Improving Spatial Visualization and Mental Rotation Using FORSpatial Through Shapes and Letters in Virtual Environment

Yun Zhou, *Member, IEEE*, Tao Xu, *Member, IEEE*, Han Yang, and Shiqian Li

***Abstract*—Existing research on spatial ability recognizes the critical role played by spatial visualization and mental rotation. Recent evidence suggests that external visualization and manipulation can boost spatial thinking. The virtual environment provides an exciting opportunity so that many spatial ability training tasks based on reading printed illustrations can be migrated to a highly 3-D interactive and visualized environment. However, few studies have employed virtual reality (VR) technology to improve spatial visualization and mental rotation. In addition, the design of training contents and corresponding VR applications are still lacking. In this work, we propose FORSpatial, a system mainly for spatial ability training in a virtual environment. First, in this article, we design a novel scheme and principles for generating tasks, involving spatial visualization and mental rotation through flexible combinations of shapes and letters. Based on this, we create testing questions and a FORSpatial training application in desktop VR. FORSpatial and its components are made publicly available and free to use. To evaluate the performance of spatial training, verify the usability of the FORSpatial application, and analyze learning behavior, we organized a user study with 49 .participants, including an experimental group and a control group. The comparison between experimental and control groups shows the significant improvement of spatial skills through training. The analysis of interaction logging data and subjective comments reveals how FORSpatial supports spatial thinking.**

***Index Terms*—Interaction recording and analysis, learning, spatial ability, training and development, virtual and augmented reality, virtual environments.**

I. Introduction

SPATIAL skills refer to a class of cognitive functions that involve the ability to imagine and mentally transform spatial information [1]–[3]. People create mental images in mind when they engage in daily life tasks, such as arranging the objects in a room or finding a way. They also use the spatial ability when engaging in professional work, such as designing a structure of the pipeline or working on the internal structure of the human body during surgery. Previous research has recognized that spatial ability plays a critical role in science, technology, engineering, and mathematics (STEM) [4] and schooling [5]. Good spatial skills positively impact the achievement of people in STEM. For people who work in STEM fields, a lack of spatial abilities may result in many obstacles to work or even lead to failure for career development.

Spatial ability is not a single ability, but a complex ability composed of multiple subskills [1], [3], [6]. Carroll [6] surveyed more than 90 datasets and identified the factor structure of visuospatial ability. These factors include spatial visualization, spatial relations, closure speed, flexibility of closure, perceptual speed, and visual memory. Some studies have pointed out their training goals aiming at which subskill [7]–[9], while some have used the general term spatial ability instead of explicit subskills [10], [11]. Among these factors, spatial visualization and mental rotation underlie the fundamental of spatial ability, and they play a critical role in performing daily or professional tasks. For example, when surgeons work, they visualize a 3-D image of an organ and surrounding tissue and simultaneously rotate the organ with details partly or completely in mind. Several researchers have wonderful work on developing spatial visualization [9], [12] and mental rotation [7]. However, current studies have suffered from a lack of an investigation of these two subskills.

How to improve spatial ability is an attractive and significant challenge for researchers. The last two decades have seen a growing trend toward using computer-aided techniques to test and train spatial ability [1]. Researchers used interactive animation methods [13], [14], engineering drawings [15], [16], and the applications of the 3-D environment [9], [17] instead of early methods based on reading printed illustrations. Among computer-aided applications, virtual reality (VR), augmented reality, and virtual environment [18] show their benefits [13], [17], [19]–[21], with 3-D display and rich spatial interaction. In such VR and 3-D environments, researchers and practitioners can create objects and scenes challenging to build in reality. Due to its low cost and convenience [22], [23], desktop VR has been regarded as ideal for creating a learning environment [11]. However, few studies have explored the design and development of desktop VR applications to train spatial visualization and mental rotation.

Manuscript received 7 June 2020; revised 12 August 2021; accepted 23 April 2022. Date of publication 28 April 2022; date of current version 5 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grants 62077036 and 61703259, and in part by the National Key Research and Development Program of China under Grant 2018AAA0100501. *(Corresponding author: Tao Xu.)*

Yun Zhou, Han Yang, and Shiqian Li are with the Faculty of Education, Shaanxi Normal University, Xi'an 710062, China (e-mail: zhouyun@snnu.edu.cn; yanghan@snnu.edu.cn; qc@snnu.edu.cn).

