

Article

Spatial Ability Performance in Interior Design and Architecture: Comparison of Static and Virtual Reality Modes

Ji Young Cho ^{1,*}  and Joori Suh ² ¹ Department of Housing and Interior Design, Kyung Hee University, Seoul 02447, Republic of Korea² School of Architecture and Interior Design, University of Cincinnati, Cincinnati, OH 45221-0016, USA; joori.suh@uc.edu

* Correspondence: jcho@khu.ac.kr

Abstract: Recent advancements in virtual reality (VR) technology have enabled its integration into learning diverse aspects of spatial components and relationships in the field of spatial design, as well as designing, communicating, collaborating, and managing complex building projects. With the growing interest in incorporating VR technology in spatial design, examining whether people understand, perceive, and perform spatial tasks in the same way in VR as they do in static modes is essential. Thus, the purpose of this study was to compare spatial ability performance in a conventional static paper–desktop mode and an interactive VR mode. Thirty students completed the Architecture and Interior Design Domain–Specific Spatial Ability Test in both modes. Their visual cognitive style was measured with the Object–Spatial Imagery Questionnaire, and their responses to the usability of the VR mode were analyzed. The results revealed: (a) significant difference in performance between static and VR modes, including better performance in three spatial visualization subconstructs in static mode than in VR; (b) no gender difference in VR mode; (c) a tendency of spatial visualizers to benefit from VR mode; and (d) a tendency of people with high spatial ability to be more susceptible to test mode. Overall, the results contribute to expanding our understanding of spatial ability performance in different test modes and provide insights concerning the integration of VR into the development of spatial ability tools and education.



Citation: Cho, J.Y.; Suh, J. Spatial Ability Performance in Interior Design and Architecture: Comparison of Static and Virtual Reality Modes. *Buildings* **2023**, *13*, 3128. <https://doi.org/10.3390/buildings13123128>

Academic Editors: Oleg Kapliński, Wojciech Bonenberg, Agata Bonenberg and Jan Słyk

Received: 10 November 2023
Revised: 6 December 2023
Accepted: 15 December 2023
Published: 17 December 2023



Copyright: © 2023 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/>).

Keywords: design education; interactive; interior design; spatial ability; test mode; virtual reality

1. Introduction

Spatial ability is one of the key cognitive abilities required to perform design tasks in the field of interior design and architecture. In creating new spaces and buildings, designers exercise their imaginations, visualizing the composition and volumetric relationships among spatial components. In addition, to communicate spatial properties and their relationships, visual representation methods involving two-dimensional (2D) and three-dimensional (3D) drawings are used among designers, engineers, and constructors. Thus, two of the primary skills imbedded in education in interior design and architecture are spatial awareness and versatility in communicating spatial information.

With the development of virtual reality (VR) technology, tremendous changes have taken place in communicating spatial objects and components, especially in education in interior design and architecture. The high sense of presence and the sense of immersion with the VR technology allow students to learn content more vividly in a virtual environment (VE). Furthermore, various spatial interactions, such as teleporting and manipulating objects in VE, are now possible. Real-time modification and iterations of design can be made within VE, and collaboration among designers and multiple stakeholders from multiple locations is also possible in VR. Researchers have shown the benefit of using VR technology in design studios, including facilitating creativity in the student design process [1], enhanced perception of solid–void characteristics of a model [2], and motivation [1].

With growing interest in incorporating VR technology in various aspects of the spatial design process, the following question must be asked: Do people understand, perceive, and perform spatial tasks in the same way in VR as they do in static mode? Since the measurement of spatial ability innately involves 3D space, the potential of VR technology in spatial ability has been investigated by researchers [3–5]. Specifically, Lochhead et al. (2022) investigated differences in mental rotation (MR) performance in 2D images versus 3D stimuli, both in VR mode, and found better participant performance in 3D stimuli [3]. In measuring spatial ability, however, examining what type of spatial ability is measured, what test mode must be used, and how the spatial ability questions should be presented is important since they may impact spatial ability performance depending on test mode. Moreover, recent neuroscience studies revealed that responses to the same task in virtual reality and desktop mode vary considerably [6]. Different parts of the brain are activated, and since the networks of the brain regions are dissimilar, deep understanding of spatial ability performance in interactive VR versus conventional static mode is necessary. Whether performance in VR mode resembles that in the original test modes has rarely been reported; consequently, researchers have pointed out the strong need to understand spatial ability performance in various test modes [3].

Thus, the purpose of this study was to identify the characteristics of spatial ability performance in interactive VR mode in comparison with that in a static paper–desktop mode, using the Architecture and Interior Design Spatial Ability Test (AISAT), developed to measure domain-specific spatial ability in the field of architecture and interior design. Test data and self-reported information were analyzed to examine potential factors related to the test mode effect on the spatial ability test. Performance in the two modes was compared using a statistical program. Two modes were examined; therefore, any impact that might have occurred as a result of the order of performing the tasks was examined. Performance in relation to individual differences, such as gender and one’s visual cognitive style were also examined. The Object–Spatial Imagery Questionnaire (OSIQ) [7] was administered to identify participants’ tendency toward a particular visual cognitive style and its relationship with performance in the two test modes of the AISAT. Performance in VR mode by high vs. low spatial ability score in static mode was examined to identify who benefits from using the VR mode. Moreover, a usability survey was also included to evaluate the tool used for the VR version of the AISAT.

The main research questions were:

- Is spatial ability performance on the AISAT similar or different in the static mode and VR mode?
- Does the order of taking the tests in a particular mode (VR mode first vs. static mode first) have an impact on performance?
- Does gender difference exist in spatial ability performance in the two modes?
- Does visual cognitive style matter in spatial ability performance?
- Do individuals with low spatial ability benefit from the VR test mode?
- How do participants evaluate the usability of the VR AISAT?

2. Background

2.1. Spatial Ability Measurement

Spatial ability is the proficiency needed to understand spatial conditions among objects and environment and problem-solving capacity regarding spatial information. It is defined as “skill in representing, transforming, generating, and recalling symbolic, nonlinguistic information” ([8], p. 1482) and “the ability to visualize, manipulate and interrelate real or imaginary configurations in space” ([9], p. 3). It is a significant cognitive capacity required in everyday problem-solving, including wayfinding, design, technical drawings, and graphic visualization [10]. Existing instruments for measuring SA were developed mainly in the 1970s by researchers in the fields of cognitive and developmental psychology with the primary purpose of predicting successful academic or professional performance. Spatial ability can be categorized as ability in mental rotation, spatial visualization (SV),

and spatial perception [8]. Some of the well-known and validated tools include mental rotation, developed in 1978 [11]; paper folding, introduced in 1976 [12]; and a test for perspective-taking and spatial orientation, introduced in 2004 [13]. These tools were originally developed in static mode in a paper-and-pencil version; later some of them were converted from static to an interactive computer version or interactive VR version by other researchers [3]. In the domain of architecture and interior design, examples of tools include the 3D Ability Test (3DAT), developed in 2010 [14]; and the Urban Layout Test, Indoor Perspective Test, and Packing Test, introduced in 2021 [15].

In the domain of interior design and architecture, spatial ability can be understood as “the mental manipulative skills required to perform mental processes such as the rotation of objects, the understanding of how objects appear in different positions, and the conceptualization of how objects relate to each other in space” ([16], p. 2). Unlike in other domains, practitioners of interior design and architecture deal with spatial relationships of human-scale or large environment-scale components as well as the translation of 2D drawings into 3D representation of spatial components within habitable spaces. In architecture and interior design, designers handle spaces that are more specific and concrete as well as abstract. The AISAT was developed with the intention to measure domain-specific spatial ability in architecture and interior design, different from general spatial ability; its reliability and validity were investigated and reported in a prior study [17]. In the AISAT, instead of using small abstract objects typical in general spatial ability tools, large-scale spatial forms relevant to environmental design were used to test mental rotation and spatial visualization. The predictive potential of the developed AISAT for design performance and creativity has been investigated and reported [18,19].

2.2. Nature of VR and Its Application for Spatial Ability Measurement

VR technology is a medium, comprising computer simulations, that allows the viewer to experience full immersion within the simulations [20]. In a VE, spatial representation as well as human–object interaction becomes more realistic, allowing the user to experience the sense of the scale of the environment. VR simulation also makes spatial objects appear closer to their real form with an increased sense of depth and three-dimensionality. With the provision of constantly changing perspective views, the viewer can experience more extensive perception of the spatial forms from many different angles, which is impossible with a static mode.

VR has the effect of creating a sense of immersion and presence as users interact with visual stimuli in a 3D space. Sense of presence and fidelity are key characteristics of VR, the former being the degree to which users feel that they are present in the VE—or that virtual objects are present—without being aware of their virtual presence. Fidelity refers to the degree to which users experience interactions within the VR as if they were interacting in the real world [21]. A review of previous studies has shown that VR experiences help to expand cognitive abilities and improve spatial perception along with increasing immersion, confidence, and motivation [22]. Since the development of VR technology, the design process, modifications of design, and the way designers collaborate have been influenced by VR technology across domains. Along with more extended application of the immersive experience encompassing the virtual and physical worlds, many new approaches to design and management methods using the eXtended Realty (XR) platform have been introduced. For example, Banfi et al. (2019) proposed informative models for architectural heritage using the scan-to-building information modeling (BIM) process for VR and augmented reality (AR) that allows an immersive experience of built heritage to discover hidden historical values [23]. Other examples include a platform for immersive and interactive experiences for cultural heritage [24] and a platform for reducing the rework of design changes in BIM using VR and AR [25]. Alizadehsalehi and Yitmen (2021) proposed a digital twin-based automated construction progress monitoring system that allows a user to collaborate to create, analyze, manage, and visualize construction progress using BIM and XR technologies [26]. With the potential of integrating artificial intelligence and

machine learning within the VR-based educational process, including design exploration and communication, deeply understanding how designers' spatial skills differ in the virtual and physical interactions has become more crucial.

