virtual worlds

MDPI

Article

Exploring Mixed-Interaction Mode in a Virtual Cockpit: Controller and Hand Gesture Integration

Yemon Lee, Andy M. Connor * and Stefan Marks

School of Future Environments, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand; yemonlee@gmail.com (Y.L.); stefan.marks.ac@gmail.com (S.M.)
* Correspondence: andy.connor.phd@gmail.com

Abstract: This paper evaluates a new interaction mode for object manipulation tasks in virtual reality (VR) utilizing an aircraft cockpit simulation. Building on prior research, this study examines the effectiveness and user experience of a mixed-interaction mode that involves the combination of handheld controllers with hand gestures. Qualitative interviews with participants provided detailed feedback on the combined input approach. The analysis highlights the strengths and challenges of the mixed-interaction mode, indicating a perceived increase in task completion efficacy and enhanced user experience. As an outcome of the research, design guidelines were developed based on participants' insights, focusing on the optimal balance of naturalness and precision for mixed interaction in VR that can also be utilized more generally. This study offers practical implications for creating immersive virtual environments and informs future research in VR interaction modes and user experience.

Keywords: virtual reality; object manipulations; human–computer interaction; user experience; multi-modality; user interfaces

check for updates

Academic Editor: Christos J. Bouras

Received: 29 April 2025
Revised: 30 May 2025
Accepted: 5 June 2025
Published: 19 June 2025

Citation: Lee, Y.; Connor, A.M.; Marks, S. Exploring Mixed-Interaction Mode in a Virtual Cockpit: Controller and Hand Gesture Integration. *Virtual Worlds* **2025**, *4*, 28. https://doi.org/10.3390/virtualworlds4020028

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As virtual reality (VR) technologies continue to advance, creating intuitive and effective interaction modes remains a significant challenge, particularly for complex tasks such as those in cockpit environments. Traditional handheld controllers offer precision but may lack naturalness, while hand gesture devices, such as the Leap Motion from Ultraleap (Ultraleap Limited, registered in England and Wales under company number 08781720. The West Wing, Glass Wharf, Bristol, England, BS2 0EL), provide a more intuitive interaction method but often fall short in precision. The aim of this paper is to undertake a qualitative usability evaluation of a system that combines different styles of interaction.

Our prior research [1] has investigated a mixed-interaction mode combining handheld controllers with hand gesture devices, and quantitative data indicated some potential advantages as well as some drawbacks when such a mode is used in generic interaction tasks. This previously published work was intended to be a standalone study evaluating three interaction modes, namely, just controllers, just hand gestures, and the combination of the two. The work focused on quantitative measures, such as task completion time, across three different interaction tasks. However, participants were asked to use a talk-aloud protocol during the evaluation as well as completing a usability questionnaire after the evaluation tasks were completed. An initial triangulation of the data showed interesting inconsistencies, particularly exhibiting different preferences for different tasks, as well as one participant volunteering that they really wanted to see the mixed interaction

in a shooting game where the controller would provide a realistic feeling of carrying a firearm, whereas the gesture interaction would provide a more intuitive way of reloading the ammunition.

This study therefore focuses on a qualitative evaluation of this combined approach in a scenario specifically designed for mixed interaction, specifically an aircraft cockpit simulator where the sidestick is manipulated with the controller while hand gestures are used to interact with other controls on the instrumentation panel. The cockpit simulation has similarities to the shooting game suggested by the participant in our first study but is less polarising to participants. The primary aim is to gain a deeper insight into the user experience from the qualitative data to assess the effectiveness and applicability of the combined input approach, providing insights into user preferences and challenges. This contributes to addressing the overarching research question to determine under what circumstances the mixed-interaction mode provides advantages for object manipulations in VR. The contributions of this research lie in its exploration of the mixed-interaction mode in VR and its potential to inform future design decisions related to interactions in VR.

By analysing participants' feedback, we propose design guidelines that balance naturalness and precision in VR interfaces. Our findings offer practical implications for creating more immersive virtual environments and contribute to the broader field of VR interaction modes and user experience.

2. Related Works

VR can potentially transform training methods using operational experience to improve understanding of complex tasks and procedures [2]. However, designers and developers often face challenges when creating solutions for complex and inconsistent interaction modes involving multiple input and output devices [3,4]. This research is inspired by day-to-day life activities that are rarely performed using the same tools with both hands, e.g., using two knives when eating. Instead, humans typically use a combination of different tools, such as a knife and a fork, each serving a unique purpose and complementing the other. This observation forms the basis for exploring the proposed mixed-interaction mode in VR, which at the time this research started had not been addressed in previous studies.

In the existing literature, various single-mode interactions in VR have been explored, such as hand gesture interaction [5–7], haptic interaction [8,9], eye-gaze interaction [9–11], and speech-based interaction [12,13]. However, there is limited research on combining different interaction modes into a mixed-interaction approach.

The concept of mixed interaction in this study refers to a multimodal interaction method that combines controller-based and hand gesture modes simultaneously within the same virtual environment. This approach allows users to leverage their natural movements and hand gestures while also achieving precise control using a controller [4,5]. Mixed interaction can be particularly useful in tasks that require a combination of different interaction modes, such as using a controller-based mode for precise actions and hand gestures for more expressive interactions. This could potentially provide an opportunity to enhance the user experience within VR.

Some studies have started exploring mixed interaction in VR. For example, Sundén et al. [14] introduced a Hybrid VR touch table that incorporates an interactive multi-touch table combined with an HMD and an optional wireless controller. Their findings highlight the system's flexibility in catering to users who are interested in or prefer either of the exploratory options, as well as those with specific needs or limitations that restrict them to only one option. Wolf et al. [15] conducted a study comparing object manipulation tasks using a combination of speech-and-gesture-based approaches to a typical menu-oriented interface. The comparison focused on three measures: flow, usability, and presence, to

evaluate the effectiveness of the speech-and-gesture-based system. The results show that the speech-and-gesture-based approach had the highest usability rating, and the fewest user errors compared with the menu-oriented interface. Additionally, the speech-and-gesture-based interface was considered the most intuitive and efficient when comparing the strengths and weaknesses of 2D UI, 3D UI, and Speech UI [16].