Tao Xu is with the School of Software, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: xutao@nwpu.edu.cn).

Digital Object Identifier 10.1109/TLT.2022.3170928

This article presents the design, implementation, and user study of FORSpatial, a system for spatial ability testing and training in a virtual environment. First, we design a novel scheme and principles for generating tasks, involving spatial visualization and mental rotation through flexible combinations of shapes and letters. Then, we create testing questions based on this scheme and principles and propose a FORSpatial training application with 3-D interaction in a desktop VR environment. By changing the combination, we can adjust the difficulty of the tasks. FORSpatial is downloadable and free to use, and its components are available on the Unity asset store. Letter F, O, R, Arrow, and eight shapes underlie the current version of the FORSpatial training application.

The salient aspects of FORSpatial are as follows.

1) The scheme and principles underlying FORSpatial are designed mainly for spatial visualization and mental rotation. Based on shapes and letters, FORSpatial focuses on testing and training three facets: three-dimensional shape folding ability, rotation of mental objects, and rotation of patterns on the surface of the mental object.
2) FORSpatial is flexible. Through a customized combination of shapes and letters, FORSpatial can create tasks for testing and training the user's spatial skills levels. The testing and training difficulty level can be adjusted to meet the needs of specific groups, making it suitable for primary, secondary, and college students. Unify testing and training into the same framework enables test scores to reflect training performance more accurately.
3) With desktop VR, current FORSpatial takes advantage of the high interactivity of VR and considers pedagogical aspects. It helps users to visualize and rotate objects in the learning area with 3-D interaction. The practice with feedback supports the user to reflect, fix the error in the cognitive structure in mind, and learn with goals. High interactivity and feedback-based exercises enable students to maintain high levels of engagement in training.

To evaluate the performance of spatial training, verify the usability of the FORSpatial application, and analyze learning behavior, we organized a user study with 49 participants, including an experimental group and a control group. In this user study, we collect the data through interaction logging, questionnaire, and interviews. The comparison between experimental and control groups shows a significant improvement of spatial skills through training. Results from interaction logging data, questionnaires, and interviews reveal how FORSpatial supports spatial thinking. Finally, we discuss the limitations and future possibilities. The findings contribute to spatial skills training in desktop VR.

This article contains three research purposes:

1) to investigate the design rationale for spatial visualization and mental rotation upgrading;
2) to explore how desktop VR can be combined with the pedagogical aspects and together serves as the design foundation of spatial training;
3) to evaluate the effects of such VR training on enhancing spatial ability.

II. Related Work

Most of the research shares a consensus on spatial thinking: spatial thinking is about the shapes and arrangements of objects and spatial processes [24]. Spatial thinking skills link to spatial visualization and strategies [1], [24], involving the ability to imagine and mentally transform spatial information. A great deal of previous research into spatial thinking has focused on spatial ability tests and spatial thinking tasks. This section will briefly present and discuss spatial visualization and mental rotation, training in spatial thinking studies, and desktop VR for promoting learning.

A. Spatial Visualization and Mental Rotation

The spatial ability is complex and consists of multiple subskills. Researchers have tried to figure out potential components of spatial ability through the factor-analytic approach. Linn and Petersens' work divided spatial ability into three factors: spatial visualization, mental rotation, and spatial perception. Hegarty and Waller [3] reviewed research on spatial ability and listed the factors based on Carroll's work [6] as spatial visualization, spatial relations, closure speed, flexibility of closure, perceptual speed, and visual memory.

Although the definition of spatial visualization has not yet reached a unified description, researchers agreed on an understanding of this factor, that is, spatial visualization is identified as a process to imagine spatial movements of objects mentally [3], [6], [25]. To identify essential components of spatial ability, Hegarty [24] examined what spatial ability tests measure. For example, paper folding and cube comparisons [26] focus on spatial visualization. The task in the paper folding test [26] is to indicate which of the five diagrams on the right shows how the paper will look when being unfolded.