One example of the VR version of spatial ability-measuring tools is the Immersive Mental Rotations Test by Lochhead et al. (2022), who fit the mental rotation test originally developed by Vandenberg and Kuse in 1978 to VR head mounted display (HMD)—Oculus Quest [3]. They did not, however, compare performance in VR and static modes. Instead, they compared performance of 2D images versus 3D objects, both in VR mode and found that participants performed better in terms of accuracy and speed in “stereo 3D stimuli than with 2D images of those stimuli” (p. 1). In addition, they pointed out that conventional paper-and-pencil MR measures two things—one is the ability to visualize 2D to 3D, and the other is the ability to rotate that 3D—and the immersive VR can alleviate the cognitive load of the first stage by providing 3D objects in VR. Finally, they argued that with the emergence of new technologies such as VR, new questions on utility, interface, application, and users in VE must be asked.

2.3. The Development and Test Modes of AISAT

The original version of the AISAT was available only in static mode—on paper or a computer screen. The VR version of the AISAT was developed to explore the potential of implementing VR technology in measuring spatial ability in the domain of interior design and architecture. Implementation of VR technologies using large-scale simulated environments may have benefit in terms of providing a better sense of scale and enhancing the understanding of spatial relationships among the components presented at an environmental scale, but each mode has its own strengths and weaknesses (See Table 1).

Table 1. Strengths and weaknesses of two modes of the AISAT.

	Static Paper—Computer Mode AISAT	Interactive VR Mode AISAT
Strengths	Display of the entire space is possible. Convenience of preparing for the test	Realistic simulated view of the space Dynamically coordinated perspective views according to the viewer's positions High resolution images Interaction within the space (e.g., teleporting, changing direction of views, and rotating objects using controller)
Weakness	Printing of each page required. Low resolution Limited interaction Lack of realistic experience	Cumbersome and heavy HMD Requires longer preparation time and training for familiarity Discomfort (e.g., dizziness)

The paper-and-pencil version has benefits in ease of use and distribution. It is also economical since no specific equipment is needed. The structure of each question entails one question and four multiple choices. One can look at all options and choose one answer. Its weakness, however, lies in the resolution of question images and details due to the difficulty in conveying realistic and high-resolution images on a small piece of paper. In addition, it is not sustainable in the long run since questions must be printed each time the test is given. As a result of the low resolution and limitation in detail on the paper version, the spatial visualization part of the AISAT was prepared for the computer. Questions were prepared in an online survey format (qualtrics.com, accessed on 1 June 2021), and the link was shared with participants. In the paper–desktop mode, all questions and answer options were provided in static mode. In contrast, in the VR version of the AISAT, all questions and answer options appear in interactive mode.

The VR version of the AISAT was developed using the questions from the paper–desktop test mode. Three-D modelling was created using SketchUp 2020.2 and Rhino 6 software. The 3D modelling was converted to a fbx file, an extension that can be used on the Unity game engine. To test the AISAT in VR, user interface and user experience design

of the virtual platform were required to induce appropriate behaviors and experiences for users. In the VR version of the AISAT, test takers can look around the environment in question, rotate the objects in the 3D test space (on the MR test), look at 3D simulations from diverse angles by shifting their bodies and by using the controller, and move within the 3D simulation space using the controller and teleport function. Unity, a game engine, was used for the software development and HTC VIVE was used as the HMD for the VR experience.

2.4. Spatial Ability, Gender, and VR

Gender difference has been frequently reported, particularly men's outperformance in MR (e.g., [8,27–29]). In our prior studies, in static mode AISAT, no gender difference was observed in any of the SV subconstructs [18], but gender difference was found in MR [17]. Reporting no gender difference in VR can be found in literature. For example, Park (2009) found that in VR, gender did not affect spatial ability in MR and SV [30]. Samsudin et al. (2011) also found that in MR, no significant gender difference occurred in the VR version [31]. Thus, we assumed that no gender difference would be observed in the VR mode of the AISAT.

2.5. Spatial Ability and Visual Cognitive Style

“Visual cognitive style” is a term that refers to “an individual's cognitive capacity to process information in terms of object and spatial images” ([32], p. 198). The intent of the OSIQ, developed by Blajenkova et al. (2006), was to measure tendencies in visual cognitive style. One's visual cognitive style can be divided into object visualization style and spatial visualization style; those processing styles were called object visualizers and spatial visualizers, respectively [7]. Object visualizers tend to be good at processing object information, such as colorful and pictorial images of objects, and spatial visualizers tend to excel in processing the spatial relationships of objects. The OSIQ was used to identify its relationship with performance in the two modes of the AISAT and any differences in advantaging in a certain mode. The literature has shown that spatial visualizers who have high spatial scores tend to have higher spatial ability than object visualizers [32,33].

2.6. Spatial Ability Level and Benefit of VR

A review of prior studies showed the potential benefit of VR in people with different levels of spatial ability. For example, in a comparison of learning performance in PowerPoint slides versus desktop VR in people with high spatial ability (HSA) and low spatial ability (LSA) [34], the performance of people with LSA increased in VR but that of people with HSA did not, indicating that VR benefits people with LSA. An electroencephalography (EEG) study by Sun et al. (2019) showed that VR helped reduce the cognitive load for people with LSA [35]. In fact, some researchers have argued for a compensator hypothesis [36,37], meaning that those with LSA benefit since VR can alleviate cognitive load, but others have argued that those with HSA benefit under the enhancer hypothesis [37]. Investigating those who benefit in VR mode according to one's spatial ability is worthwhile.

3. Research Method

The primary focus of this research was to understand performance in a static mode and the VR mode of the AISAT. Based on key findings from our previous research and literature review, we hypothesized the following for the current research:

H1: Performance in the static mode of AISAT and the VR mode of AISAT will be similar.

H2: The order of taking the tests (VR mode first vs. static mode first) will not influence spatial ability performance.

H3: No gender difference will be observed in the performance of the VR AISAT.

H4: Spatial visualizers in visual cognitive style will perform better in VR mode than object visualizers.

H5: Those with LSA will benefit from VR mode more than those with HSA.

Figure 1 shows a graphic summary of the research framework and hypothesis of the current study.

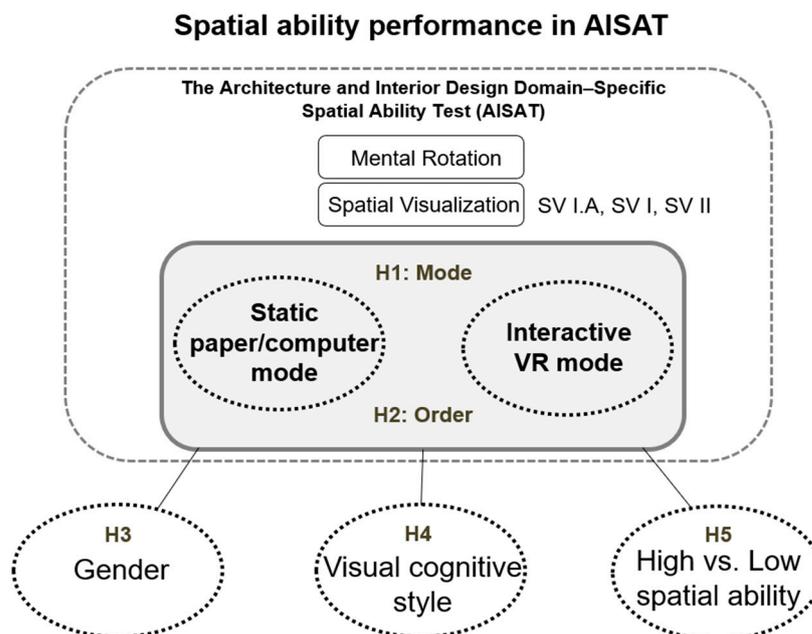


Figure 1. Research framework and hypothesis.

3.1. Procedure

Data were collected from June 2021 to June 2022. A total of 46 college students (7 men and 39 women) enrolled in an interior design program at one university in Seoul, Republic of Korea signed up to participate. The dominance of women in interior design programs is commonly observed in diverse countries, such as in the US [38] and UK [39]. An official report in 2020 indicated that the proportion of female interior design students was 83% between 2014 to 2018 [39]. Although no official report on the proportion of gender in interior design programs is available in Korea, one may reasonably say the dominance of females in the current research represents the tendency in the discipline. Participants were randomly assigned to two groups, and each group performed the two modes of tests in a different order in a one-week period to ensure no impact from the sequence of solving the tasks.

For the static mode, participants sat in one classroom together and completed the task individually. One researcher administered the AISAT. Fourteen minutes were given for the static mode; including the tutorial and sample questions, about twenty minutes were used. For the VR mode, each participant performed the VR AISAT individually. Participants went to another room, where they stood while wearing HMD and completed the VR version. The questions themselves required 23 min. However, including the tutorial and sample questions, about 40 min were needed. The rationale behind the time allotted derived from the results of the pilot study, in which the 10 participants taking the test in VR mode needed more time to attain a score similar to what they earned in static mode. Thus, we determined the time allotted would be what participants needed to earn similar scores in both modes. Once finished, participants returned to a waiting area and completed the usability test questionnaire.

