable insights into how participants felt about their

TABLE 4 Regression analysis on procedural learning outcomes (post-training test).

Variable	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	95% CI
Model 1					
Constant	5.21	0.66	7.92	<0.001	[3.9, 6.5]
Pretraining test	0.34	0.09	3.66	<0.001	[0.16, 0.53]
Model 2					
Constant	3.52	0.90	3.90	<0.001	[1.72, 5.32]
Pretraining test	0.35	0.09	3.82	<0.001	[0.17, 0.52]
Cognitive Load	0.06	0.02	2.63	0.010	[0.02, 0.11]
Model 3					
Constant	2.24	1.08	2.08	0.04	[0.09, 4.38]
Pretraining test	0.343	0.09	3.89	<0.001	[0.17, 0.52]
Cognitive Load	0.07	0.02	2.93	0.004	[0.02, 0.12]
Spatial Ability	0.11	0.05	2.08	0.04	[0.004, 0.21]

Note. CI, confidence interval; SE, standard error; N = 79.

learning experience and their readiness to apply what they had learned.

Next, participants completed a practical assessment using either the desktop PC or HMD VR that lasted exactly 10 min. During this assessment, participants were tasked with identifying faults or damages on various aircraft components in a clockwise manner, consistent with the procedure used during the training phase. Participants indicated the identified faults by pointing at them using a laser pointer controlled via either a mouse (in the PC condition) or hand controllers (in the VR condition). A researcher directly observed the participants during the task, monitoring their actions in real time. The researcher manually recorded the participants' responses, and this data was stored for later review and analysis. This practical evaluation was scored on accuracy and adherence to inspection procedures, allowing the researcher to gauge each participant's ability to apply their skills in a realistic context. Finally, after completing the practical assessment, participants took a post-knowledge test designed to measure their performance outcomes in terms of knowledge retention and application. This test included various question formats such as multiple-choice, fill-in-the-blank and scenario-based questions that assessed their understanding of pre-flight procedures. At the end of the experiment, qualitative feedback (debrief) was obtained to gather participants' opinions about the training experience, their perceptions of VR, and their views on the effectiveness of instructorless training methods.

4 Results

4.1 Data pre-screening

Data from four participants were excluded from the analyses for the following reasons: one participant encountered technical difficulties, another experienced severe simulation sickness, and two withdrew due to a lack of interest in continuing the study. These exclusion criteria were applied to ensure that external factors, such as technical issues or discomfort from simulation sickness did not compromise the

TABLE 5 Descriptives and correlations between technology factors and affective factors.

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
1. Interactivity	0.49	0.50	-							
2. Immersion	0.53	0.50	-0.04	-						
3. Self -Efficacy	15.97	4.28	-0.06	-0.02	-					
4. Motivation	13.28	4.48	-0.095	0.15	0.40**	-				
5. Cognitive Load	26.29	7.65	-0.12	0.07	0.31**	0.37**	-			
6. Pre-training	6.73	2.06	0.09	-0.11	-0.08	-0.02	-0.009	-		
7. VR Assessment	7.13	2.29	-0.06	0.197	-0.03	0.19	0.12	0.10	-	
8. Post-training	7.51	1.83	-0.14	-0.08	-0.007	0.07	0.26*	0.39**	0.20	-

* $p < .05$. ** $p < .01$. $N = 79$.

accuracy and reliability of the data. As a result, the subsequent analyses include data from 79 participants.

4.2 Test scoring criteria

Scoring for the pre-knowledge test, practical assessment, and post-knowledge test was done on a 15-point scale. For the pre-knowledge test and post-knowledge test, multiple-choice and fill-in-the-blank questions were scored dichotomously, with correct responses receiving one point and incorrect responses receiving zero. Scenario-based questions were evaluated using a rubric that specified required words and phrases, awarding one point for a fully accurate response and 0.5 points for a partially correct answer. For the practical assessment, participants received one point for correctly identifying a flaw in an aircraft component and zero points for incorrect or missed identifications.

4.3 Descriptive statistics

A total of 79 students participated in the study (Table 3 provides details on demographics), and they were randomly assigned to one of four experimental groups, which varied based on immersion level (PC or VR) and interactivity level (Passive or Active) and stratified by participant sex. The sample predominantly consisted of individuals aged 18 to 21 ($n = 69$), with a sex distribution of 43 females and 36 males, and participants were mostly right-handed ($n = 72$).

Means, SDs, and bivariate correlations of the variables of interest can be found in Table 4. The analysis focused on the relationships between the technology factors (immersion and interactivity), performance measures (pre-knowledge test scores, practical assessment scores, and post-knowledge test scores), and affective mechanisms (cognitive load, motivation, and self-efficacy). Self-efficacy was found to have a positive correlation with both motivation, $r(77) = 0.403$, $p < 0.001$ and, cognitive load, $r(77) = 0.311$, $p = 0.005$. Motivation was positively correlated with cognitive load, $r(77) = 0.373$, $p < 0.001$. Cognitive load also demonstrated a significant positive correlation with post-knowledge test scores, $r(77) = 0.435$, $p < 0.001$ while pre-knowledge test scores were positively correlated with post-knowledge test scores, $r(77) = 0.266$, $p = 0.018$.

4.4 Equivalence assessment across immersion and interactivity

A 2 (Immersion: PC vs VR) x 2 (Interactivity: Passive vs Active) analysis of variance (ANOVA) was conducted to verify that the random assignment of participants between the conditions of immersion and interactivity worked. Across all variables of interest including spatial ability, personality traits and pre-knowledge test scores, there were no significant differences (Table 5 provides details on the equivalence test). For sex, there were roughly equal men and women in each group. For immersion, female participants had a distribution of 21 in the PC condition and 22 in the VR condition, while male participants had 16 in the PC condition and 20 in the VR condition. Regarding interactivity, female participants were distributed as 21 in the passive condition and 22 in the active condition, whereas male participants had 19 in the passive condition and 17 in the active condition. Hence, we concluded that random assignment worked based on sex, pre-knowledge test, spatial ability and personality traits.