People often rotate objects mentally to change their perspective. Mental rotation refers to the ability to rotate an object in a space mentally. In Shepard and Metzler's experiment [27], participants were presented with pairs of 2-D representations of 3-D objects. Participants' tasks were to determine whether the objects were identical except orientation. Results showed that the participants rotated the object in a 3-D space in their minds. Card rotations [26], Vandenberg mental rotations tests [28], and the Purdue visualization of rotations test [29] focus on the test of mental rotation. The task in the card rotations test is to identify which of the items on the right are the one given item on the left.

Concerning spatial ability, a STEM task requires several subskills [1], [24], among which spatial visualization and mental rotation contribute the most. For example, when performing a task requiring spatial ability, an individual needs to represent a 3-D image of an object and rotate it with details in mind. Therefore, this study aims at training spatial visualization and mental rotation as a group.

B. Spatial Ability Training

Previous studies have reported that spatial ability is malleable and can be improved through several means. Three training categories have been classified by Uttal *et al.* [1]: courses, video games, and spatial tasks. The course of spatial ability training generally lasts for one or several semesters [16] or lessons. Video game training involves playing a video or computer game [7]. For example, Cherney [7] investigated how 3-D and 2-D computer games improve performance on mental rotation tasks. Participants completed the training over two weeks or within one week. The findings suggest that even very minimal computer game practice may improve performance on mental rotation tasks. Due to the high interactivity, 3-D applications [11], VR applications [9], [12], and tangible interaction applications [8], [10] have drawn attention for spatial ability improvement. Wang *et al.* explored the effects of 2-D- versus 3-D-based media representations on the influence of spatial visualization ability [11]. They conducted a pretest/post-test comparison-group experiment with 23 participants involved. The intervention was within a 35-min period. Lee *et al.* used a V-FrogTM software to explore spatial learning effects in desktop VR-based environment [11]. They employed a quasi-pretest-post-test experimental design, and the training took about 1.5 h. Their findings indicated a significant positive effect of the desktop VR-based learning environment on the performance outcome in biology education for high and low spatial ability learners. Yu *et al.* proposed a spatial ability training and test through tangible interaction with an EasySRRobot [8]. Chang *et al.* proposed a tangible VR system and evaluated it using pre- and post-tests. Their results showed statistically significant improvements for enhancing spatial cognition [10].

Usually, courses and video games train spatial ability through indirect tasks that impact spatial ability do not follow the same rationale as those used in the tests. Contreras *et al.* [16] investigated whether the indirect training in technical drawing course improved the spatial visualization ability of architecture students. The subjects in their experiment were freshman students enrolled in technical drawing courses and discrete mathematics courses. These indirect pieces of training involve more complex spatial tasks, requiring task decomposition and rule-based reasoning to augment visualization [24].

Spatial ability can also be trained directly through the tasks that follow the same rationale as those used in the tests for ability measurement, such as in [8]. Direct pieces of training are more about imagery strategies and analytic strategies [24]. In this work, we focus on direct training with spatial tasks supported by desktop VR applications.

C. Desktop VR Supported Training

VR simulates an environment similar to or completely different from the real world. VR computer simulations can be explored and interacted with by people. There is a growing trend to use VR-based learning to help absorb knowledge and improve skills in schools and colleges. Different types of VR applications bring various experiences and meet specific requirements [30]. Zhou *et al.* have classified four types of VR: desktop semi-immersive VR, mobile semi-immersive VR, fully immersive VR room, and fully immersive headset supported VR [31]. Lee *et al.* categorized VR into two types: immersive VR and nonimmersive VR (desktop VR) [11]. Regardless of how VR is classified, people have a consistent view of desktop VR. Desktop VR environment, in which 3-D graphics performance is delivered by 2-D screen, supports rich spatial interaction, such as rotating, by using the mouse, keyboard, touch screen, and standard interactive devices [11], [31], [32]. Although the interaction and immersive experience provided by desktop VR are less than the immersive type of VR, it is an alternative to immersive VR and has its benefits. It costs a lower price, can be easily deployed in the classroom, and is convenient for collaborative learning and larger groups [22], [23]. A 3-D display and spatial operation underlie the educational use of desktop VR [33], [34]. For example, Arloon series applications [33] support children to interact with virtual scenes and objects on a mobile phone or tablet 3-D and learn about plant ecosystems, geometry, chemistry, solar system, etc. The work in [35] proposed a collaborative VR learning environment to facilitate 3-D geometric problem-solving. Similarly, desktop VR can provide 3-D display and rich spatial manipulation to reach spatial ability training. In addition, an advantage of using VR environments lies in the possibilities of automatically recording users' interactions. However, using off-the-shelf VR applications to train spatial ability could not meet such advanced requirements to collect and analyze users' behaviors.