3.2. Instrument

Participants completed two modes of the AISAT, the OSIQ, a demographic information questionnaire, and the usability test questionnaire for the VR version.

3.2.1. AISAT

The constructs of the static mode and the VR mode of the AISAT and time given for each construct of questions appear in Table 2.

Table 2. AISAT constructs in two modes.

Subconstruct of Spatial Ability	Description	Static Mode		Interactive VR	
		Number of Questions	Time	Number of Questions	Time
Mental rotation	MR The ability “to mentally rotate 3D spatial forms and visualize them rapidly”	7	4 min	7	2 min
Spatial visualization	SV I.A (2D → 3D in abstract information) The ability “to read 2D information and convert it to 3D and find correct location for the correct viewpoint with respect to the orientation of test takers’ own body”	7	2.5 min	7	4 min
	SV I (2D → 3D) The ability “to read 2D spatial information (e.g., floor plan drawing), expand it into volumetric forms, mentally proceeding through various possibilities to locate the correct 3D exterior or interior shapes of the building”	10	5 min	10	10 min
	SV II (3D → 2D) The ability “to read 3D volumetric information, compress complex volumetric information, convert it to 2D information, and find the correct 2D floor plan”	5	2.5 min	5	7 min
Total		29	14 min	29	23 min

Sample questions for each construct in the static mode and the VR mode of the AISAT appear in Table 3.

The AISAT comprises of two main constructs—MR and SV—and three subconstructs within SV: SV I.A, SV I, and SV II. MR measures the ability to mentally rotate 3D spatial forms and visualize them rapidly. The questions in MR comprise “abstract environmental information, such as vertical dividers and horizontal elements familiar from an isometric architectural perspective and furniture drawings” ([17], p. 16). SV I.A and SV I measure the ability to translate 2D information into 3D information. The difference between SV I.A and SV I is that SV I.A uses “abstract environmental information with appropriate human-eye level” ([17], p. 16) while SV I uses more concrete information, such as a 2D floor plan and 3D interior/architectural perspectives. SV II measures the ability to translate 3D information into 2D information with concrete spatial elements, such as 3D interior/architectural perspectives and 2D floor plan.

The AISAT was prepared as two sets (Set A and Set B). Each set has the same number of questions and participants were randomly assigned one of the sets. Within each set, the order of the questions was randomized to counter for possible learning and carry-over effects. The VR mode provides a tutorial on how to use the button in the system and a few sample example questions for users, allowing them to familiarize themselves with the system. Onscreen text as well as verbal instructions were delivered by the research moderator. After the introduction and tutorial sessions, an overview of the entire procedure was introduced before they began the test. In addition, the moderator asked participants whether they understood how to use the equipment, had any issue regarding the HMD, or had any questions so that all participants could begin with sufficient understanding of the test tool and feel comfortable.

Table 3. Sample questions from the AISAT static mode and VR mode.

Subconstruct of Spatial Ability	Static Mode (Paper-Computer)	Interactive Mode (VR)
MR		
	Paper version	
	Users are asked to imagine rotating the objects along a vertical axis, and find two images that are the same as the model aside from its orientation.	Users are asked to rotate an object to match the provided object in question.
SV I.A		
	Paper version	
	Users are asked to look at the 2D plan, imagine that they are standing inside the circle looking in the direction of the arrow, and find the correct perspective among options.	Users are asked to look at 2D drawing and find the correct 3D abstract models of spatial information
SV		
SV I	Computer version	
	Users are asked to look at the floor plan, imagine its three-dimensional condition, and select the correct 3D perspective among the options.	Users are asked to look at a 2D drawing and select the correct 3D architectural model among options.
SV II		
	Computer version	
	Users are asked to look at the perspective image, imagine its two-dimensional condition, and select the correct 2D drawing among options.	Users are asked to explore 3D space and select correct 2D drawing among options.

3.2.2. OSIQ

After completion of the VR mode of AISAT, participants' visual cognitive style was measured using the OSIQ. The original OSIQ is a self-reported measurement consisting of 30 statements with responses recorded on a 7-point Likert scale, ranging from 1 (totally disagree) to 7 (totally agree). In this study a simple version comprising 10 statements was used (five questions for spatial visualizer and five for object visualizer).

3.2.3. Demographic Information Questionnaire

The questionnaire on demographic information included gender, age, year in design major, VR experience, and experience with VR games.

3.2.4. Usability Test Questionnaire

The usability test comprised 34 statements with responses recorded on a 5-point Likert scale, ranging from 1 (totally disagree) to 5 (totally agree). Table 4 shows the categories and sample statements appearing on questionnaire for the usability of the VR AISAT.

Table 4. Questionnaire categories and sample statements.

Item	Sample Statements	Number of Questions
Usability	I'm more comfortable with a traditional paper test. I am more comfortable with the virtual reality method. The controls in the program are easy to use.	7
Easy to understand tutorial	The tutorial for the first type, MR, helped me understand how to solve the problem.	4
Task performance	I understood exactly what the goal of the task was.	4
Exploration and navigation	The menus of the interface were easy to navigate.	4
Satisfaction	Doing the virtual AISAT is a valuable experience for me.	5
Immersion	It was immersive while solving the virtual AISAT.	3
Discomfort	I felt dizzy while using Virtual AISAT.	7
Open-ended questions	Please feel free to share your thoughts on solving the problems. Tell us how we can improve AISAT-VR	2

3.3. Participants

Participants were recruited from students enrolled in an interior design program at one university in Seoul, Republic of Korea. Students voluntarily participated in the study and received a gift card valued at \$20 as compensation for their time. Data were collected ethically, following approved Institutional Review Board protocol at the university.

Originally, a total of 46 students signed up for the research. However, as a result of COVID-19, only 30 participants (12 participants who performed the VR mode first and 18 participants who performed the static mode first) completed both modes, which were used for the comparison analysis.

4. Results

In the static mode, MR and SV I.A questions were in pencil-and-paper format, and SV I and SV II questions were on the computer screen display format. Scores for the paper version were recorded by the researchers. In the computer screen display version, their responses were recorded in Qualtrics. On the VR AISAT, participants' selection of answers and response time were recorded on a desktop, and correctness was automatically measured. In this research, only the accuracy score was measured and analyzed since each test had a time limit, and thus the response time was not analyzed.

4.1. Demographic Characteristics Analysis

Participants included 6 men and 24 women (average age = 22.36). A total of 3 of them were sophomores, 21 were juniors, and 6 were seniors. A total of 46.7% of them had no prior experience in VR, 26.7% had one, 20% had two, and 6.7% had four VR experiences. Regarding VR game experience, 53.3% answered no experience; 30%, one; 6.7%, two, and 10%, three times. Approximately half the participants had no prior experience in either VR or VR games. When analyzing any difference in performance due to participants' past VR experience and VR game experience, no statistical difference was observed. So, both modes of the AISAT could be used as a measuring tool without past VR-related experiences.

4.2. Performance Comparison by Test Mode (Test of H1)

For comparison purposes, scores have been converted based on a maximum score of 100. When comparing participants' performance in static mode and VR mode, overall, they performed better on SV in static mode but better on MR in VR mode. When examining the performance score in each subconstruct, the pattern between static mode and VR showed differences; in static mode, scores were high in SV I.A, MR, SV I, and SV II in that order; scores in VR mode were high in MR, SV I.A, SV I, and SV II in that order. Overall, participants performed better in SV I.A and MR but worse in SV I and SV II. The average score of each mode was 66.78 (SD = 14.59) in static mode and 56.09 (SD = 12.80) in VR mode out of 100. See Figure 2.

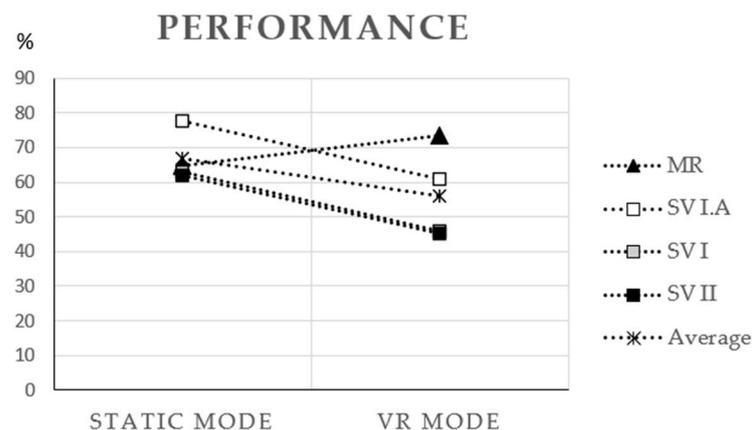


Figure 2. Observed scores (% out of 100) by static and VR mode.