4.5 Effect of immersion and interactivity on procedural learning outcomes

We employed a 2 (Immersion: PC vs VR) x 2 (Interactivity: Passive vs Active) between-subjects analysis of covariance (ANCOVA) to investigate the main effects and interaction of immersion and interactivity on procedural learning outcomes (measured using the practical assessment and post-knowledge test scores) while controlling for pre-knowledge test scores. The scoring range for all tests (pre-knowledge, practical assessment, and post-knowledge) was theoretically 0 to 15 points. However, in practice, participant's actual scores ranged from 2 to 11.5 for pre-knowledge, 2 to 12 for the practical assessment scores and 1 to 11 for the post-knowledge scores.

Before conducting the analysis, we evaluated the normality of the model residuals using Q-Q plots and the assumptions of normality were met.

For the practical assessment scores, the means for the levels of immersion were PC ($M = 6.62$) and VR ($M = 7.58$). The main effect of immersion was not statistically significant albeit close, $F(1, 74)$

$=3.41$, $p = .07$, $\eta p2 = .04$, as measured by practical assessment scores between students in desktop PC and VR conditions. Furthermore, the means for the levels of interactivity were—passive ($M = 7.25$) and active ($M = 6.96$). The main effect of interactivity was also not significant $F(1, 74) = 0.31$, $p = 0.58$, $\eta p2 = 0.004$ between the active and passive conditions.

Additionally, the interaction between immersion and interactivity (Figure 7 details the ANCOVA results visually) was not statistically significant $F(1, 74) = 0.44$, $p = 0.51$, $\eta p2 = 0.006$. Overall, these results indicate that neither immersion nor interactivity, independently or in combination, had a significant effect on procedural learning outcomes as measured by the practical assessment scores. The small effect sizes further support the lack of substantial impact of these factors on the learning outcomes.

In the post-knowledge test, the means for the levels of immersion were—PC ($M = 7.36$) and VR ($M = 7.15$). The main effect of immersion was not statistically significant, $F(1, 74) = 0.12$, $p = 0.74$, $\eta p2 = 0.002$ between students in desktop and VR conditions, as measured by the post-knowledge test scores. Similarly, the means for the levels of interactivity were—passive ($M = 7.39$) and active ($M = 7.12$), the main effect of interactivity was not significant, $F(1, 74) = 0.07$, $p = 0.79$, $\eta p2 = 0.001$. and statistical analysis indicated no significant difference in scores between the active and passive conditions. The interaction between immersion and interactivity (Figure 8 details the ANCOVA results visually) was also not statistically significant, $F(1, 74) = 1.10$, $p = 0.299$, $\eta p2 = 0.02$. The results suggest that immersion and interactivity, both independently and in combination, do not significantly impact procedural learning outcomes in this context.

4.6 Effect of sex, spatial ability and personality trait on procedural learning outcomes measured with practical assessment scores

We conducted a stepwise multiple regression analysis to examine the predictors of procedural learning outcomes (measured with the practical assessment scores). We measured individual sex differences (coded: 0 = female, 1 = male), spatial ability, and personality traits. Analysis was performed using IBM SPSS regression and explore for evaluation of assumptions. Assumptions for normality, linearity, and homoscedasticity of residuals were met. The criteria for variable entry and removal were set at a probability of F to enter $p < 0.05$ and a probability of F to remove $p > 0.10$. The final model retained sex as a significant predictor of VR performance, $F(1, 77) = 17.42$, $p < 0.001$, explaining 18.5% of the variance in practical assessment scores ($R^2 = 0.185$, Adjusted $R^2 = 0.174$). The unstandardized regression coefficient ($B = 1.96$, $SE = 2.08$) indicated that males scored, on average, 1.96 points higher than females on the practical assessment (95% confidence interval [1.03, 2.90]). The standardized regression coefficient ($\beta = 0.43$) suggested that sex had a moderate positive effect on practical assessment scores, relative to other variables.

Spatial ability and personality traits were excluded from the final model because they did not significantly contribute to the prediction of practical assessment scores ($p > 0.10$).

4.7 Effect of sex, spatial ability and personality trait on procedural learning outcomes measured with post-knowledge test scores

A stepwise multiple regression analysis was performed with procedural learning outcomes as the outcome variable measured with post-knowledge test scores and sex, spatial ability and personality traits as the predictor variables (Table 6 details the multiple regression results).

Analysis was performed using IBM SPSS regression and explore for evaluation of assumptions. Assumptions for normality, linearity, and homoscedasticity of residuals were met. The criteria for variable entry and removal were set at a probability of F to enter $p < 0.05$ and a probability of F to remove $p > 0.10$. In the first model, pre-knowledge test scores were entered as the sole predictor, explaining 14.8% of the variance in post-knowledge test scores ($R^2 = 0.148$, Adjusted $R^2 = 0.137$), $F(1, 77) = 13.41$, $p < 0.001$. The unstandardized coefficient ($B = 0.343$, $SE = 0.094$, $t = 3.66$, $p < 0.001$) and standardized coefficient ($\beta = 0.385$) indicated that higher pre-knowledge test scores were significantly associated with better post-training performance. In the second model, cognitive load was added, which contributed an additional 7.1% of the variance ($\Delta R^2 = 0.071$, $F(1, 76) = 6.91$, $p = 0.010$). Together, pre-knowledge test scores and cognitive load explained 21.9% of the variance ($R^2 = 0.219$, Adjusted $R^2 = 0.199$). The unstandardized coefficient for cognitive load ($B = 0.064$, $SE = 0.024$, $t = 2.63$, $p = 0.010$) and standardized coefficient ($\beta = 0.266$) suggest that greater cognitive load during training is positively associated with higher post-training performance. In the third model, spatial ability was added as a significant predictor, contributing an additional 4.3% of the variance ($\Delta R^2 = 0.043$, $F(1, 75) = 4.32$, $p = 0.041$). The final model explained 26.2% of the variance in post-knowledge test scores ($R^2 = 0.262$, Adjusted $R^2 = 0.232$). Spatial ability ($B = 0.11$, $SE = 0.05$, $t = 2.08$, $p = 0.04$) was a significant predictor, indicating that higher spatial ability was associated with better post-knowledge test scores.