It is not enough to focus on interaction aspects to better design a specific VR-supported learning and training experience. Pedagogical aspects should also be taken into consideration [31]. The user must be attracted and engage in important information in training. Practice plays a vital role in raising engagement. The practice has been proved to be beneficial to learning by a great deal of previous research [36]. Providing feedback for practice can affect learning and reinforce the learning effect [37]. Therefore, designing a specific desktop VR application using practice with feedback can enhance the learning experience and improve spatial ability.

III. Design Rationale and Implementation of FORSpatial

A few previous studies have employed VR technology to improve spatial visualization and mental rotation. The design of training contents and how to scaffold training in VR are still lacking. Therefore, this study was motivated due to the missing studies on designing an application for training spatial visualization and mental rotation in a VR environment. FORSpatial focuses on the overall performance of people's 3-D representation, 3-D object rotation, and letter rotation in mind. As shown in Fig. 1, the FORSpatial system contains two parts: testing and training application. The scheme and principles describe the ways that we create the testing. Then, we design and implement the training application based on the proposed scheme and principles using VR technology. The training application supports visualization, direct manipulation, and instant practice, allowing users to drag, rotate, observe 2-D–3-D dynamic transformation, and self-test with feedback.

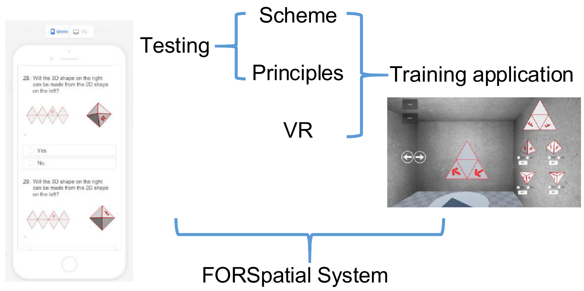


Fig. 1. Architecture of the FORSpatial system.

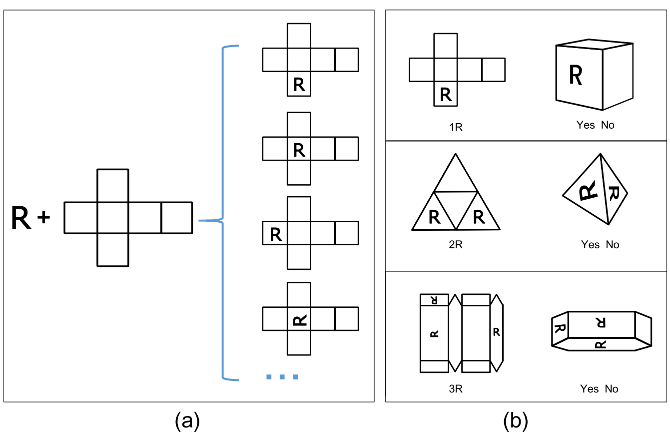


Fig. 2. Design scheme based on a flexible combination of shapes and letters in FORSpatial. (a) Combination of one letter and one shape, and the samples of 2-D shape from this combination. (b) Three sample questions for FORSpatial test.

A. Design Rationale

The design scheme is based on a flexible combination of shapes and letters. Through the combination of selected shapes and letters, we obtain multiple groups of 2-D shapes with letter patterns [see Fig. 2(a)]. Each shape has its corresponding 3-D style, and different rotation angles make the 3-D object look different. Based on the FORSpatial scheme, we create 3-D representation tasks to investigate how easily users can see and manipulate the shapes and letters spatially in their minds. As shown in Fig. 2(b), each item in the FORSpatial scheme has a drawing of a flat shape (unfolded 2-D shape) on the left, with one, two, or three same letters on surfaces of this shape. On the right, an object is shown, which might be made by folding the flat shape along the line. The task is to determine whether or not each given folded 3-D shape with letters can be made from an unfolded 2-D shape. For each question, users give answers by choosing yes or no. Therefore, it requires the user to build 3-D object representation and direction in mind and work well on mental letter rotation.