Examination of statistical differences using a paired sample *t*-test revealed significant differences in three SV items with better performance in static mode. A statistically significant difference was found in SV 1.A performance between static mode ($M = 77.62$, $SD = 14.86$) and VR mode ($M = 60.95$, $SD = 26.51$) with $t(29) = 3.22$, $p = 0.003$, $d = 0.78$; SV I performance between static mode ($M = 63.00$, $SD = 20.03$) and VR mode ($M = 46.00$, $SD = 15.67$) with $t(29) = 4.39$, $p = 0.000$, $d = 0.95$; SV II performance between static mode ($M = 62.00$, $SD = 28.45$) and VR mode ($M = 45.33$, $SD = 28.74$), with $t(29) = 4.09$, $p = 0.000$, $d = 0.58$; and the average of all subconstruct score between static mode ($M = 66.78$, $SD = 14.59$) and VR ($M = 56.09$, $SD = 12.80$), with $t(29) = 4.61$, $p = 0.000$, $d = 0.78$. Results indicated that participants performed significantly better in the three SV tests in static mode. The three SV subconstructs measure the following spatial visualization abilities: SV I.A (2D to 3D in abstract information), SV I (2D to 3D in concrete information), and SV II (3D to 2D in concrete information). The dimensionality crossing between 2D and 3D in SV seemed to impact performance in VR more. The SD of the average score of static mode was larger than that in VR, indicating larger individual differences in static mode performance than in VR.

4.3. Performance Comparison by the Order of Mode (Test of H2)

Of 30 participants, 12 performed VR mode first, and 18 performed static mode first. When comparing any difference by the order, the independent *t*-test result shows no statistically significant difference in the performance by the order except SV II in VR, meaning the order of test mode did not bring effect on performance in most cases but it did on performance of VR SV II. Those who completed VR first performed better in VR SV II than those who did static mode first. See Figure 3.

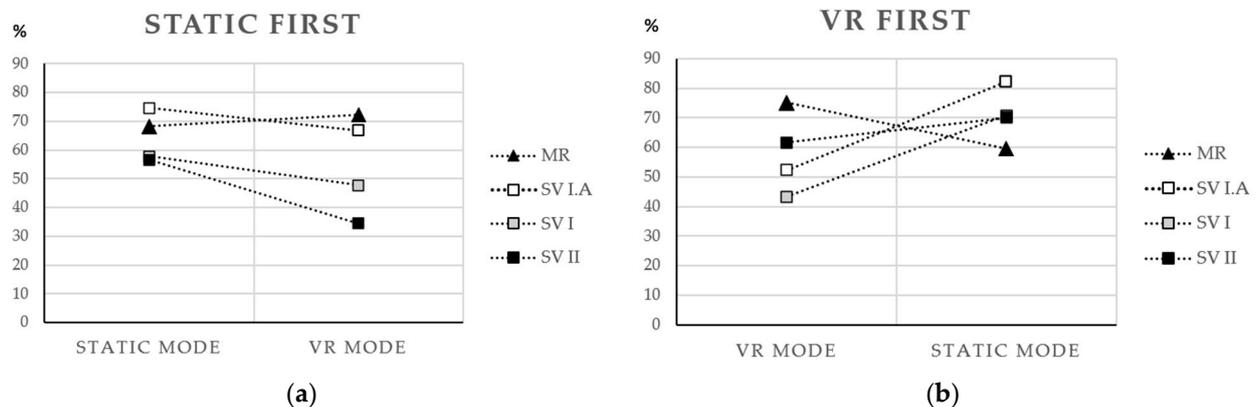


Figure 3. (a) Observed scores by static first group; (b) Observed scores by VR first group.

When calculating the difference between the score in static mode and that in VR mode and then conducting an independent *t*-test using that score, significant differences were found in SV I.A and SV I. A significant difference occurred in the SV I.A scores between static first group ($M = 7.94$, $SD = 25.09$) and VR first group ($M = 29.76$, $SD = 28.87$) with $t(28) = 2.20$, $p = 0.036$ and in the SV I scores between static first group ($M = 10.00$, $SD = 19.40$) and VR first group ($M = 27.50$, $SD = 20.01$); $t(28) = 2.39$, $p = 0.024$. In the two subconstructs, those who completed VR first had larger difference in that score, meaning the VR first group performed better in static mode later compared to static first group, indicating the potential VR training effect on performance.

4.4. Performance Comparison by Gender (Test of H3)

The participants included 6 men and 24 women. Although the proportion is somewhat unbalanced for comparison, any difference according to gender was examined to confirm that VR AISAT does not measure general SA but domain-specific SA. Overall, men performed better than women in all items except MR in VR. When examining any difference using an independent *t*-test, significant difference was observed only in MR in static mode between men ($M = 88.10$, $SD = 14.05$) and women ($M = 58.93$, $SD = 23.95$) with $t(28) = -2.84$, $p = 0.008$ and in the average score of static mode between men ($M = 79.31$, $SD = 9.75$) and women ($M = 63.65$, $SD = 14.03$) with $t(28) = -2.57$, $p = 0.016$. This result indicates that performance on most components of the AISAT in either static or interactive VR modes is not influenced by gender.

Figure 4 shows that compared with men, for whom all scores in VR mode decreased, women had higher MR scores in VR mode, indicating for women VR mode was much easier than the static mode. The reason for better performance in interactive MR seemed to be the reduced cognitive load in VR mode since participants were able to see the rotated view of the shape quickly by clicking the button until it matched the shape suggested in the question. In addition, for women, the slopes of the graph of three SV subconstructs tended to be much gentler than for men, meaning that women did not receive much impact in performance in VR compared to men.

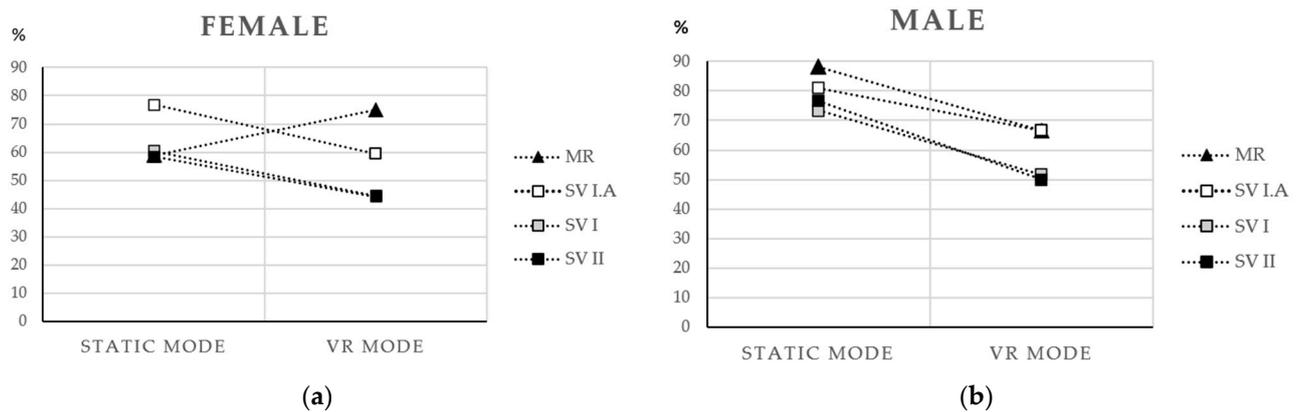


Figure 4. (a) Observed scores by female; (b) Observed scores by male.

4.5. Performance Comparison by the Visual Cognitive Style (Test of H4)

Based on the OSIQ score, participants were categorized into two groups: a spatial visualizer group ($N = 13$), who had high spatial visualization scores, and an object visualizer group ($N = 13$), who had high object visualization score. Overall, the spatial visualizer group performed better than the object visualizer group. The independent t -test result showed that significant difference was observed in static SV II, VR SV I, and VR SV II. Significant difference was observed in SV II between the spatial visualizer group ($M = 72.31$, $SD = 23.86$) and the object visualizer group ($M = 46.15$, $SD = 28.73$) with $t(24) = 2.53$, $p = 0.019$; in VR SV I between the spatial visualizer group ($M = 53.08$, $SD = 14.94$) and the object visualizer group ($M = 36.92$, $SD = 14.37$) with $t(24) = 2.81$, $p = 0.01$; and in VR SV II between the spatial visualizer group ($M = 56.92$, $SD = 25.62$) and the object visualizer group ($M = 30.77$, $SD = 29.00$) with $t(24) = 2.44$, $p = 0.023$. In the three items, the spatial visualizer group performed significantly better than the object visualizer group. The better performance in two of SV subconstructs in VR indicated the potential of spatial visualizers' benefit in VR. See Figure 5.

4.6. Performance Comparison by High vs. Low Spatial Ability (Test of H5)

Participants were divided into two groups based on the sums of their scores on performance in static mode: high spatial ability group (HSA, $N = 10$, static sum is 22 or more out of 29) and low spatial ability group (LSA, $N = 8$, static sum is 17 or less out of 29); their performance was compared using an independent t -test. In the VR mode, HSA performed better than the LSA in SV I and SV II, meaning that people with high spatial ability perform well in SV I and SV II in VR but performance in MR and SV I.A in VR mode was not influenced by spatial ability. See Figure 6.

Examination of the statistical differences of performance in static mode and VR mode using a paired sample t -test within each group revealed different patterns in people with HSA and LSA. A significant difference was observed in three SV items for HSA with better performance in static mode. This indicated that those with HSA tended to be more influenced by and sensitive to the test mode than those with LSA. However, for LSA most differences disappeared, but only difference in SV I was observed; and the MR score of LSA was even higher in VR mode than that in static mode although statistically not significant. In addition, the degree of decline from SV scores in static to that in VR mode was smaller in those with LSA. This indicated the potential benefit of VR for LSA. For HSA, statistically significant difference was found in SV 1.A between static mode ($M = 87.14$, $SD = 8.11$) and VR mode ($M = 57.14$, $SD = 25.20$) with $t(9) = 3.37$, $p = 0.008$, $d = 1.60$; SV I performance between static mode ($M = 75.00$, $SD = 17.16$) and VR mode ($M = 55.00$, $SD = 7.07$), with $t(9) = 3.35$, $p = 0.008$, $d = 1.52$; and SV II performance between static mode ($M = 86.00$, $SD = 16.47$) and VR mode ($M = 62.00$, $SD = 23.94$), with $t(9) = 3.67$, $p = 0.005$, $d = 1.17$. For LSA, statistically significant difference was found in SV II between static mode ($M = 37.50$, $SD = 22.52$) and VR mode ($M = 22.50$, $SD = 24.93$) with $t(7) = 2.39$, $p = 0.048$, $d = 0.63$.