Wagner et al. [17] evaluated interaction modes for a data manipulation study using virtual hand interaction with a grabbing and stretching action and a virtual ray pointer with actions assigned to controller buttons and a mixed-interaction mode. The results from 15 participants show that mixed interaction did not significantly increase workload or decrease system usability or task ease. Nevertheless, 60% of participants preferred using the mixed-interaction mode for various low-level tasks over the other two modes, while 40% expressed that the mixed-interaction mode could be confusing. Due to the small sample size, the results for data manipulation did not demonstrate a significant effect from using mixed interaction. The researchers suggested that designers should select interaction modes that favour specific tasks and believe that integrating different interaction modes is necessary for data manipulation in the future to overcome the limitations of specific interaction modes.

Mixed Interaction

The concept of mixed interaction refers to a multimodal interaction method that combines a controller and a hand gesture device simultaneously within the same virtual environment. Olmedo et al. [18] emphasised that introducing multimodal interaction, including speech and gesture, significantly enriched the user experience in VR applications. The authors underscored the need for standardised multimodal systems to enhance usability and accessibility. This approach allows users to leverage their natural movements and hand gestures while achieving precise control using a controller [19,20]. Mixed interaction can be particularly useful in tasks that require a combination of different interaction modes, such as using a controller for precise actions and hand gestures for more expressive interactions. This could potentially provide an opportunity to enhance the user experience within VR.

Huang et al. [19] aimed to explore the effectiveness of combining hand gesture and controller inputs in immersive environments. The study involved 22 participants, and the results indicated that the efficiency of task execution using hybrid inputs was comparable to using bimanual controllers. Specifically, the combination of holding a controller in the dominant hand and performing hand gestures with the non-dominant hand showed potential for efficient performance. The research provided the following tasks and design guidelines for the hybrid input:

1. Task Suitability: Hybrid inputs are suitable for certain types of tasks. Users show a greater willingness to use hybrid inputs when facing complicated or bimanual tasks, while they may be less interested in using them for simple tasks.
2. Device Assignment: Assigning the appropriate devices for each action is crucial for the effectiveness of hybrid inputs. Operations requiring stability and precise control, such as HOLD and AIM actions, should be assigned to controllers due to their better tracking capabilities.
3. Providing Hints: Users may make incorrect device and action pairings when using hybrid inputs. Providing hints, such as adjusting the position of game objects, can guide users to use the appropriate device for each action, enhancing performance.
4. Consider Hand Dominance: When the correct pairing of device and action is used, the influence of dominant and non-dominant hands becomes evident. Users often

prefer using their dominant hand for actions like MOVE. Designers should consider hand dominance when designing interactions in hybrid input systems.

Despite these findings, only a few studies in the existing literature explore the potential of combined interaction, specifically for controllers and hand gestures. Huang et al. [19] focused primarily on game-like tasks within immersive environments, which presented preliminary results that justify further research in the area.

This study aims to contribute to this field by investigating the potential impacts of different VR interaction tasks on user performance, engagement, and experience. It extends the evaluation beyond game-like tasks to explore the effectiveness of hybrid inputs in other domains. This approach allows for a more comprehensive understanding of the usability, efficacy, and user satisfaction associated with hybrid inputs across diverse contexts. By conducting evaluations in these simulated scenarios, the research builds upon the work conducted by Huang et al. [19] and contributes to the ongoing development of knowledge regarding mixed interaction in VR.

3. Method

This study was designed as an evaluation of a novel mode of interaction for virtual environments that combined the use of a single handheld controller in conjunction with gestures. Our previous study [1] undertook a quantitative evaluation of this mode of interaction on generic interaction tasks which led to mixed outcomes. This study therefore focuses on using specific tasks that are designed for the mixed interaction approach and focuses on qualitative data to provide an in-depth understanding of the user experience. The following sections describe the design of the environment, the tasks, data collection, and data analysis methods.

3.1. Cockpit Model Design

The intention of this study is not to provide realistic flight deck training. Instead it is to evaluate the novel interaction mode that combines controllers and gestures using the cockpit environment as a specific case of a task designed in such a way that the combination may provide advantages.

In order to maximise the potential of the novel interaction mode, the strategic placement of controls and tools at logical locations is crucial [21]. For example, in this scenario, a handheld VR controller can be used to operate the sidestick, which is a joystick used to control the aircraft's movement, while optical hand-tracking technology enables users to manipulate buttons and knobs on the main instrument panels using hand gestures. Aligning the design of the environment with the types of interactions that suit the mixed interaction facilitates a robust evaluation of the mode of interaction.

The cockpit environment was created based on the A350 Airbus flight deck system and is shown in Figure 1. This specific airplane model was chosen as schematics are readily available to facilitate the development of a reasonably realistic environment as a context in which to define generic interaction tasks. The intention was not to create a fully working flight deck simulator. The cockpit was modelled in Unity3D with the textures of the main instrument panel, glareshield, and pedestal derived from photographs of the interior of a real A350. This approach ensures a high level of detail and realism in the VR cockpit, which is intended to enhance the user's immersive experience. The textures not only contribute to the visual appeal but also play a crucial role in user interaction. Users can interact with these textured elements in a manner similar to a real flight deck, thereby improving the authenticity of the simulation.

Figure 1. Cockpit model used as the basis for the interaction tasks.

Figure 2 presents a view of the mixed-interaction mode being used with the cockpit simulator, as seen from the perspective of the user. The gloved hand is a representation for the VR controller mode being used to manipulate the sidestick on the left side. The semi-transparent hand is a representation of the real hand being optically tracked that is being used to manipulate the various cockpit elements such as knobs, switches, and levers.