FORSpatial system has excellent flexibility. F in FORSpatial refers to the letter that we use in practice with feedback in the application. O refers to the letter that we employ in the learning area with 3-D interaction in the application. R refers to the letter that we use in the test. These letters can be selected, and the shapes can be designed according to the usage of the FORSpatial system. The rich combination of letters and shapes makes it

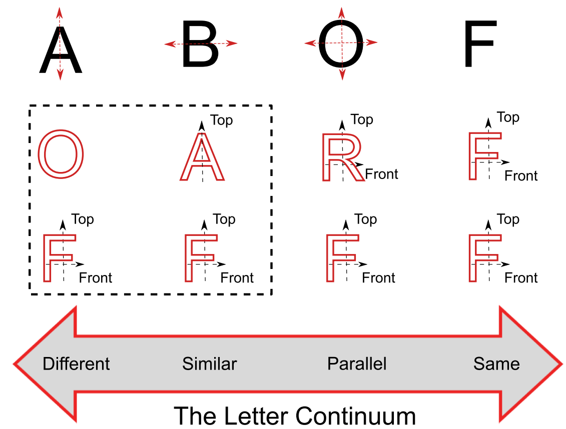


Fig. 3. Letter continuum and sample letters.

possible to adjust the difficulty of the test and learning application in the system, making it suitable for primary, secondary, college students, etc. In this work, we have designed 24 unfolded 2-D shapes in the FORSpatial training application, including eight shapes with one, two, and three letters.

1) Shapes: In FORSpatial, the shapes are created based on basic shapes, such as triangles, squares, and circles. The sample shapes in the test [38] inspired the creation of some shapes in our work. As shown in Fig. 2(b), here are some shape examples in the current FORSpatial application, including triangles, squares, and their combination. The mental folding difficulty varies among shapes.

2) Letters: We regard letters as patterns. In FORSpatial, we put one, two, or three same letters on the surface of shapes to increase the difficulty of spatial thinking. The letters can have zero, one, or two lines of symmetry. The lines of symmetry are shown in Fig. 3 in dotted lines with the arrow. A letter having one line of symmetry can be divided into two congruent parts, which are mirror patterns of each other. For example, A and B have one line of symmetry. The A has a vertical line, while the B has a horizontal line. The O has two lines. The F and R have zero lines of symmetry and cannot be folded in half in any way with the parts matching up. The spatial visualization and reasoning difficulty vary among these letters. An increased number of letters also increases the difficulty.

As shown in Fig. 3, the letter continuum gives the similarity between two letters. Concerning the number of lines of symmetry, the O is quite different from F. R is similar to F, and F is the same as F. Since the asymmetry of the letters increases the difficulty of the test, we use F and R in the current instance of FORSpatial system. We use R in the test, and each item has one, two, or three letters on the surface of eight shapes. For the practice with feedback in the application, we use F to create questions.

3) Principles of Questions/Tasks Set up in FORSpatial Test/Application: If users want to answer correctly, they need to succeed in the following aspects.

1) Fold 3-D object from 2-D flat figures, do some simple rotation, identify this object from any angle, and associate the faces and sides of the object to the parts of the shape in mind.

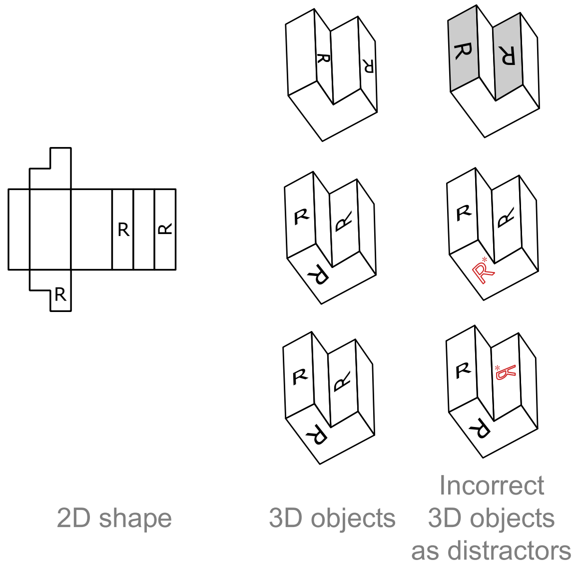


Fig. 4. Sample of 3-D object distractors. (Top) Letter appears on the wrong face. (Middle) Letter has an incorrect orientation on the face. (Down) Letter has incorrect handedness.