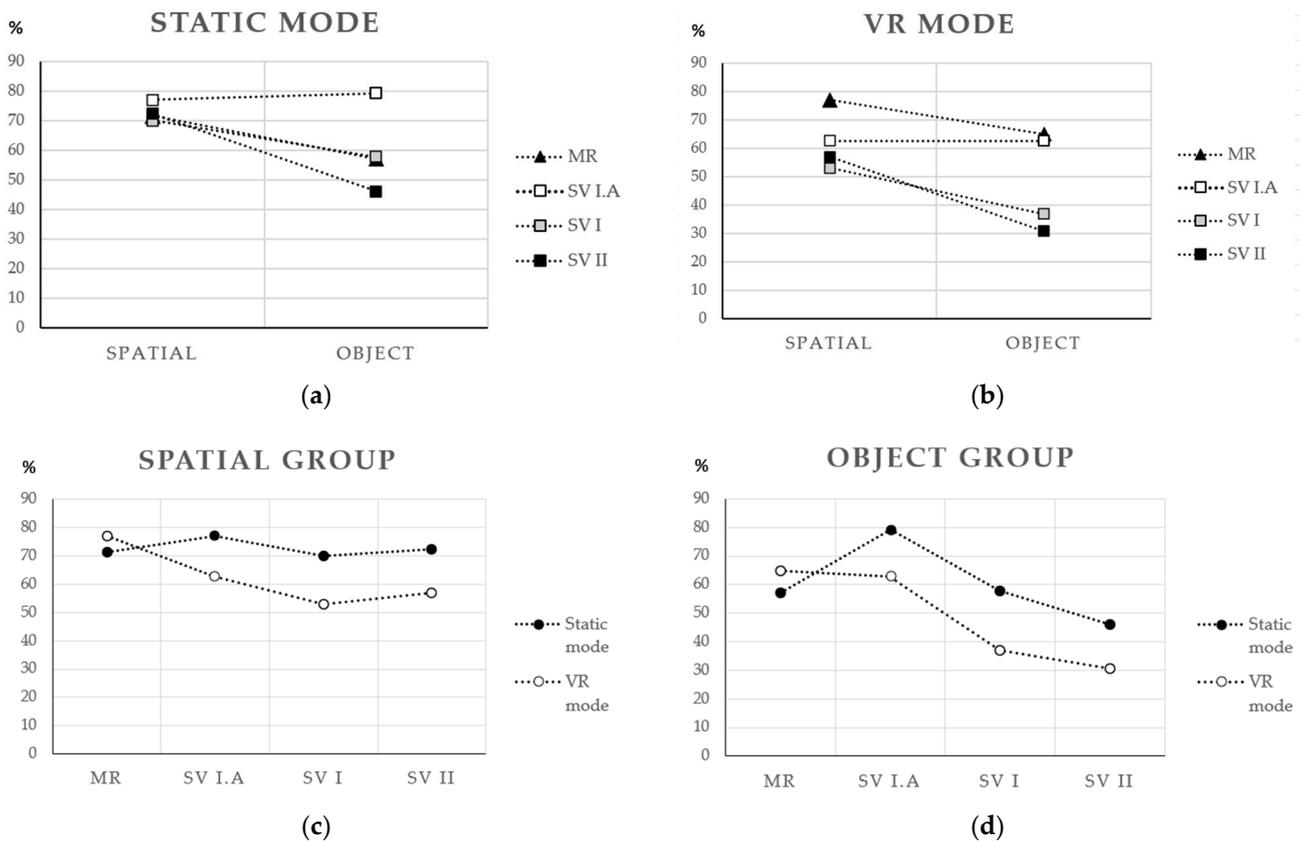


Figure 5. (a) Observed scores by OSIQ score in static mode; (b) Observed scores by OSIQ score in VR mode; (c) Observed scores by spatial group; (d) Observed scores by object group.

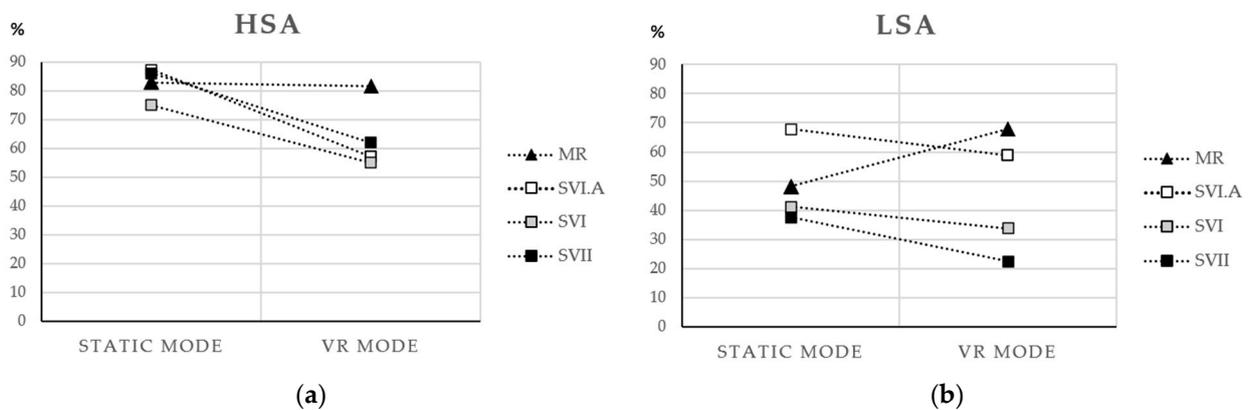


Figure 6. (a) Observed scores by HSA; (b) Observed scores by LSA.

4.7. Correlation between Static and VR Mode Performance

A correlation among static and VR mode performance and visual cognitive style scores was conducted to understand their relationship. The Pearson correlation analysis results showed that the average of all subconstructs in static mode correlated with that in VR ($r = 0.577$). However, only the performance in SV II in static mode shows statistically significant correlation with that in VR mode. The result implies that the overall performance in each of the two modes is related. However, when measuring the detailed subconstruct of spatial ability, differences in performance may be brought about due to the different test modes. As for the visual cognitive style, spatial scores on the OSIQ correlated with VR SV I ($r = 0.463$) and VR average score ($r = 0.447$) while object scores negatively correlated with

the paper version SV I.A ($r = -0.395$). The spatial visualizer group seems to have benefited from VR mode. See Table 5.

Table 5. Correlation between paper and VR AISAT performance.

	MR	SV I.A	SV I	SV II	Static_Average	VR MR	VR SV I.A	VR SV I	VR SV II	VR Average	Spatial Score	Object Score
MR	1	0.096	0.267	0.323	0.674 **	0.054	0.283	0.481 **	-0.031	0.358	0.246	-0.046
SV I.A		1	0.250	0.179	0.464 **	-0.182	0.152	0.216	0.428 *	0.246	-0.112	-0.395 *
SV I			1	0.401 *	0.780 **	0.250	-0.208	0.314	0.367 *	0.290	0.256	0.143
SV II				1	0.704 **	0.499 **	-0.076	0.514 **	0.695 **	0.687 **	0.262	-0.292
Static_average					1	0.264	0.031	0.574 **	0.500 **	0.577 **	0.284	-0.147
VR MR						1	-0.388 *	0.280	0.229	0.491 **	0.295	-0.074
VR SV I.A							1	0.180	-0.105	0.350	0.010	-0.119
VR SV I								1	0.447 *	0.819 **	0.463 **	-0.120
VR SV II									1	0.633 **	0.273	-0.259
VR average										1	0.447 *	-0.245
Spatial score											1	0.369 *
Object score												1

Note. * Significant at the 0.05 level; ** Significant at the 0.01 level (2-tailed).

4.8. Usability Test of VR AISAT

Participants' opinions on the usability of VR AISAT were analyzed. Table 6 summarizes the mean score of each question, reliability with Cronbach's alpha, and item average. To check the internal consistency of the questions in each category, Cronbach's alpha was computed. The alpha for tutorial helpfulness, task performance, exploration and navigation, satisfaction, and discomfort formed a scale that has reasonable internal consistency reliability with over 0.66.

Overall, participants perceived static mode as easier to handle than VR. They perceived the traditional mode as 'more comfortable' ($M = 3.67$) than VR mode ($M = 2.73$). Regarding tutorial helpfulness, participants indicated that understanding the tutorials on the four types of VR spatial ability tasks was 'helpful in completing the test' ($M > 4.10$). Regarding task performance, their answers indicated that 'they knew task goals and procedures' ($M > 3.90$), but since task results were not provided to users in the system, 'seeing the results' was rated relatively low ($M = 3.37$). Regarding exploration and navigation, although participants perceived the 'interface menus as easy to navigate' ($M = 4.17$), 'free exploration', 'knowing the exact location in the simulated environment', and 'free manipulation of the virtual objects' were not rated particularly high ($M < 3.40$). Overall, participants were satisfied with the VR AISAT experience ($M > 4.00$). They said 'doing the VR AISAT was fun and a good experience' ($M = 4.23$). Participants reported 'high immersion' ($M = 4.03$). In terms of discomfort, reports on 'feeling dizzy while wearing HMD' were higher than average ($M = 3.44$), and 'heavy headset' also received high marks ($M = 4.13$).