Figure 2. User perspective of the mixed-interaction mode with different hand styles representing the controller and gesture-based styles of interaction.

3.2. Setup and User Tasks

This user study is designed to explore the user experience of using mixed interaction in a context where the interaction tasks are designed around a mixed mode of interaction. In the cases of this research, an aircraft flight deck was selected as the environment; however, the tasks were suitably abstracted on purpose as the intent was to evaluate the interaction mode, not determine its potential value for pilot training. Participants were asked to sit in the middle of the room, and they were given an HTC Vive headset with a Leap Motion

mounted at the front and one HTC Vive controller before entering the virtual environment. Figure 3 shows a participant with the controller in their left hand and the Leap Motion attached to the front of the headset. The HTC Vive used in this study was the Vive Pro model, running on SteamVR version 1.18.5. The experiment was conducted in a study room at the AUT campus, approximately 4 m × 3 m in size. Two base stations were positioned diagonally in opposite corners at a height of about 2.2 m to provide stable tracking coverage and minimise occlusion.

Figure 3. The combination of controller and gestures in use.

The tracking data from the controller and Leap Motion were processed within the same Unity3D environment. Both devices were connected to a single computer to support simultaneous input. Occasional latency was observed when participants switched quickly between hand gestures and controller use, typically due to the limited tracking range of the Leap Motion or minor changes in lighting conditions. These issues were infrequent and were mitigated through recalibration and by asking participants to keep their hands within the visible tracking area.

The participants were seated in the physical space at a location that was mapped to the captain's seat on the left-hand side of the cockpit, as seen in Figure 4.

This user study was designed to assess the participants' experience in a cockpit scenario, which was intentionally abstracted from real flight desk tasks as none of the participants were commercial pilots. The participants were asked to perform a series of tasks, such as manipulating virtual knobs, pressing virtual buttons, and pulling virtual levers using a Leap Motion controller with their right hand. The scenario also focused on the participants' ability to use the HTC Vive controller to balance the horizon with their left hand.

Figure 4. Captain Seat on the Left in the Virtual Cockpit Environment.

All the participants were able to navigate through the six steps in the cockpit training simulator with verbal guidance. The tasks were explained to them, and they were shown how to use the controller and Leap Motion device, but no practice rounds were conducted. This approach was chosen to assess the intuitiveness of the mixed-interaction mode for novices in a complex environment. The tasks were as follows:

1. Turn on the orange integration button (Bottom right of front panel)
2. Rotate one of the knobs to the right (Front panel)
3. Flick a white switch (Pedestal panel on the right side)
4. Push the lever (On the pedestal panel)
5. Use the sidestick to balance the position of the plane (Left-hand side)
6. Press the green autopilot button on the front panel (Middle of the front panel)

A more detailed view of the flight controls and their association with the six tasks is shown in Figure 5.

Figure 5. Location of controls associated with the six interaction tasks in the preceding list.

3.3. Participants

Eight participants, all with prior VR experience and who had been involved in the previous quantitative study [1], were individually invited to participate in this study. The fact that the participants all had previous VR experience, and indeed had all utilized the mixed-interaction mode in our previous work, does lead to potential bias. If nothing else, generalizing from this group to a wider population is potentially problematic. Despite this, as the work is exploratory in nature, no attempt was made to account for any bias. The small number of participants was due to the exploratory nature of this study, focusing on in-depth qualitative data rather than quantitative data. Each participant attended a 30 min session. This study primarily aimed to assess the proficiency of interaction techniques in a new context. Future studies could consider a common baseline for VR technology abilities among participants for more controlled results.

The number of participants is relatively low but suitable for an explorative usability study. Some authors argue that just 5 users are able to highlight 80% of usability problems in prototypes [22]. Others find that 10 ± 2 participants are sufficient [23]. Schmettow [24] also argues that there is no 'magic number' and that a great discrepancy between general and expert users can be demonstrated. However, the number of participants is still considered a limitation and this will be discussed later in this paper.

3.4. Data Collection

Qualitative data was gathered through the use of semi-structured interviews with the users. These interviews allowed for an in-depth exploration of users' experiences, perceptions, and thoughts, providing rich, contextual insights that served to enhance the findings. Being semi-structured in nature, the core questions provided consistency in data collection whilst also allowing follow-up questions to probe more deeply and add richness to the data. The core interview questions were as follows:

Q1: How did you find the interaction task?
Q2: How demanding, both physically and mentally, did you find the task?
Q3: How successful do you think you were in conducting the task?
Q4: How did the mixed-interaction mode contribute to your success?
Q5: What will be the difference if you were given a set of controllers or using only the Leap Motion?
Q6: What sort of use cases can you see as benefiting from this type of mixed interaction?

The follow-up questions typically included probes such as: "Why?", "Can you elaborate?", "Can you tell me more about that?", and "how do you feel?".

3.5. Data Analysis

The data analysis approach used in this study was the Reflexive Thematic Analysis (RTA) method proposed by Braun and Clarke [25], which was implemented following the worked example provided by Byrne [26], which consists of six phases. After familiarization with the data, the main phases consist of generating initial codes, generating themes, reviewing themes, and finally defining the themes. These four phases are not a linear process, and whilst iterative in nature, it is still necessary to respond to changes that emerge in the process.

RTA was specifically chosen over alternative thematic analysis approaches that adopt either codebooks or coding reliability measures as the coding and data analysis was conducted by the lead author alone, and RTA acknowledges the role of the researcher and it embraces the subjectivity of the approach, allowing themes to be produced at the intersection of the researcher's theoretical assumptions, their analytic resources and skill, and the data itself [26].