4.8 Exploratory analysis

We conducted an exploratory analysis to examine the correlations between variables in our study. The analysis focused on the relationships between the technology factors (immersion and interactivity), performance measures (pre-knowledge test scores, practical assessment scores, and post-knowledge test scores), and affective mechanisms (cognitive load, motivation, and self-efficacy). Bonferroni correction was not applied in the correlation analysis because it can be overly conservative in exploratory research, potentially masking meaningful relationships (Streiner and Norman, 2011). The correlation analysis (see Table 5 details the correlation analysis) identified several significant relationships between the variables studied. Self-efficacy was found to have a positive correlation with both motivation, $r(77) = 0.403$, $p < 0.001$ and, cognitive load, $r(77) = 0.311$, $p = 0.005$. This suggests that individuals with higher self-efficacy tend to experience greater motivation and cognitive load. Additionally, motivation was positively correlated with cognitive load, $r(77) = 0.373$, $p <$

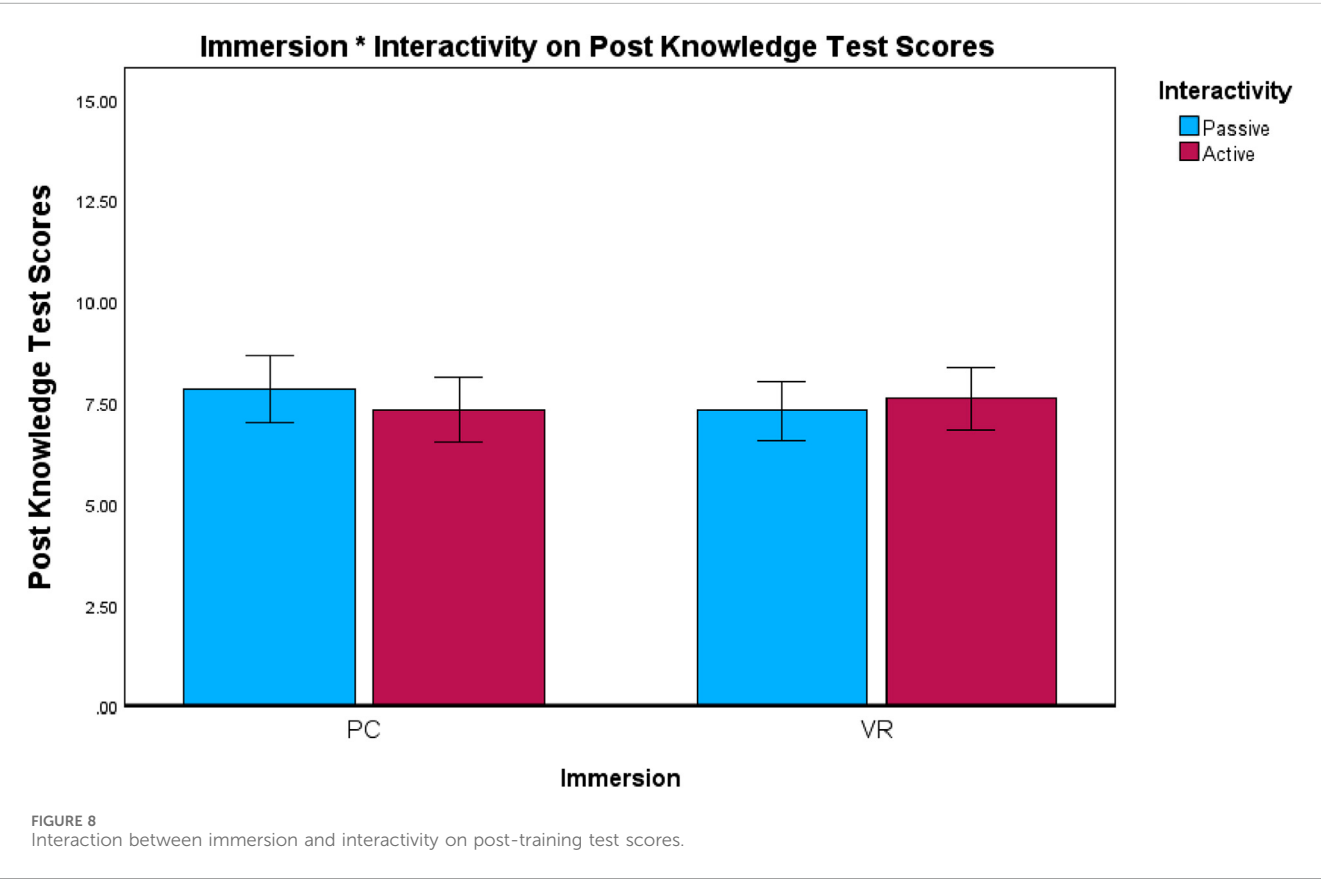
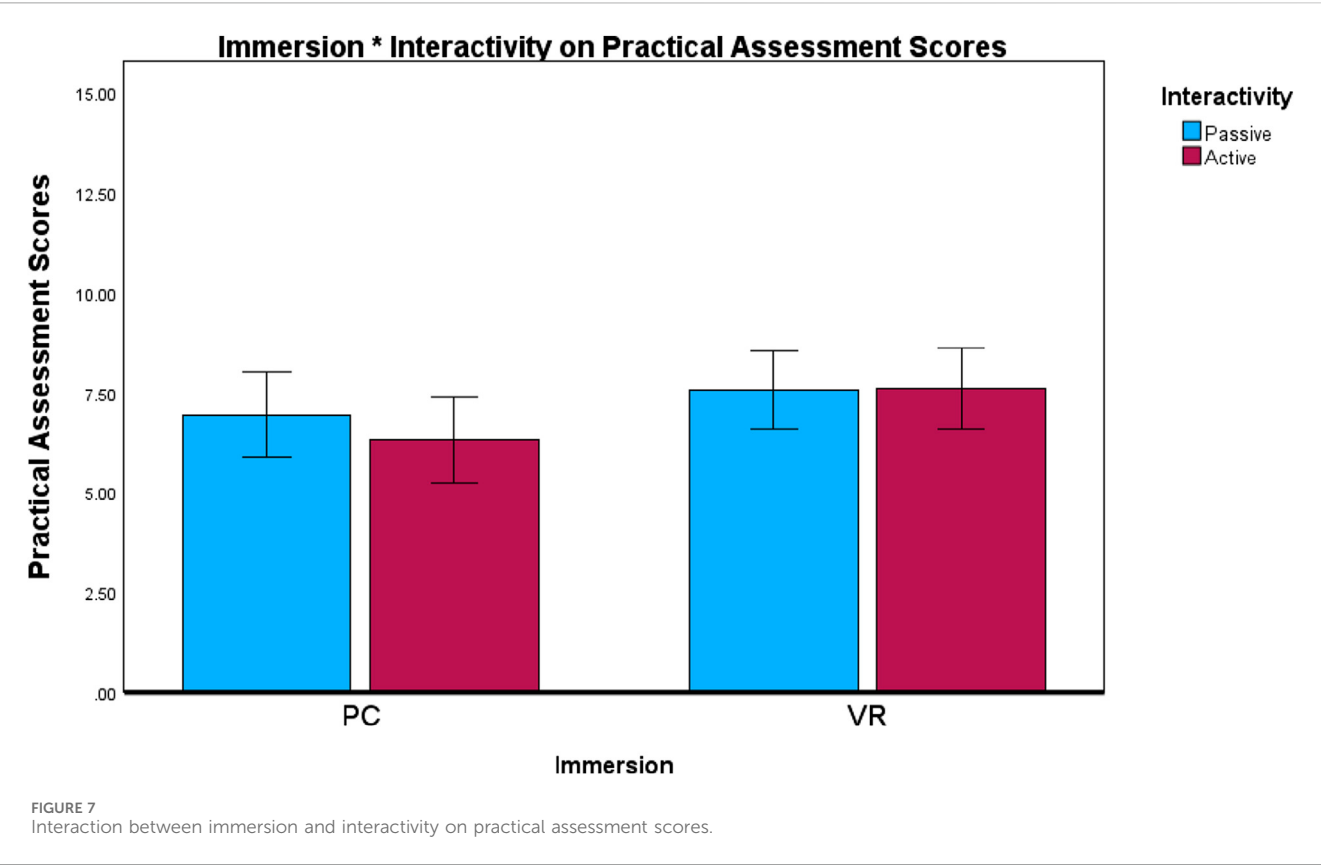