In addition to the structured usability questionnaire, participants were asked two open-ended questions: "Freely talk about your feelings while solving problems on the VR AISAT" and "Please suggest anything to improve the VR AISAT". Table 7 is a summary of participants' responses.

The main feelings that participants experienced while solving the VR AISAT were quite diverse: negative comments highlighted dizziness and uncomfortable HMD. Positive comments included 'helpful' and 'fun experience.' For improvement, participants suggested improving graphics for clarity and allowing more freedom of movement.

Table 6. Usability of VR AISAT item and mean score.

Subcategory	Statements	Mean of Each Question	Cronbach's Alpha	Item Average
Usability	I am more comfortable with a traditional paper test.	3.67	-	3.67
	I am more comfortable with the virtual reality method.	2.73	-	2.73
	The controls in the program are easy to use.	3.83		
	The eye-level position in the program is appropriate.	3.67		
	The brightness of the lighting in the program is adequate.	4.07	(0.371)	3.86
	The level of graphics in the program is appropriate.	3.67		
	The size of the objects, such as models and buttons, is appropriate.	4.07		
Tutorial helpfulness	The tutorial for the first type, MR, helped me understand how to solve the problem.	4.10		
	The tutorial on the second type, SV I. A, helped me understand how to solve the problem.	4.14		
	The tutorial on the third type, SV I, helped me understand how to solve the problem.	4.13	0.891	4.14
	The tutorial on the third type, SV I, helped me understand how to solve the problem.	4.17		
Task performance	I understood exactly what the goal of the task was.	4.20		
	I was able to decide what I wanted to do in the task.	4.00		
	I could easily see the results of my tasks.	3.37	0.710	3.88
	I could see exactly what the next task was.	3.93		
Exploration and Navigation	I could freely explore the experimental spaces.	3.37		
	I knew exactly where I was	3.03		
	The menus of the interface were easy to navigate.	4.17	0.661	3.49
	I was able to manipulate the virtual objects freely.	3.40		

Table 6. Cont.

Subcategory	Statements	Mean of Each Question	Cronbach's Alpha	Item Average
Satisfaction	I think I can measure my spatial ability through the virtual AISAT.	4.07	0.760	4.08
	I think the information and knowledge I gained from the virtual AISAT will help me improve my spatial ability in real life.	4.03		
	The virtual AISAT was fun and a good experience.	4.23		
	Doing the virtual AISAT is a valuable experience for me.	4.27		
	I would highly recommend the Virtual AISAT to others.	3.80		
Immersion	It was immersive while solving the virtual AISAT.	4.03	-	4.03
	I wish there were more environmental devices for immersion.	3.93	-	3.55
	I wish I could see my body for immersion.	3.17		
Discomfort	I wish I could see my body for immersion.	3.17	0.788	2.86
	I felt dizzy while using Virtual AISAT.	3.43		
	I felt nauseous while using Virtual AISAT.	2.93		
	I felt motion sickness while repositioning.	2.63		
	I felt motion sickness while rotating my head.	2.80		
	The headset hurts where it touches my skin.	2.37		
	The headset is heavy.	4.13		
The controller is difficult to use.	1.73			
Open-ended questions	Please feel free to share your thoughts on solving the problems. Tell us how we can improve the AISAT-VR			

Note. Grey highlight means reverse item; that is, a high score indicates high negativity.

Table 7. Summary of open-ended responses.

Question	Answers	Freq. (Multiple Answers Possible)	Question	Answers	Freq. (Multiple Answers Possible)
Freely talk about your feelings while solving problems on the VR AISAT.	Dizziness	9	Please suggest anything to improve the VR AISAT.	Graphics	6
	Helpful, fun	9		Freedom of movement	6
	Uncomfortable HMD	6		Dizziness	2
	Need graphics improvement	4		Accidental click on button resulting in moving forward to next question	2
	Difficult questions	4		Focus is not clear	2
	Problems are intuitive and easily understood.	4		Uncomfortable with VR devices	2
	Unfamiliarity	2		Adding background music	1
	Lack of freedom of movement	1		How to check your view is uncomfortable	1
	3D implementation is well done	1		Changing the way tutorials work	1
		Adding one's location		1	
		Show remaining time		1	
		Hard to see the bottom of the field of view		1	
		Difficult to see the point of view in the SV I view		1	
		Change MR model and arrow button layers	1		
	Immersive experience	1			

In summary, the questions concerning how research results supported our hypotheses are reported below:

H1: Performance in the static mode of AISAT and the VR mode of AISAT will be similar. → Mostly rejected. A significant difference was revealed in the performance in three SV constructs with better performance in static mode. Although no statistical difference was found in MR, MR performance was better in VR mode.

H2: The order of taking the tests (VR mode first vs. static mode first) will not influence spatial ability performance. → Mostly supported. Difference was observed only in SV II.

H3: No gender difference will be observed in the performance of the VR AISAT. → Supported. No gender difference was observed in any of the constructs in the VR mode.

H4: Spatial visualizers in visual cognitive style will perform better in VR mode than object visualizers. → Partially supported. Spatial visualizers performed better in VR SV I and VR SV II out of four constructs.

H5: Those with LSA will benefit from VR mode more than those with HSA. → Partially supported. The difference in performance between static mode and VR mode was observed in only one construct with those with LSA but three with HSA. The degree of decline from scores in static to that in VR mode was smaller in LSA, and HSA seems to be more susceptible to test mode.

5. Discussion

5.1. Differences in Spatial Ability Performance in Static and VR Modes

The results of this study showed a difference in performance on the AISAT in static mode and VR mode. Participants generally performed better in static mode than interactive VR except in MR. Possible reasons behind such differences in the test takers' performance could be (a) challenges caused by the VR system itself and (b) potential issues caused by slightly different procedures designed for the AISAT VR software. One of the challenges caused by the VR system itself was cybersickness. General discomfort, including dizziness, uncomfortable HMD, and lack of movement as indicated in the open-ended responses may relate to the test takers' low performance in VR. Researchers have theorized the potential causes of cybersickness. Among them, sensory conflict theory explains that cybersickness occurs with a mismatch between visual senses and the vestibular senses [40]. The teleporting option added to SV II was intended to provide options to navigate to view the space from many different positions, allowing the viewer to have a holistic understanding of the given space. However, the sudden positional change of the users might have caused them to experience cybersickness.

Another possible reason for performance difference in the two modes might be the slightly different procedures designed for the AISAT VR software: First, an additional action was required in taking the test of each subconstruct of spatial ability in the VR mode. For example, compared to the static mode, in VR more steps in the thinking procedure were required to solve problems. In VR SV I.A, the user's position in relation with the 3D objects matters, so the four answer options were in four directions in a radial organization, requiring users to rotate their bodies 90 degrees to see each option. The test takers had to use their hands to manipulate the controller. Second, differences existed in the field of vision in the two test modes. The ease of grasping the entirety of the space in question in the static mode compared with the VR mode allowed the user to perform better, especially in the spatial visualization test. Encoding the geometry of a room in VR was reported difficult [41] perhaps as a result of the field of view. To take advantage of the benefit of immersive quality of VR technology, real-world scale was applied to the simulated models in question in AISAT VR mode. Therefore, compared to the static mode, where test-takers could see the spatial objects in question and multiple choices together, VR users needed to explore the space and the answer choices to fully understand the question, thereby causing more cognitive load. Last, although we provided a training tutorial with a few sample questions before they took the test in VR mode, more exercises might have been needed

for some of the test takers. Providing enough exercises and a substantial training period is critical in VR [21].

In fact, prior studies investigating the effect of VR on spatial ability revealed inconsistency in findings regarding the benefit of VR. For example, better SV performance was observed in a 2D floor plan-based task than in VE [42]. In other studies, however, better SV performances were reported in VR than non-VR [43] and in printed materials [31]. In contrast, previous research comparing MR performance showed more consistent results, reporting better performance in VR than non-VR mode [5,44], with printed materials [31], and in 2D [4]. Although no statistical difference was found, the results of the current study support the previous study on MR performance regarding better performance in VR than that in static mode. In MR in VR mode, users could strategically rotate the option until they could equate it with the shape suggested in the question, reducing the test takers' cognitive load. Especially with the function that allows the test taker to rotate the object in real time, the participants had the advantage of quickly grasping the rotated forms to solve the problem.

The factors noted above did not diminish the benefit of using VR for developing spatial ability measuring tools. In VR mode, users could experience the sense of scale that is part of spatial skills not formally recognized in previous spatial ability studies. The ease of viewing the entirety of the space in question in the static mode benefited users allowing them to perform better in the static mode. However, the static mode did not provide a sense of presence and immersion. Based on the results of this study, we believe the perceptual experience of spatial dimensions and solving spatially relevant problems cannot be considered the same cognitive activities. Therefore, this research highlights the necessity of careful consideration in designing and developing VR test modes for spatial ability measurement.