According to Braun and Clarke [25], thematic analysis is used to analyze the interview data, where codes are developed from the transcripts and categorized into initial themes, which are then further analyzed to produce main themes and potentially sub-themes. This is generally an iterative process where the raw data is coded and used to identify meanings, which are then organized into patterns. These patterns are then further interpreted and refined into themes relevant to the study context [27]. The analysis can reveal the participants' opinions, attitudes, and beliefs and provide insights into the data to help answer the research question.

4. Results

The interviews in the cockpit study were intended to gather qualitative data about the user experience of the mixed-interaction mode under evaluation. Details of the analysis process and the outcomes are described in the following sections.

4.1. Analysis Process

In the initial code generation, the transcribed data of each interview was analyzed with interesting statements coded to begin with. This process was repeated several times until the un-coded data could be confirmed to be of no interest to the study. The coded text was then manually arranged by similarity, again an iterative process, which led to defining initial themes based on the clustered. The initial thematic map is shown in Figure 6.

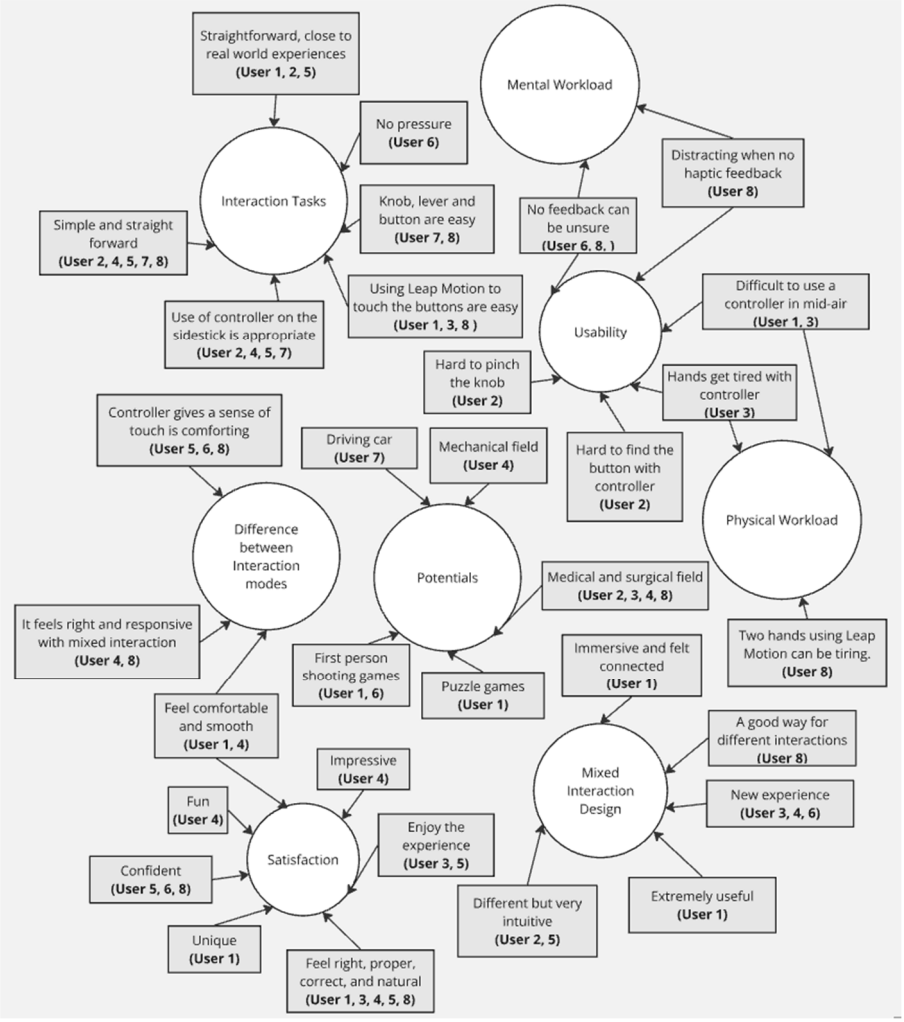

Figure 6. Initial thematic map.

The themes were primarily aligned as might be expected, by focusing on areas such as experience with interaction tasks, the mental and physical workload in the mixed-interaction mode, user satisfaction, usability, innovation experience, and potential for the mixed-interaction mode. However, it can be seen that this initial thematic map had a number of overlapping themes; for example, "new experience" was related to mixed interaction design, whilst "unique" was related to satisfaction. The initial themes were then reviewed using the key questions proposed by Braun and Clarke [28] to question whether the themes were at the correct level of granularity and whether themes and codes needed to be moved, swapped, removed, or had new ones added.

The process of reviewing the initial themes therefore involved determining how similar codes could be combined and sorted to form the final themes [29]. The final thematic map is shown in Figure 7, in which each of the themes has been refined into several sub-themes that encompass the codes, which are not shown in this diagram.

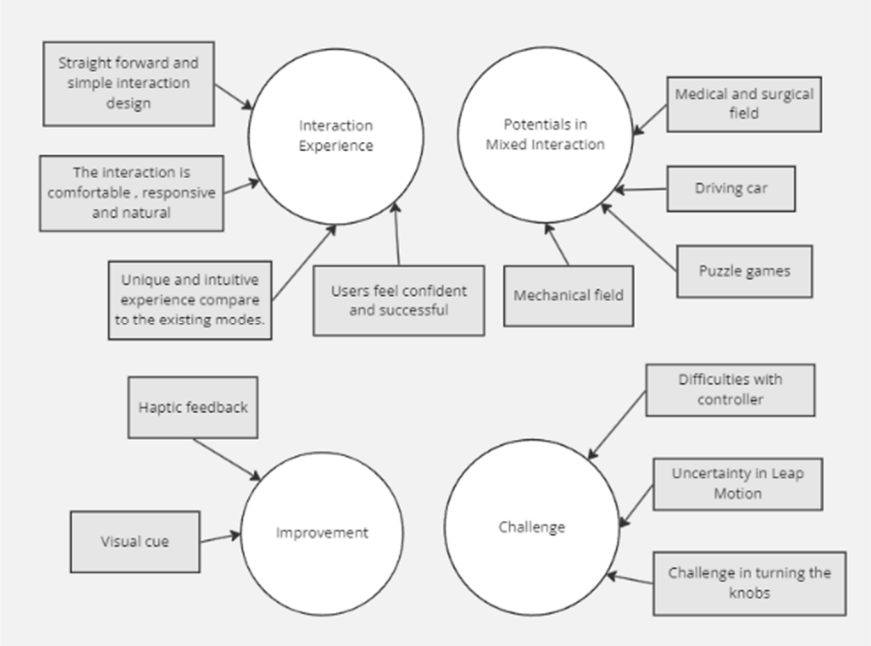

Figure 7. Final thematic map.