TABLE 6 Assignment check across immersion and interactivity.

Variable	Immersion/Interactivity	N	M
Spatial ability	PC	37	10.44
	VR	42	10.53
	Passive	40	10.76
	Active	39	10.21
Extraversion	PC		5.34
	VR		4.80
	Passive		4.65
	Active		5.49
Agreeableness	PC		5.64
	VR		4.81
	Passive		5.28
	Active		5.17
Conscientiousness	PC		4.41
	VR		5.06
	Passive		4.54
	Active		4.92
Emotional stability	PC		5.77
	VR		5.77
	Passive		5.41
	Active		4.79
Openness	PC VR		4.79
	Passive Active		4.45
			4.88
			4.35
Pretraining test	PC		6.95
	VR		6.53
	Passive		6.57
	Active		6.92

Note. $p < .05$; $N = 79$.

0.001, indicating that as motivation increases, so does cognitive load. Cognitive load also demonstrated a significant positive correlation with post-knowledge test scores, $r(77) = 0.435$, $p < 0.001$. This implies that a higher cognitive load is associated with better performance on the post-knowledge test. Furthermore, pre-knowledge test scores were positively correlated with post-knowledge test scores, $r(77) = 0.266$, $p = 0.018$, suggesting that individuals who performed well before training were likely to maintain their performance after training. No other significant correlations were found among the variables.

5 Discussion

5.1 Empirical contribution

The goal of the present study was twofold: first, to examine the influence of the technological affordances of VR (immersion and interactivity) on procedural learning; second, to test the influence of individual differences factors (sex, spatial ability, and personality traits) on procedural learning. The results revealed no significant main effects or interactions of immersion and interactivity on learning outcomes; even if the means trended in the direction as predicted in the hypotheses. The findings thus did not support our hypothesis that higher immersion and higher interactivity would influence procedural learning outcomes in virtual reality. This was consistent with Johnson et al. (2022) finding that students performed equally well in learning a

mechanical maintenance procedure using both DVR and IVR. These results suggest that deploying a HMD to learn procedures in VR may not be strictly necessary if those same skills can be effectively learned using a desktop PC, keyboard, and mouse. However, further research is needed to confirm these findings and explore the underlying mechanisms.

Several studies have explored the impact of technological features, such as immersion and interactivity on learning in virtual reality have yielded insights consistent with our findings (Makransky and Petersen, 2021; Petersen et al., 2022; Yang et al., 2023). According to the cognitive-affective model of immersive learning proposed by Makransky and Petersen (2021), effective learning in VR requires the integration of cognitive and affective factors, not just immersion and interactivity because these variables in themselves do not directly affect learning. Similarly, Makransky et al. (2019) and Petersen et al. (2022) demonstrated that while immersion can enhance feelings of presence and enjoyment in the virtual learning environment, it does not necessarily lead to better learning outcomes. These findings suggest that the relationship between technological affordances and learning outcomes is complex and mediated by additional factors, such as cognitive load, engagement, and individual learner characteristics.