5.2. Less Gender Difference in VR

The results of this study regarding the gender difference in VR align with prior research. Gender difference was found only in MR in the static mode. However, no gender difference was found in any of the subconstructs in VR mode. Lochhead et al. (2022) found that less gender difference in 3D stimuli MR in VR than 2D image stimuli in VR may have occurred since the 3D stimuli do not require the 2D to 3D transformation step, "alleviating the cognitive burden of dimensionality crossing" ([3], p. 15), imposed by 2D representation of a 3D object. The difference in strategies used by participants of each gender in solving MR problems may be another possible reason behind less gender difference in performance in VR mode. Functional brain imaging in a study by Jordan et al. (2002) showed that activation patterns in men and women differed when they solved MR problems [45]: The brain areas known to be involved in the cognitive processing of object-part identification, object categorization, and spatial analysis were activated more in women than in men; whereas brain areas involving more concrete imagined rotation similar to hands-on approach were activated in men (p. 2406). The function in the VR version AISAT that allowed the test takers to view the rotated forms of the given object in real time by one click may have advantaged female test takers to quickly grasp the rotated forms and expedited their cognitive processing of object-part identification and decision making based on spatial analysis (i.e., matching between rotated form and the image in question). For male test takers the way MR test is designed in VR may have forced them to shift the way to solve mental rotation task using different types of spatial analysis focused more on matching skills, disadvantaging their skills of imagined mental rotation using primary motor cortex. Nonetheless, previous research findings also showed the lack of gender difference in MR tested in VR [30,31]. Less gender difference in VR modes indicates the potential of VR to benefit female users in spatial learning and thinking using VR.

5.3. Who Benefits from VR

In the current study those with HSA generally received higher scores than those with LSA regardless of test mode. When comparing their performance in each subconstruct of spatial ability, however, the spatial visualization scores of those with HSA were lower when the test was conducted in the VR mode, but for LSA most of the differences were diminished. In fact, only a difference in SV I was observed. In addition, the MR scores of those with LSA were higher in VR mode compared to that in static mode. This suggests that individuals with high spatial ability as assessed in the static mode do not consistently benefit from the VR test mode in spatial visualization, and those with HSA tend to be more susceptible to test mode than those with LSA. In addition, this study revealed spatial visualizers with visual cognitive style benefit from the VR mode. This suggests the potential challenges for those with an object-oriented cognitive style when using VR technology.

5.4. Implications

This study shows the potential of innovative and practical application and provides following implications to the community of spatial design researchers, practitioners, and educators. First, providing a means of holistic understanding of the spatial information in question in VR mode is necessary for better performance. Although the same questions were used for both test modes, when the static image of a space is converted into a simulated volumetric spatial form in a VE, a perceptual alteration may occur. In static mode, test takers can see the entire space in question, allowing them to grasp a holistic view of the space. When the real scale is applied in the VE and a viewer is put in front of the simulated building or inside the space, however, due to the limited field of vision in the VE, the viewer may need more time to reach a holistic understanding of the given space.

Second, making the procedure in solving problems in VR as simple as possible is essential. The complexity of the VEs and wearing HMD itself could bring about a sense of overwhelming and discomfort. In the AISAT, compared to the simple procedure and instruction in static mode, in VR a more complex procedure and learning curve occurred with each construct of SV. In addition, per each construct slightly different actions were needed, such as clicking the option button, looking at options in a larger scale, and teleporting. Such a difference in procedure in each construct might require more working memory, cognitive load, and effort to adapt to each type of question. This might have been the reason that the participants performed better in the static mode irrespective of the benefit of experiential quality becoming more realistic in the VE. Whether the static version or the VR version of the spatial ability test is closer to the spatial ability required to solve spatial problems in the real world is unknown.

Third, using VR mode for training purposes and static mode for measuring purposes is encouraged. Although findings were limited, one result shows the VR first group performed better in SV I.A and SV I static mode later compared to static first group, indicating the potential benefit for developing a VR version for training of spatial ability. Previous research [16,46] also has shown the benefit of virtual simulation technology in aiding spatial learning. For example, after using extended reality technology in a design project, participants' spatial ability improved [16]. Although measuring spatial ability in static mode and in VR mode is not perfectly compatible, the use of the VR mode of the AISAT as a training tool may have a potential benefit in enhancing at least spatial visualization skills.

Fourth, educators must be cautious and avoid the hasty development of VR tools for spatial learning and education without a thorough consideration of the complex nature of VR technology and its interaction with individual differences. Despite the advantages of VR technology, the user's perceptual understanding of spaces in VEs may differ from that in static test modes.

5.5. Limitations

The first limitation is the sample size (30 participants) and the imbalance between the number of female and male students. The participants were interior design students in Korea; whether the findings in this research are consistent in other cultures and more skilled interior designers is unknown. Even though we planned a total of 46 participants, some completed only one mode due to COVID-19 and could not be included for analysis. In addition, due to the time-consuming nature of the AISAT designed to measure domain-specific spatial ability and the limited number of VR HMDs, voluntary participation was low. However, the gender proportion in interior design programs is female dominant at 85% in UK, not much different in Korea. Thus, the disparity of the gender proportion in this study is likely to represent the status quo of the population. The second limitation regards the VR HMD used in the study. Since the VR AISAT was developed for HTC VIVE specifically, participants reported its heaviness as part of the inconvenience of the test mode. Since the up-to-date lighter devices have been introduced, applying them to different types of HMD seems necessary. The third is the experimental design limitations. In terms of time allotted for VR mode, ours was based on a pilot study of 10 students, necessitating more investigation into the time needed with a greater number of participants.

6. Conclusions and Future Works

In this study we explored the influence of test modes on domain-specific spatial ability performance in the field of architecture and interior design. We examined two distinct test modes: a static paper–desktop mode and an interactive VR mode. Conclusions are as follows:

1. Our findings reveal that participants demonstrated better performance in the spatial visualization test when using the static mode as opposed to the VR mode. Conversely, mental rotation performance was better in the VR mode.
2. The order of test mode minimally influenced performance.
3. The gender difference typically observed in mental rotation tests was not evident in the VR mode, indicating potential advantages for female users in employing VR for spatial learning related to mental rotation.
4. Individuals with a strong spatial visualizer cognitive style excelled, particularly in the SV I and SV II conducted in VR mode.
5. People with high spatial ability tended to be more influenced by test mode than those with low spatial ability.
6. Spatial visualizers and low spatial ability groups tended to benefit from VR mode.
7. Regarding usability, participants expressed positive reactions to VR AISAT, but challenges due to the nature of virtual reality occurred, including dizziness and participants' perception that solving spatial problems in the static mode was easier than in the VR mode.

These findings highlight the intricate nature of spatial ability across test modes, further influenced by individual differences.

In future, the following activities should be carried out by researchers: First, reexamining the time allotted for each construct in the VR mode of the AISAT is necessary with more participants and can result in higher a level of correlation in each construct of the static and VR modes. Second, improving graphic quality and the usability of the AISAT based on participants' responses, especially using lighter HMD for ease of use and comfort would be of considerable value. Third, making a greater number of questions in each construct with the aid of rule-based question generation will be helpful in the development of the AISAT. Artificial Intelligence technology can also be used to customize the question types and number of questions in real time to measure more detailed spatial ability specific to an individual. Fourth, improving the VR mode of AISAT and recruiting a larger number of participants from different cultures and with diverse levels of experience will facilitate an understanding of how and what areas the VR mode benefits; generalizability will follow. In addition, conducting a study with bio and brain sensors, such as an eye-tracker and an EEG, to understand the strategies of solving space-related problems involving high vs. low

spatial ability and the cognitive and neural mechanism of spatial thinking and visualization would be highly valuable.

Author Contributions: Conceptualization, J.Y.C. and J.S.; methodology, J.Y.C. and J.S.; formal analysis, J.Y.C.; validation, J.Y.C. and J.S.; investigation, J.Y.C. and J.S.; resources, J.Y.C.; data curation, J.Y.C.; writing—original draft preparation, J.Y.C. and J.S.; writing—review and editing, J.Y.C. and J.S.; visualization, J.Y.C.; funding acquisition, J.Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded the BK21 plus program “AgeTech-Service Convergence Major” through the National Research Foundation (NRF) funded by the Ministry of Education of Korea [5120200313836] and by the NRF grant funded by the Korea government (MSIT) [2020R1A2C1009689].

Institutional Review Board Statement: The study was approved by the Institutional Review Board of Kyung Hee University IRB (protocol code: KHSIRB-21-146 (NA) and date of approval: 6 April 2021 and 6 April 2022).

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

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to [privacy].