The four main themes of "Interaction experience", "Improvement", "Challenge", and "Potentials in mixed interaction" are described in the following sections. Whilst this final thematic map does not include the codes, the way that the data from different participants contributed to the final themes is similar to what can be seen in the initial thematic map, with concepts generally being visible in at least two of the participants' data. The corpus of data is relatively small, however, so no further attempt has been made to question the prevalence of codes that have led to the themes.

4.2. Theme 1: Interaction Experience

Participants shared their perceptions of the interaction tasks in the user study, focusing on interaction design, user experience, comparison with existing modes, and satisfaction. Frequently used descriptors included "easy, straightforward and simple", "natural and responsive", "confident and successful", "unique and intuitive", and "fun and enjoyable" (see Table 1).

Table 1. Codes and Representative Quotes for Theme 1.

Codes	Representative Quotes
Easy, straightforward, and simple	The leap interaction is very very very very good I could see and map the movement of my hand with the movement of the representation in the game, so it was easy to use, it was not hard to try to do things and try to find how to press things or to move things.
	I think it [successful task completion] was contributed because being able to have controller in one hand and able to have the other hand free, was I think much easier than using two controllers for me personally like having like rather trying to focus on two physical objects in your hand once having one hand like freely move, the other hand being able to hold controller and do stuff as well.
Natural and responsive	The interactions, they work surprisingly well to be honest and I especially for it just tracking my hand and being able to interact with the environment that so much like it's a strange experience at first, but it almost is like second nature in a way it feels and like I'm actually there it's super cool everything works really well.
Confident and successful	I didn't need to use any mental capacity to do any of the tasks. . .Pretty successful. . .like pressing levers and turn knobs are pretty easy.
	With only controllers maybe I would feel a little more in control but I feel more comfortable with the mixed interaction.
Unique and intuitive	It was very unique experience to have . . . it's already immersive sitting in the cockpit and actually touching and moving the controls. The right hand, the ghost hand and there was very connected.
Fun and enjoyable	It was quite fun actually being able to interact with like being able to use my entire hand also like having the other hand ready on the joystick to actually move it around. . .everything else felt really smooth and it was really impressive to see it actually just know my hand was there but immediately work with it.

Regarding the interaction mode, participants found the hand gesture and the controller-based modes easy and straightforward to use. Participants praised the Leap Motion device for its naturalness and responsiveness as it allows tasks like touching buttons to be completed intuitively. The handheld controllers were similarly viewed as straightforward in the usage scenario for which it was used, namely the moving of the cockpit sidestick.

Core tasks like turning knobs, pulling levers, and pressing buttons, designed to mimic real-world actions, were described as simple and confidence-boosting. Similarly, emotional responses were overwhelmingly positive, with terms like "unique", "intuitive", and "natural" frequently mentioned.

When compared with existing modes, the mixed-interaction mode stood out as both unique and intuitive, offering an engaging and user-friendly experience.

4.3. Theme 2: Challenge

Participants discussed the mental and physical workload of the interaction tasks, focusing on challenges with the controller, uncertainty with Leap Motion, and difficulties turning knobs. Commonly mentioned phrases included "no feedback", "hands get tired with the controller", and "hard to pinch the knob" (see Table 2). Interestingly these comments somewhat contradict earlier observations around ease of use.

Some participants reported fatigue when using both the controller and Leap Motion, though others indicated the opposite. The Leap Motion's lack of visual and haptic feedback also clearly posed challenges for tasks like holding a knob.

Table 2. Codes and Representative Quotes for Theme 2.

Codes	Representative Quotes
No feedback	I don't know if I'm holding a knob yet, unless is like a very significant visual indicator that it's been triggered.
	With the leap motion when you need like grabbing something it feels like everything is just weight nothing. . .. It's like if I had no touch or anything it's like trying to pretend that you're turning the knob or something.
Tiredness	It can be a little stressing or tiring, so the leap motion would need some more type of feedback or guideline of feedback or something to compliment that lack of physicality.
Difficulties	Pushing buttons it's probably the most difficult was that the stick is my hand was a bit shaky because it's in the mid air. . .The Vive controller it sometimes . . . I've played with those in the past, so it gets like real finicky trying to pick things up or my hand gets tired whereas this didn't feel like any fatigue was on my hand at all, so it was it is a lot easier to actually use my real hand.

4.4. Theme 3: Improvement

The participants were asked about the usability and possible improvements of the interaction task that was given in the user study. Participants defined the usability differently. The improvements therefore focused primarily on haptic feedback and visual cues, which unsurprisingly correlated well with the challenges theme. The delineation of these two themes was, at times, difficult, and arguably they could be merged into a single theme. However, certain comments were made that not only illustrated the challenges with mixed interaction but provided valuable feedback for potential improvements in the user study. Text such as this was generally coded into two sections with the frequently used words "No feedback" and "Distracting" being used to develop the challenges theme, whilst other words such as "haptic" and "visual cues" were used to develop the improvements theme (see Table 3).

Table 3. Codes and Representative Quotes for Theme 3.