We also found sex and spatial ability as significant predictors of procedural learning. Previous studies have shown that sex, spatial ability, and personality traits can significantly influence learning performance across different modalities. For example, individuals with higher spatial ability are often better equipped to understand and manipulate 3D environments, giving them a distinct advantage in VR learning tasks (Keehner et al., 2008). Conversely, Lee and Wong (2014) found that individuals with low spatial ability performed well in DVR. Similarly, sex differences in spatial cognition, such as men often outperforming women on mental rotation tasks, may partially explain variations in performance outcomes in spatially demanding VR tasks. Research suggests that this disparity may stem from differences in cognitive strategies. Men are more likely to use holistic, spatial visualization approaches that align with the demands of mental rotation tasks, whereas women often adopt more analytical, step-by-step strategies, which may be less effective (Voyer et al., 1995). Another contributing factor to disparities in spatial ability is experience. Experience plays a critical role in developing spatial skills. Men often have greater exposure to activities that enhance spatial reasoning, such as video games, sports, and construction-related tasks, which provide repeated practice and foster the development of spatial visualization abilities. For personality traits, our results showed no significant effect of any of the items in the five-factor model on learning outcomes. However, research has found conscientiousness and openness to experience to be associated with learning transfer in VR (Thorpe et al., 2023). Conscientious learners may be more likely to engage with task instructions and maintain focus, leading to better procedural skill acquisition, while individuals with high openness to experience may benefit from the novelty and engagement offered by immersive VR environments. Overall, these findings emphasize the importance of considering individual differences when designing and implementing VR across any training system, by tailoring VR environments to address diverse learner needs such as providing adaptive support for those with low spatial ability. Future research

should further explore these interactions to develop personalized VR environments that maximize learning for all users.

5.2 Practical contribution

By examining the effects of technological affordances (immersion and interactivity) and individual differences (sex, spatial ability and personality traits), this research suggests strategies for optimizing the design and use of VR learning systems. This study found no significant effect of immersion or interactivity on learning outcomes, suggesting that the decision to implement VR should be driven by the specific requirements of the learning task rather than the assumption that VR inherently enhances learning. For tasks that do not require physical immersion or high interactivity, cost-effective desktop-based systems using a keyboard and mouse may suffice, offering similar outcomes for procedural learning. Training organizations can use these insights to allocate resources more strategically and avoid unnecessary expenditures on high-end VR setups. Furthermore, the significant effects of individual differences such as sex, spatial ability, and cognitive load on learning outcomes suggest the need for personalized and adaptive training in VR which artificial intelligence can provide. Learners with low spatial ability may benefit from additional guidance, such as interactive tutorials, scaffolding, or simplified spatial tasks to reduce cognitive load and improve performance. Additionally, the practical assessment results showed that VR is particularly beneficial for tasks that require physical practice or spatial navigation. Industries where hands-on skill acquisition is critical, such as healthcare, aviation, and construction, can benefit from using VR for procedural training in safe and controlled environments.

6 Limitations and suggestions for future research

The study identified that learning occurs in VR and individual differences like spatial ability and sex could affect procedural learning outcomes (Petersen et al., 2022). However, a key limitation of this study was the small number of participants, which makes it difficult to generalize the findings to a larger population. The limited sample size also prevented the use of more advanced statistical methods that could have provided deeper insights into how the variables relate to each other. This study was designed to detect larger, more pronounced effects rather than subtle, smaller effects, which may have gone unnoticed due to the constraints of the sample size. Also, the effects observed were relatively small, meaning the results should be interpreted with caution, as the study may not have been adequately powered to capture more nuanced relationships. This raises questions about the generalizability of the findings to broader populations or different contexts. Future research with larger samples could help strengthen these findings, enabling the detection of both large and small effects and allowing for more detailed and sophisticated analyses that better capture the complexity of these relationships.

Also, several key constructs were not explored in this study. Factors such as presence, agency, and representational fidelity,

which are known to play significant roles in VR experiences, were not measured. These variables may have provided deeper insights into how users interact with and experience VR environments, potentially influencing learning outcomes. For instance, presence which is the feeling of being physically there in the virtual environment (McCeery et al., 2013; Schroeder, 2002) could enhance engagement and procedural learning.

Agency and representational fidelity which refers to the degree to which VR accurately represents real-world experience (Bonfert et al., 2024) could also impact the effectiveness of VR training. Future studies should include these constructs to assess their influence on procedural learning outcomes.

Furthermore, stepwise multiple regression has been criticized for several key methodological weaknesses. It often capitalizes on chance, leading to overfitting and the inclusion of spurious predictors that may not generalize to other datasets. The method is also inherently unstable, with small changes in the data producing drastically different models (Antonakis and Dietz, 2011).

Additionally, stepwise regression ignores theoretical considerations, focusing solely on statistical criteria, which can result in models lacking interpretability and meaningfulness (Antonakis and Dietz, 2011). Moreover, the iterative selection process inflates the risk of Type 1 errors and biases the estimation of coefficients, making them unreliable. Also, it fails to address fundamental assumptions of regression, such as homoscedasticity and measurement error, further undermining the validity of its results (Antonakis and Dietz, 2011). Future research should replicate the results of this study to validate its conclusions.

In addition, although cybersickness was not systematically assessed using validated instruments such as the Simulator Sickness Questionnaire (SSQ), no participants explicitly reported symptoms like nausea, fatigue, or eye strain during or after the VR session. The only participant who experienced discomfort had been feeling unwell prior to the study and was excluded from the analysis. A 10-minute break was also provided between the 20-minute VR training and the 10-minute assessment, reducing the likelihood of continuous exposure-related effects. While no adverse effects were observed in this study, future research could benefit from including standardized cybersickness assessments and documenting post-session experiences to refine exposure durations and system design so as to optimize learning experiences and outcomes.