Acknowledgments: We acknowledge all participants in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Obeid, S.; Demirkan, H. The influence of virtual reality on design process creativity in basic design studios. *Interact. Learn. Environ.* **2023**, *31*, 1841–1859. [CrossRef]
2. Ceylan, S. Using virtual reality to improve visual recognition skills of first year architecture students: A comparative study. In Proceedings of the 12th International Conference on Computer Supported Education (CSEDU), Prague, Czech Republic, 2–4 May 2020; Volume 2, pp. 54–63.
3. Lochhead, I.; Hedley, N.; Çöltekin, A.; Fisher, B. The immersive mental rotations test: Evaluating spatial ability in virtual reality. *Front. Virtual Real.* **2022**, *3*, 820237. [CrossRef]
4. Passig, D.; Eden, S. Virtual reality as a tool for improving spatial rotation among deaf and hard-of-hearing children. *CyberPsychol. Behav.* **2001**, *4*, 681–686. [CrossRef] [PubMed]
5. Yurt, E.; Sünbül, A.M. Effect of modeling-based activities developed using virtual environments and concrete objects on spatial thinking and mental rotation skills. *Educ. Sci. Theory Pract.* **2012**, *12*, 1987–1992.
6. Kim, N.; Gero, J.S. Neurophysiological responses to biophilic design: A pilot experiment using VR and EEG. In *Design Computing and Cognition*; Gero, J.S., Ed.; Springer International Publishing: Cham, Switzerland, 2022; pp. 235–253. [CrossRef]
7. Blajenkova, O.; Kozhevnikov, M.; Motes, M.A. Object-spatial imagery: A new self-report imagery questionnaire. *Appl. Cogn. Psychol.* **2006**, *20*, 239–263. [CrossRef]
8. Linn, M.C.; Petersen, A.C. Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Dev.* **1985**, *56*, 1479–1498. [CrossRef] [PubMed]
9. Gaughran, W. Cognitive modelling for engineers. In Proceedings of the 2002 American Society for Engineering Education Annual Conference and Exposition, Montreal, QC, Canada, 16–19 June 2002; American Society for Engineering Education: Washington, DC, USA, 2002. [CrossRef]
10. Hegarty, M.; Waller, D. Individual differences in spatial abilities. In *The Cambridge Handbook of Visuospatial Thinking*; Shah, P., Miyake, A., Eds.; Cambridge University Press: Cambridge, UK, 2005; pp. 121–169. [CrossRef]
11. Vandenberg, S.G.; Kuse, A.R. Mental rotations, a group test of 3-dimensional spatial visualization. *Percept. Mot. Ski.* **1987**, *47*, 599–604. [CrossRef] [PubMed]
12. Ekstrom, R.B.; French, J.W.; Harman, H.H.; Dermen, D. *Manual for Kit of Factor Referenced Cognitive Tests*; Educational Testing Service: Princeton, NJ, USA, 1976.
13. Hegarty, M.; Waller, D. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* **2004**, *32*, 175–191. [CrossRef]
14. Sutton, K.; Williams, A. Implications of spatial abilities on design thinking. In *Proceedings of the Design and Complexity: DRS International Conference 2010, Montreal, QC, Canada, 7–9 July 2010*; Durling, D., Bousbaci, R., Chen, L., Gauthier, P., Poldma, T., Roworth-Stokes, S., Stolterman, E., Eds.; Design Research Society: London, UK, 2010; Available online: <http://www.drs2010.umontreal.ca/data/PDF/115.pdf> (accessed on 10 September 2023).
15. Berkowitz, M.; Gerber, A.; Thurn, C.M.; Emo, B.; Hoelscher, C.; Stern, E. Spatial abilities for architecture: Cross sectional and longitudinal assessment with novel and existing spatial ability tests. *Front. Psychol.* **2021**, *11*, 4096. [CrossRef]
16. Darwish, M.; Kamel, S.; Assem, A. Extended reality for enhancing spatial ability in architecture design education. *Ain Shams Eng. J.* **2023**, *14*, 102104. [CrossRef]
17. Cho, J.Y.; Suh, J. The architecture and interior design domain-specific spatial ability test (AISAT): Its validity and reliability. *J. Inter. Des.* **2022**, *47*, 11–30. [CrossRef]

18. Cho, J.Y.; Suh, J. Understanding spatial ability in interior design education: 2D-to-3D visualization proficiency as a predictor of design performance. *J. Inter. Des.* **2019**, *44*, 141–159. [[CrossRef](#)]
19. Suh, J.; Cho, J.Y. Linking spatial ability, spatial strategies, and spatial creativity: A step to clarify the fuzzy relationship between spatial ability and creativity. *Think. Ski. Creat.* **2020**, *35*, 100628. [[CrossRef](#)]
20. Sherman, W.R.; Craig, A.B. *Understanding Virtual Reality*; Morgan Kaufman: San Francisco, CA, USA, 2003.
21. Waller, D.; Hunt, E.; Knapp, D. The transfer of spatial knowledge in virtual environment training. *Presence* **1998**, *7*, 129–143. [[CrossRef](#)]
22. Ahn, J.M.; Cho, J.Y. The potential of VR and AR in improving designers' spatial ability. *Conf. Korean Inst. Inter. Des.* **2018**, *20*, 47–49.
23. Banfi, F.; Brumana, R.; Stanga, C. Extended reality and informative models for the architectural heritage: From scan-to-BIM process to virtual and augmented reality. *Virtual Archaeol. Rev.* **2019**, *10*, 14–30. [[CrossRef](#)]
24. Silva, M.; Teixeira, L. Developing an extended reality platform for immersive and interactive experiences for cultural heritage: Serralves museum and coa archeologic park. In Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Recife, Brazil, 9–13 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 300–302.
25. Panya, D.S.; Kim, T.; Choo, S. An interactive design change methodology using a BIM-based Virtual Reality and Augmented Reality. *J. Build. Eng.* **2023**, *68*, 106030. [[CrossRef](#)]
26. Alizadehsalehi, S.; Yitmen, I. Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX). *Smart Sustain. Built Environ.* **2023**, *12*, 200–236. [[CrossRef](#)]
27. Ju, J.E.; Kim, A.Y. Gender differences in spatial ability and using strategies for solving spatial problems. *Korean J. Woman Psychol.* **2010**, *15*, 829–851.
28. Maeda, Y.; Yoon, S.Y. A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT: R). *Educ. Psychol. Rev.* **2013**, *25*, 69–94. [[CrossRef](#)]
29. Reilly, D.; Neumann, D.L. Gender-role differences in spatial ability: A meta-analytic review. *Sex Roles* **2013**, *68*, 521–535. [[CrossRef](#)]
30. Park, J.-R. The Analysis of Spatial Ability Difference According to Learning Style and Gender in Virtual Reality Based Learning. Master's Thesis, Korea National University of Education, Cheongju, Republic of Korea, 2009. Available online: <https://www-riss-kr.openlink.khu.ac.kr/link?id=T11590539> (accessed on 10 September 2023).
31. Samsudin, K.; Rafi, A.; Hanif, A.S. Training in mental rotation and spatial visualization and its impact on orthographic drawing performance. *Educ. Technol. Soc.* **2011**, *14*, 179–186.
32. Cho, J.Y. An investigation of design studio performance in relation to creativity, spatial ability, and visual cognitive style. *Think. Ski. Creat.* **2017**, *23*, 67–78. [[CrossRef](#)]
33. Kozhevnikov, M.; Kozhevnikov, M.; Yu, C.J.; Blazhenkova, O. Creativity, visualization abilities, and visual cognitive style. *Br. J. Educ. Psychol.* **2013**, *83 Pt 2*, 196–209. [[CrossRef](#)] [[PubMed](#)]
34. Lee, E.A.L.; Wong, K.W. Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Comput. Educ.* **2014**, *79*, 49–58. [[CrossRef](#)]
35. Sun, R.; Wu, Y.J.; Cai, Q. The effect of a virtual reality learning environment on learners' spatial ability. *Virtual Real.* **2019**, *23*, 385–398. [[CrossRef](#)]
36. Hays, T.A. Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *J. Educ. Comput. Res.* **1996**, *14*, 139–155. [[CrossRef](#)]
37. Huk, T. Who benefits from learning with 3D models? The case of spatial ability. *J. Comput. Assist. Learn.* **2006**, *22*, 392–404. [[CrossRef](#)]
38. Meneely, J.; Portillo, M. The adaptable mind in design: Relating personality, cognitive style, and creative performance. *Creat. Res. J.* **2005**, *17*, 155–166.
39. British Institute of Interior Design. Ground-Breaking Diversity Analysis of Interior Design Students. 10 December 2020. Available online: <https://biid.org.uk/resources/ground-breaking-diversity-analysis-interior-design-students> (accessed on 20 November 2023).
40. LaViola, J.J., Jr. A discussion of cybersickness in virtual environments. *ACM Sigchi Bull.* **2000**, *32*, 47–56. [[CrossRef](#)]
41. Kimura, K.; Reichert, J.F.; Olson, A.; Pouya, O.R.; Wang, X.; Moussavi, Z.; Kelly, D.M. Orientation in virtual reality does not fully measure up to the real-world. *Sci. Rep.* **2017**, *7*, 18109. [[CrossRef](#)] [[PubMed](#)]
42. Schnabel, M.A.; Kvan, T. Spatial understanding in immersive virtual environments. *Int. J. Archit. Comput.* **2003**, *1*, 435–448. [[CrossRef](#)]
43. Chen, C.J. Are spatial visualization abilities relevant to virtual reality? *E-J. Instr. Sci. Technol.* **2006**, *9*, n2.
44. Molina-Carmona, R.; Pertegal-Felices, M.; Jimeno-Morenilla, A.; Mora-Mora, H. Virtual reality learning activities for multimedia students to enhance spatial ability. *Sustainability* **2018**, *10*, 1074. [[CrossRef](#)]

45. Jordan, K.; Wüstenberg, T.; Heinze, H.J.; Peters, M.; Jäncke, L. Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia* **2002**, *40*, 2397–2408. [[CrossRef](#)]
46. Jiang, E.; Laidlaw, D.H. Practicing in Virtual Reality Improves Mental Rotation Ability: Lower Scorers Benefit More. 2019. Available online: https://cs.brown.edu/media/filer_public/26/02/2602a1b3-b630-4e19-9cf6-5f077f5f3271/jiangelaine.pdf (accessed on 10 September 2023).

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.