Codes	Representative Quotes
Haptic	Some sort of haptic feedback with vibrations which may go like a bit further and selling the experience.
	Because there is no like haptic feedback for the hands so you only have to rely on visuals.
Visual cues	The response, maybe it needs a little bit of more visual flair, it suppose like visual indication of when you're interacting with your right hand. . .with the leap, it needs more hover effect, with something that indicates that you are actually interacting with. . . sometimes just seeing the thing moving or especially the buttons like are you pressing it or not, so a more visual distinct visual representation would be better.

Several participants suggested the incorporation of haptic feedback, such as vibration, to provide users with real-time information on the current state of the task or to alert users to potential obstacles. Here we can see that the sub-themes are again not entirely distinct, but in terms of thinking about how to improve the experience, this could be achieved using either improved visuals, or haptic feedback, or a combination of both. From an interaction standpoint, participants highlighted the need for visual cues, such as the hover

effect from Leap Motion, to help users keep track of their progress and provide feedback on their performance.

4.5. Theme 4: Potentials in Mixed Interaction

Participants were asked to provide suggestions of potential applications they considered applicable for adapting the mixed-interaction mode (see Table 4). Participants identified several potential applications, particularly in VR games, training, and the medical field. In VR games, especially shooting games, they saw the benefit of using Leap Motion to reduce issues like controller collisions at close distances as well as other opportunities for how mixed interaction could improve the gaming experience with more intuitive control methods that might just be more fun.

Table 4. Codes and Representative Quotes for Theme 4.

Codes	Representative Quotes
Games	First person shooters if you bring your second controller up and you'll have to hold the grip and pull the slide back but you commonly smash the two controllers together because that too close. . . and I think it's huge that will be hugely beneficial for certain types of games probably puzzle games too, where instead of having to have the controllers you can slide things around and connect wisely.
	A magic fantasy base like you have a sword in one hand and you have like your right hand doing spells like doing little finger motions to do the magic oh like doing your finger movements.
Training	Mechanical stuff like teaching people how to use certain things even like cooking like having like the whole like from one hand having like the left doing the cutting would be really cool to have like that sort of stuff.
Medical	Surgery, if you have the left hand for like surgical scissors like physical objects that you hold right in the right hand for like moving stuff around the patient and like picking objects.

In training scenarios, such as cooking or driving simulators, participants suggested that mixed interaction could provide a more realistic experience by allowing users to use both hands for different tasks. Finally, in the medical and surgical field, they saw the advantage of using mixed interaction for precision tasks, like holding surgical tools with one hand while performing other actions with the other. This suggests that a mixed-interaction mode could be beneficial for tasks requiring accuracy, such as surgery, though of course it is worth noting that none of the participants had surgical experience so this requires further investigation.

5. Discussion

The cockpit study presented above, conducted with a limited number of participants, provided initial insights into the potential benefits of the mixed-interaction mode in VR. These potential benefits have been extracted from the interview transcript data, with a representative set of comments included in the previous section. Participants reported that combining a controller with hand gestures was comfortable and allowed them to successfully complete the tasks they were set. However, these findings are preliminary and based on a small sample size. Further research is needed to confirm these results and explore the mixed-interaction mode's effectiveness across different tasks and user groups with varying levels of VR experience.

5.1. Insights and Contribution

The interview results indicated a generally positive user experience with the mixed-interaction mode. This is of particular interest as our previous quantitative evaluation [1] resulted in a more varied reception, which perhaps suggests that the mixed mode is more suitable for environments and tasks that have been specifically designed for combining controllers and gestures. Participants found the user interface intuitive and easy to use, and they appreciated the natural interaction enabled by combining Leap Motion and a controller. This approach also allowed access to both physical and VR environments without significantly compromising the user experience. Despite these positive outcomes, some participants highlighted challenges, such as confusion when switching between devices, suggesting a need for customizable designs to accommodate user preferences.

Participants emphasized the importance of appropriate visual cues to enhance task performance and reduce fatigue, particularly in mid-air interactions. The mixed-interaction mode reduced physical demand by requiring only one controller, but it also introduced the potential for increased cognitive load when alternating between devices. These findings align with research by Kang et al. [30] and underscore the importance of designing adaptive systems that balance usability with physical and cognitive demands.

This research expands on Huang et al.'s [19] findings by evaluating the mixed-interaction mode's potential beyond game-like tasks, offering insights into its usability and user satisfaction in diverse VR contexts. While the results suggest advantages such as improved accuracy and naturalness, further comparative evaluation is required to fully understand the mode's benefits and limitations.

This study highlights the potential of using a mixed-interaction mode in VR, particularly when the results are triangulated with the quantitative data from an earlier study [1], and in doing so, provides directions for future research. The findings, though preliminary, offer valuable guidance for designing VR interactions that are intuitive, efficient, and adaptable to user needs.

5.2. Limitations

Whilst this research has indicated that the mixed mode of interaction has potential to provide a positive user experience in situations that have been designed around the use of combining controllers and gestures, there are several limitations to this study. Firstly, whilst the tasks utilized were abstracted from actual cockpit training scenarios, it was still a limited generalization of interaction tasks that provided a specific context based around a relatively small number of movements. Further work would be needed to understand the full range of movements and tasks that benefit from the approach.

Another consideration related to generalizability is the relatively small and homogeneous set of participants. Much larger studies that incorporate a greater diversity of participants would be needed to understand how broadly the potential benefits of mixed interaction could be realized. Such studies could utilize quantitative measurements, similar to our previous work [1], or adopt a different approach to coding qualitative data that is suitable for a larger size, for example, eschewing a reflexive approach to coding and utilising either a codebook approach or introducing reliability measures into the coding process.

The hardware used in this study was also a limitation, as, for example, the Leap Motion is a relatively old sensor. Since this study was undertaken, commercial manufacturers of VR hardware have started providing the option to combine controllers and hand gestures (Refer to: https://developers.meta.com/horizon/documentation/unity/unity-multimodal/ (accessed on 11 June 2025)), which removes barriers and concerns over latency be-

tween multiple systems that would facilitate a better understanding of the potential of this approach.