Another limitation is the content of the training itself, which was aviation-based and targeted novice participants. The specific context of aviation training may limit the generalizability of these findings to other domains. Moreover, since participants were novices, their unfamiliarity with the subject matter may have influenced the results. Future research could explore how prior experience or expertise in the training content affects procedural learning outcomes in VR. Exploring whether novices and experts benefit differently from VR training could provide insights into how to tailor VR environments to meet the needs of different learners.

Finally, future studies should examine how combinations of technological features and individual differences influence learning across various domains. Such research will help refine VR applications for training and ensure they deliver maximum value to learners.

7 Conclusion

We explored the technology features (immersion and interactivity) of VR and the individual differences factors that are predicted to influence procedural learning in virtual reality. We found no significant difference to suggest that higher immersion or higher interactivity meant better procedural learning. However, sex and spatial ability were significant predictors of procedural learning. It shows that while VR offers unique advantages for practical assessments and physical skill development, its value is task-specific and dependent on individual learner characteristics. Training organizations should carefully evaluate the necessity and cost-effectiveness of VR for their training objectives, ensuring that technological solutions align with the needs of the learners. By adopting targeted, evidence-based strategies, VR can be effectively integrated into learning environments to optimize procedural knowledge and ensure equitable access to training opportunities.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by University of Central Florida Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

FD: Investigation, Writing – review and editing, Conceptualization, Writing – original draft, Methodology, Formal Analysis. VS: Writing – review and editing, Supervision, Methodology. FJ: Software, Resources, Validation, Writing – review and editing, Supervision, Methodology.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was in part supported by the US Department of Transportation/Federal Aviation Administration (FAA) collaborative research agreement

References

Akindele, N., Taiwo, R., Sarvari, H., Oluleye, B., Awodele, I. A., and Olaniran, T. O. (2024). A state-of-the-art analysis of virtual reality applications in construction health and safety. *Results Eng.* 23, 102382. doi:10.1016/j.rineng.2024.102382

692M151940002 and 692M152440003; program manager: FAA ANG-C1, the NextGen Human Factors Division; program sponsors: FAA AVS, the Aviation Safety Office, and FAA AFS-280, the Air Transportation Division—Training & Simulation Group. The views expressed herein are those of the authors and do not reflect the views of the United States (U.S.) Department of Transportation (DOT), FAA, or the University of Central Florida.

Acknowledgments

We sincerely thank the members of the Team Performance Laboratory—Nathan A. Sonnenfeld, Blake Nguyen, and Alex Alonso—for their support and contributions to the simulation design. We also extend our gratitude to the simulation developers—John Benge, Tom Bennett, Adam Lawton, Scott Malo, and David Philipos.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. The author(s) confirm that Generative AI was utilized in the preparation of this manuscript. Specifically, Grammarly was employed to enhance readability while preserving the original content provided by the authors.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frvir.2025.1601562/full#supplementary-material>

Andersen, M. S., and Makransky, G. (2021). The validation and further development of the multidimensional cognitive load scale for physical and online lectures (MCLS-POL). *Front. Psychol.* 12, 642084. doi:10.3389/fpsyg.2021.642084