At this stage, no attempt has been made to interpret the findings in relation to variations in the participants, which may include factors such as hand dominance, gender, age, VR experience, and so on. There is potential to triangulate existing quantitative data that includes such factors with the qualitative data in this paper; however, at this stage, this is left for future work.

5.3. Design Guidelines

This research contributes to the understanding of VR interaction design by developing guidelines to inform future design decisions. The guidelines aim to enhance user experience and usability, particularly for mixed-interaction modes.

The increasing reliance on technology highlights the need to design intuitive and user-friendly interfaces. Bossavit et al. [31] noted that design decisions greatly influence usability, and providing multiple interaction modes for object manipulation in VR is critical. Building on Huang et al. [19], who compared controller-based, gesture-based, and mixed-interaction modes, this research evaluates these modes across various VR tasks, focusing on their impacts on performance, engagement, and experience. The following design guidelines are proposed:

- Flexible Interaction Options:
  Design systems that allow for flexibility in the use of interaction modes. Rather than predetermining which hand uses the controller or Leap Motion, consider designing systems that can adapt to the user's preferences dynamically during tasks, aligning with Olmedo et al. [18], on standardized multimodal systems.
- Extended Training Phase:
  Longer training helps reduce cognitive load and improves task performance, as supported by Kang et al. [30]. Enhanced tutorials or visual cues aid users in transitioning to a mixed-interaction mode.
- Real-World Behavior Alignment:
  Simulated scenarios should mirror real-world tasks, enhancing intuitiveness and efficiency, as demonstrated in the cockpit study.
- Visual Cues:
  Use hover effects or similar indicators to compensate for the lack of haptic feedback, improving accuracy and user confidence.
- Interaction Mode Selection Based on Task Complexity:
  Match modes to task complexity; controllers suit precision tasks, while Leap Motion excels in quicker object rotation [1].
- User Comfort and Physical Demand Considerations:
  Minimize device weight and design for physical comfort to reduce fatigue. Participants in the cockpit study found the controller tiring due to its weight. Similarly, Lou [31] noted that prolonged controller use can cause hand and wrist strain.

These guidelines are directly informed by the findings of this research and represent a critical contribution to the evolving field of mixed-interaction modes in VR. While valuable, they remain propositional, requiring validation through future studies. However, they are a generalization of observations made during this research and as such can be used to help design future interaction tasks used in larger scale studies that investigate mixed interaction. Prioritizing user preferences ensures VR applications are adaptive, engaging, and aligned with evolving needs and technologies.

6. Conclusions

This research primarily contributes to the exploration of a mixed-interaction mode, combining controller and hand gesture inputs as an alternative to traditional single-mode methods for tasks requiring precision and flexibility. The findings, while insightful, are exploratory and task-specific, limited to a cockpit scenario. Generalizing these results to broader VR applications would be premature, necessitating further research to assess applicability across diverse environments. However, to some extent, it is possible to compare our findings with our previously published work and conclude that mixed interaction is better received when the environment is specifically designed for mixed interaction.

The design guidelines derived from this study provide preliminary considerations for implementing mixed-interaction modes in VR. These recommendations mark an important step forward and can be used to inform the design of tasks and environments for future studies but remain subject to refinement and validation through future research.

Future studies should investigate the mixed-interaction mode in broader contexts and for more complex tasks. Incorporating sensory inputs such as haptic feedback, voice commands, and eye-tracking could enhance immersion and interaction efficiency. Expanding participant diversity and sample size, including individuals with varying VR experience, age, and physical abilities, would yield a more comprehensive understanding of this interaction method. Addressing limitations such as learning effects and task sequencing in larger-scale studies is also crucial.

Long-term effects, including user fatigue and performance over extended use, merit exploration. Advancements in hand tracking systems and VR hardware may resolve technical challenges encountered in this study, enabling more accurate data collection and analysis.

In conclusion, this research highlights the opportunities offered by mixed-interaction modes, particularly for object-manipulation tasks in VR. While combining controller and hand gesture inputs shows promise, the findings are context-specific and require further validation. The proposed design guidelines offer a valuable foundation for future developments in VR interaction design, opening new avenues for enhancing user experiences across various fields.

Author Contributions: Conceptualization, Y.L.; methodology, Y.L. and A.M.C.; validation, Y.L.; formal analysis, Y.L.; investigation, Y.L.; resources, Y.L. and S.M.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L., A.M.C. and S.M.; supervision, A.M.C. and S.M.; project administration, A.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This research has been approved by the Auckland University of Technology Ethics Committee (AUTEC) on 13 June 2022, AUTEC Reference number 22/149.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors report there are no competing interests to declare.