- Antonakis, J., and Dietz, J. (2011). Looking for validity or testing it? The perils of stepwise regression, extreme-scores analysis, heteroscedasticity, and measurement error. *Personality Individ. Differ.* 50 (4), 409–415. doi:10.1016/j.paid.2010.09.014
- Baddeley, A. (1992). Working memory. *Science* 255 (5044), 556–559. doi:10.1126/science.1736359
- Barrett, R. C. A., Poe, R., O'Camb, J. W., Woodruff, C., Harrison, S. M., Dolguikh, K., et al. (2022). Comparing virtual reality, desktop-based 3D, and 2D versions of a category learning experiment. *Plos One* 17 (10), e0275119. doi:10.1371/journal.pone.0275119
- Bonfert, M., Muender, T., McMahan, R. P., Steinicke, F., Bowman, D., Malaka, R., et al. (2024). The Interaction Fidelity Model: A Taxonomy to Communicate the Different Aspects of Fidelity in Virtual Reality. *International Journal of Human-Computer Interaction*, 41 (12), 7593–7625. doi:10.1080/10447318.2024.2400377
- Bowman, D. A., and McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer* 40 (7), 36–43. doi:10.1109/MC.2007.257
- Buttussi, F., and Chittaro, L. (2017). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Trans. Vis. Comput. Graph.* 24 (2), 1063–1076. doi:10.1109/TVCG.2017.2653117
- Chandler, P., and Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition Instr.* 8 (4), 293–332. doi:10.1207/s1532690xci0804_2
- Chang, M. M., and Hung, H. T. (2019). Effects of technology-enhanced language learning on second language acquisition. *J. Educ. Technol. and Soc.* 22 (4), 1–17. Available online at: <https://www.jstor.org/stable/26910181>.
- Coban, M., Bolat, Y. I., and Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: a meta-analysis. *Educ. Res. Rev.* 36, 100452. doi:10.1016/j.edurev.2022.100452
- Conrad, M., Kablitz, D., and Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: a systematic review of findings. *Comput. and Educ. X Real.* 4, 100053. doi:10.1016/j.cexr.2024.100053
- Cummings, J. J., and Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychol.* 19 (2), 272–309. doi:10.1080/15213269.2015.1015740
- De Lorenzis, F., Praticco, F. G., Repetto, M., Pons, E., and Lamberti, F. (2023). Immersive Virtual Reality for procedural training: comparing traditional and learning by teaching approaches. *Computers in Industry*, 144, 103785. doi:10.1016/j.compind.2022.103785
- Ekstrom, R. B., and Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*, 1976. Princeton, NJ: Educational testing service.
- Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi:10.3758/BF03193146
- Ganier, F., Hoareau, C., and Tisseau, J. (2014). Evaluation of procedural learning transfer from a virtual environment to a real situation: a case study on tank maintenance training. *Ergonomics* 57 (6), 828–843. doi:10.1080/00140139.2014.899628
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B., Jr. (2003). A very brief measure of the big five personality domains. *J. Res. Personality* 37, 504–528. doi:10.1016/s0092-6566(03)00046-1
- Grassini, S., and Laumann, K. (2020). Are modern head-mounted displays sexist? A systematic review on gender differences in HMD-mediated virtual reality. *Front. Psychol.* 11, 1604. doi:10.3389/fpsyg.2020.01604
- Guthridge, R., and Clinton-Lisell, V. (2023). Evaluating the efficacy of virtual reality (VR) training devices for pilot training. *J. Aviat. Technol. Eng.* 12 (2), 1. doi:10.7771/2159-6670.1286
- Hamilton, D., McKechnie, J., Edgerton, E., and Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *J. Comput. Educ.* 8 (1), 1–32. doi:10.1007/s40692-020-00169-2
- Hight, M. P., Fussell, S. G., Kurkchubasche, M. A., and Hummell, I. J. (2022). Effectiveness of virtual reality simulations for civilian, *ab initio* pilot training. *J. Aerosp. Educ. and Res.* 31 (1). doi:10.15394/jaer.2022.1903
- Huk, T. (2006). Who benefits from learning with 3D models? The case of spatial ability. *J. Comput. Assisted Learn.* 22 (6), 392–404. doi:10.1111/j.1365-2729.2006.00180.x
- Jang, S., Vitale, J. M., Jyung, R. W., and Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Comput. and Educ.* 106, 150–165. doi:10.1016/j.compedu.2016.12.009
- Johnson, C. I., Bailey, S. K., Schroeder, B. L., and Marraffino, M. D. (2022). Procedural learning in virtual reality: the role of immersion, interactivity, and spatial ability. *Technol. Mind, Behav.* 3. doi:10.1037/tmb0000087
- Johnson-Glenberg, M. C. (2018). Immersive VR and education: embodied design principles that include gesture and hand controls. *Front. Robotics AI* 5, 81. doi:10.3389/frobt.2018.00081
- Jongbloed, J., Chaker, R., and Lavoué, E. (2024). Immersive procedural training in virtual reality: a systematic literature review. *Comput. and Educ.* 221, 105124. doi:10.1016/j.compedu.2024.105124
- Keekner, M., Hegarty, M., Cohen, C., Khooshabeh, P., and Montello, D. R. (2008). Spatial reasoning with external visualizations: what matters is what you see, not whether you interact. *Cognitive Sci.* 32 (7), 1099–1132. doi:10.1080/03640210801898177
- Khorasani, S., Syiem, B. V., Nawaz, S., Knibbe, J., and Velloso, E. (2023). Hands-on or hands-off: deciphering the impact of interactivity on embodied learning in VR. *Comput. and Educ. X Real.* 3, 100037. doi:10.1016/j.cexr.2023.100037
- Komaraju, M., Karau, S. J., Schmeck, R. R., and Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality Individ. Differ.* 51 (4), 472–477. doi:10.1016/j.paid.2011.04.019
- Lee, E. A. L., and Wong, K. W. (2014). Learning with desktop virtual reality: low spatial ability learners are more positively affected. *Comput. and Educ.* 79, 49–58. doi:10.1016/j.compedu.2014.07.010
- Lee, E. A. L., Wong, K. W., and Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Comput. and Educ.* 55 (4), 1424–1442. doi:10.1016/j.compedu.2010.06.006
- Lin, X. P., Li, B. B., Yao, Z. N., Yang, Z., and Zhang, M. (2024). The impact of virtual reality on student engagement in the classroom—a critical review of the literature. *Front. Psychol.* 15, 1360574. doi:10.3389/fpsyg.2024.1360574
- Liu, Z., Zhang, W., and Hu, L. (2024). Can desktop virtual reality effectively enhance academic achievement? a meta-analysis. *Innovations Educ. Teach. Int.* 62, 1–16. doi:10.1080/14703297.2024.2354753
- Li, K. C., Wong, B. T. M., Li, P., Klippel, A., and Zhao, J. (2020). Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain*, 1 (1), 28–36. doi:10.37716/HBAB.2020010105
- Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., et al. (2017). Application of virtual reality technology in clinical medicine. *American Am. J. Transl. Res.*, 9 (9), 3867.
- MacArthur, C., Grinberg, A., Harley, D., and Hancock, M. (2021). “You’re making me sick: a systematic review of how virtual reality research considers gender and cybersickness,” in *Proceedings of the 2021 CHI conference on human factors in computing systems*, 1–15. doi:10.1145/3411764.3445701
- Makransky, G., and Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: a structural equation modeling approach. *Comput. and Educ.* 134, 15–30. doi:10.1016/j.compedu.2019.02.002
- Makransky, G., and Petersen, G. B. (2021). The cognitive-affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. *Educ. Psychol. Rev.* 33 (3), 937–958. doi:10.1007/s10648-020-09586-2
- Makransky, G., Terkildsen, T. S., and Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learn. Instr.* 60, 225–236. doi:10.1016/j.learninstruc.2017.12.007
- Marougkas, A., Troussas, C., Krouska, A., and Sgourpoulou, C. (2024). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. *Multimed. Tools Appl.* 83, 18185–18233. doi:10.1007/s11042-023-15986-7
- Mayer, R. E. (2017). Using multimedia for e-learning. *J. Comput. Assisted Learn.* 33 (5), 403–423. doi:10.1111/jcal.12197
- Mayer, R. E. (2001). *Multimedia Learning*. Cambridge University Press.
- Mayer, R. E. (2021). *Multimedia learning*. 3rd ed. Cambridge University Press, 57–72.
- Mayer, R. E. (2022). “Cognitive theory of multimedia learning,” in *The Cambridge handbook of multimedia learning*. Editors R. E. Mayer and L. Fiorella 3rd ed (Cambridge University Press), 57–72.
- Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educ. Psychol. Rev.* 36 (1), 8. doi:10.1007/s10648-023-09842-1
- Mayer, R. E., Makransky, G., and Parong, J. (2023). The promise and pitfalls of learning in immersive virtual reality. *Int. J. Human-Computer Interact.* 39 (11), 2229–2238. doi:10.1080/10447318.2022.2108563
- McCreery, M. P., Schrader, P. G., Krach, S. K., and Boone, R. (2013). A sense of self: the role of presence in virtual environments. *Comput. Hum. Behav.* 29, (4), 1635–1640. doi:10.1016/j.chb.2013.02.002
- Meyer, O. A., Omdahl, M. K., and Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: a media and methods experiment. *Comput. and Educ.* 140, 103603. doi:10.1016/j.compedu.2019.103603
- Munafu, J., Diedrick, M., and Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Exp. Brain Res.* 235, 889–901. doi:10.1007/s00221-016-4846-7
- Nguyen, B., Sonnenfeld, N., Duruaku, F., and Jentsch, F. (2023). Modern training practices: studying the use of virtual reality for training procedures in flightcrew training. Available online at: <https://rosap.nsl.bts.gov/view/dot/76580>.
- Ntakakis, G., Plomariti, C., Frantzidis, C., Antoniou, P. E., Bamidis, P. D., and Tsoulfas, G. (2023). Exploring the use of virtual reality in surgical education. *World J. Transplant.* 13 (2), 36–43. doi:10.5500/wjt.v13.i2.36
- Paivio, A. (1990). *Mental representations: a dual coding approach*. Oxford University Press.