References

1. Lee, Y.; Connor, A.M.; Marks, S. Mixed Interaction: Evaluating User Interactions for Object Manipulations in Virtual Space. *J. Multimodal User Interfaces* **2024**, *18*, 297–311. [CrossRef]
2. Hoang, T.; Greuter, S.; Taylor, S. An evaluation of virtual reality maintenance training for industrial hydraulic machines. In Proceedings of the 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Christchurch, New Zealand, 12–16 March 2022; pp. 573–581. [CrossRef]
3. Dai, F. (Ed.) *Virtual Reality for Industrial Applications*; Springer Science & Business Media: Berlin, Germany, 2012.
4. Hilfert, T.; König, M. Low-cost virtual reality environment for engineering and construction. *Vis. Eng.* **2016**, *4*, 1–8. [CrossRef]
5. Khundam, C. First person movement control with palm normal and hand gesture interaction in virtual reality. In Proceedings of the 2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE), Songkhla, Thailand, 22–24 July 2015; pp. 325–330.
6. Lou, X.; Li, X.A.; Hansen, P.; Du, P. Hand-adaptive user interface: Improved gestural interaction in virtual reality. *Virtual Real.* **2021**, *25*, 367–382. [CrossRef]
7. Yang, L.I.; Dong, W.U.; Huang, J.; Feng, T.I.A.N.; Hong'an, W.A.N.G.; Guozhong, D.A.I. Influence of Multi-Modality on Moving Target Selection in Virtual Reality. *Virtual Real. Intell. Hardw.* **2019**, *1*, 303–315.
8. Perret, J.; Vander Poorten, E. Touching virtual reality: A review of haptic gloves. In Proceedings of the 16th International Conference on New Actuators (ACTUATOR 2018), Bremen, Germany, 25–27 June 2018; VDE: Berlin, Germany, 2018; pp. 1–5.
9. Rahman, Y.; Asish, S.M.; Fisher, N.P.; Bruce, E.C.; Kulshreshth, A.K.; Borst, C.W. Exploring eye gaze visualization techniques for identifying distracted students in educational VR. In Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Christchurch, New Zealand, 22–26 March 2020; pp. 868–877.
10. Adhanom, I.B.; MacNeilage, P.; Folmer, E. Eye tracking in virtual reality: A broad review of applications and challenges. *Virtual Real.* **2023**, *27*, 1481–1505. [CrossRef]
11. Plopski, A.; Hirzle, T.; Norouzi, N.; Qian, L.; Bruder, G.; Langlotz, T. The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Comput. Surv.* **2022**, *55*, 1–39. [CrossRef]
12. Dhimolea, T.K.; Kaplan-Rakowski, R.; Lin, L. A systematic review of research on high-immersion virtual reality for language learning. *TechTrends* **2022**, *66*, 810–824. [CrossRef]
13. Ironsi, C.S. Investigating the use of virtual reality to improve speaking skills: Insights from students and teachers. *Smart Learn. Environ.* **2023**, *10*, 53. [CrossRef]
14. Sundén, E.; Lundgren, I.; Ynnerman, A. Hybrid Virtual Reality Touch Table—An Immersive Collaborative Platform for Public Explanatory Use of Cultural Objects and Sites. In *Eurographics Workshop on Graphics and Cultural Heritage*; Eurographics Association: Goslar, Germany, 2017. [CrossRef]
15. Wolf, E.; Klüber, S.; Zimmerer, C.; Lugrin, J.-L.; Latoschik, M.E. "Paint That Object Yellow": Multimodal Interaction to Enhance Creativity During Design Tasks in VR. In Proceedings of the 2019 International Conference on Multimodal Interaction, Suzhou, China, 14–18 October 2019; pp. 195–204. [CrossRef]
16. Hepperle, D.; Weiß, Y.; Siess, A.; Wölfel, M. 2D, 3D, or Speech? A Case Study on Which User Interface is Preferable for What Kind of Object Interaction in Immersive Virtual Reality. *Comput. Graph.* **2019**, *82*, 321–331. [CrossRef]
17. Wagner, J.; Stuerzlinger, W.; Nedel, L. Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions for Data Manipulation in Immersive Analytics. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 2513–2523. [CrossRef] [PubMed]
18. Olmedo, H.; Escudero, D.; Cardeñoso, V. Multimodal Interaction with Virtual Worlds XMMVR: eXtensible Language for MultiModal Interaction with Virtual Reality Worlds. *J. Multimodal User Interfaces* **2015**, *9*, 153–172. [CrossRef]
19. Huang, Y.J.; Liu, K.Y.; Lee, S.S.; Yeh, I.C. Evaluation of a Hybrid of Hand Gesture and Controller Inputs in Virtual Reality. *Int. J. Hum.–Comput. Interact.* **2021**, *37*, 169–180. [CrossRef]
20. Ionescu, D.; Ionescu, B.; Gadea, C.; Islam, S. A Multimodal Interaction Method that Combines Gestures and Physical Game Controllers. In Proceedings of the 20th International Conference on Computer Communications and Networks (ICCCN), Lahaina, HI, USA, 31 July–4 August 2011; pp. 1–6.
21. Berg, L.P.; Vance, J.M. Industry Use of Virtual Reality in Product Design and Manufacturing: A Survey. *Virtual Real.* **2017**, *21*, 1–17. [CrossRef]
22. Faulkner, L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behav. Res. Methods Instrum. Comput.* **2003**, *35*, 379–383. [CrossRef]
23. Hwang, W.; Salvendy, G. Number of people required for usability evaluation: The 10 ± 2 rule. *Commun. ACM* **2010**, *53*, 130–133. [CrossRef]
24. Schmettow, M. Sample size in usability studies. *Commun. ACM* **2012**, *55*, 64–70. [CrossRef]
25. Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
26. Byrne, D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Qual. Quant.* **2022**, *56*, 1391–1412. [CrossRef]

27. Sundler, A.J.; Lindberg, E.; Nilsson, C.; Palmér, L. Qualitative Thematic Analysis Based on Descriptive Phenomenology. *Nurs. Open* **2019**, *6*, 733–739. [CrossRef]
28. Braun, V.; Clarke, V. Thematic analysis. In *APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs*; Cooper, H., Camic, P.M., Long, D.L., Panter, A.T., Rindskopf, D., Sher, K.J., Eds.; American Psychological Association: Washington, DC, USA, 2012; pp. 57–71. [CrossRef]
29. Nowell, L.S.; Norris, J.M.; White, D.E.; Moules, N.J. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *Int. J. Qual. Methods* **2017**, *16*, 1609406917733847. [CrossRef]
30. Kang, H.J.; Shin, J.; Ponto, K. A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. In Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, USA, 22–26 March 2020; pp. 275–284. [CrossRef]
31. Bossavit, B.; Marzo, A.; Ardaiz, O.; De Cerio, L.D.; Pina, A. Design Choices and Their Implications for 3D Mid-Air Manipulation Techniques. *Presence Teleoperators Virtual Environ.* **2014**, *23*, 377–392. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.