- Parong, J., and Mayer, R. E. (2018). Learning science in immersive virtual reality. *J. Educ. Psychol.* 110 (6), 785–797. doi:10.1037/edu0000241
- Petersen, G. B., Petkakis, G., and Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Comput. and Educ.* 179, 104429. doi:10.1016/j.compedu.2021.104429
- Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput. and Educ.* 147, 103778. doi:10.1016/j.compedu.2019.103778
- Rosenthal, R., Schäfer, J., Hoffmann, H., Vitz, M., Oertli, D., and Hahnloser, D. (2013). Personality traits and virtual reality performance. *Surg. Endosc.* 27 (1), 222–230. doi:10.1007/s00464-012-2424-z
- Salzman, M. C., Dede, C., Loftin, R. B., and Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual learning. *Presence Teleoperators and Virtual Environ.* 8 (3), 293–316. doi:10.1162/105474699566242
- Scorgie, D., Feng, Z., Paes, D., Parisi, F., Yiu, T. W., and Lovreglio, R. (2024). Virtual reality for safety training: a systematic literature review and meta-analysis. *Saf. Sci.* 171, 106372. doi:10.1016/j.ssci.2023.106372
- Schroeder, R. (2002). Co-presence and interaction in virtual environments: an overview of the range of issues. In *Presence 2002: Fifth International Workshop*, 274–295.
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Trans. R. Soc. Biol. Sci.* 364 (1535), 3549–3557. doi:10.1098/rstb.2009.0138
- Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., and Oliva, R. (2022). A separate reality: an update on place illusion and plausibility in virtual reality. *Front. Virtual Real.* 3, 914392. doi:10.3389/frvir.2022.914392
- Slater, M., and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoperators and Virtual Environ.* 6 (6), 603–616. doi:10.1162/pres.1997.6.6.603
- Sneider, J. T., Hamilton, D. A., Cohen-Gilbert, J. E., Crowley, D. J., and Rosso, I. M. (2015). Sex differences in spatial navigation and perception in human adolescents and emerging adults. *Behav. Process.* 111, 42–50. doi:10.1016/j.beproc.2014.11.015
- Sonnenfeld, N. A., Nguyen, B., Duruaku, F., Alonso, A., Carney, M., Smith, E., et al. (2023). Flightcrew procedures experimental training: simulation testbed design, development, and interactive demonstration. *Proc. Hum. Factors Ergonomics Soc. Annu. Meet.* 67 (1), 253–258. doi:10.1177/21695067231192919
- Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., et al. (2020). Identifying causes of and solutions for cybersickness in immersive technology: reformulation of a research and development agenda. *Int. J. Human-Computer Interact.* 36 (19), 1783–1803. doi:10.1080/10447318.2020.1828535
- Streiner, D. L., and Norman, G. R. (2011). Correction for multiple testing: is there a resolution? *Chest* 140 (1), 16–18. doi:10.1378/chest.11-0523
- Thorp, S. O., Rimol, L. M., and Grassini, S. (2023). Association of the big five personality traits with training effectiveness, sense of presence, and cybersickness in virtual Reality. *Multimodal Technol. Interact.* 7 (2), 11. doi:10.3390/mti7020011
- Urhahne, D., Chao, S. H., Florineth, M. L., Luttenberger, S., and Paechter, M. (2011). Academic self-concept, learning motivation, and test anxiety of the underestimated student. *Br. J. Educ. Psychol.* 81 (Pt 1), 161–177. doi:10.1348/000709910X504500
- Voyer, D., Voyer, S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117 (2), 250–270. doi:10.1037/0033-2909.117.2.250
- Wang, W. S., Lee, H. Y., Lin, C. J., Li, P. H., Huang, Y. M., and Wu, T. T. (2024). Enhancing students' learning outcomes in self-regulated virtual reality learning environment with learning aid mechanisms. *Br. J. Educ. Technol.* 56, 366–387. doi:10.1111/bjet.13512
- Wei, W., Chen, C., and Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: a meta-analysis. *Br. J. Educ. Technol.* 51 (6), 1991–2005. doi:10.1111/bjet.13023
- Wu, B., Yu, X., and Zhou, X. (2016). Spatial ability explains the male advantage in approximate arithmetic. *Front. psychol.* 7, 306. doi:10.3389/fpsyg.2016.00306
- Yang, H., Cai, M., Diao, Y., Liu, R., Liu, L., and Xiang, Q. (2023). How does interactive virtual reality enhance learning outcomes via emotional experiences? A structural equation modeling approach. *Front. Psychol.* 13, 1081372. doi:10.3389/fpsyg.2022.1081372