2) Identify that the letter is on which face.
3) Recognize that the letter's orientation relative to the face.

We design FORSpatial to test and enhance these aspects. In FORSpatial, the test and practice are yes–no questions, as shown in Fig. 2(b). Users need to determine each given item on the right made from an unfolded 2-D shape and choose yes or no. The distractors (incorrect answer choices) are made plausible, which are designed based on the following principles (see Fig. 4).

1) Principle 1: The letter appears on the incorrect face.
2) Principle 2: The letter has an incorrect orientation on the face.
3) Principle 3: The letter has incorrect handedness.

Through these exercises, users can reflect instantly with immediate feedback about the correctness of their choices. They can verify whether their strategies are helpful and correct. Besides, they can check whether mental representation is consistent with the visualized objects via an accurate visualization.

4) FORSpatial Training Application: The application provides 3-D visualization, direct manipulation of virtual objects, and instant exercises with feedback. It has a guided tour and scenes for manipulating and practicing. The guided tour explains how to use this application in the form of a video.

Through the mouse, users can easily interact with this desktop VR application. The current application contains 24 scenes, and each scene has the UI of navigation, learning area, and four practices with feedback (see Fig. 5). From one scene to another, users move forward or backward, controlling the learning path at their own pace. They can fold the 2-D shape to its 3-D form or unfold the shape in the interactive space via OPEN and CLOSE buttons in each scene. They can also watch the animated opening and closing process from 2-D to 3-D or vice versa. Concerning manipulation, users can drag and rotate the 2-D shape or its 3-D form to any direction to observe.

The practice with feedback supports users to reflect, fix errors in the cognitive structure in mind, and learn with goals. Otherwise, they would lack the purpose and quickly lose interest by simply playing with the 3-D model in the learning area.

In addition, FORSpatial logs users' interactions, including the time spent in each scene, the type of interaction behavior and its frequency in each scene, when the behavior occurs, when the user does the exercise, and whether the exercise is correct. The logging format is.txt, which can be easily loaded into EXCEL or read by Python applications for further analysis. This logging function assists researchers and experimenters in analyzing the behaviors of users.

5) FOR Versus FAR: Concerning the letter used in the interactive 3-D area, we have two versions, FOR and FAR (see Fig. 5). In the FOR version, we use O, which is different from the F, not having an orientation. In the FAR version, we use an arrow having the same orientation as the F. Since the arrow is similar to A, we called this version FAR.

The O and the arrow (stands for letter A) are different from the letter F in the continuum. On the one hand, we explore whether the pattern of the letter, namely, the orientation, can promote learning. On the other hand, we investigate whether the similarity between patterns in the learning area and practice area would make the training different.

B. Implementation

We develop 3-D objects used in the FORSpatial with 3Ds Max 2018 [39] and make the application with unity 2018.2.17f1 for PC [40]. The current version is rendered to the connected monitor in the desktop-3-D condition. The 3-D objects are freely available on the Unity asset store [41]. Users can install FORSpatial on PC and Mac.

The FORSpatial system allows for the training on a desktop VR, fully immersive VR environment, and web browser VR, supporting both online and offline training. In this article, we present the user study on the desktop VR version.

IV. User Study

To evaluate and examine the FORSpatial training application, we organized a user study with 49 participants and explored the following research questions (RQs).

1) RQ1: Can FORSpatial improve spatial ability?
2) RQ2: What do learners think about FORSpatial?
3) RQ3: What aspects of FORSpatial training application design need improvement?

A. Participants and Procedure

We recruited 49 participants, including 27 females and 22 males. Participants were studying at the university and compensated for their time. They aged between 19 and 28 years (M = 22.51, SD = 2.02). All participants had a normal or corrected vision.

We employed a 2 × 2 mixed design with between-subject groups, namely, an experimental group and a control group with and without training by interacting with FORSpatial. The

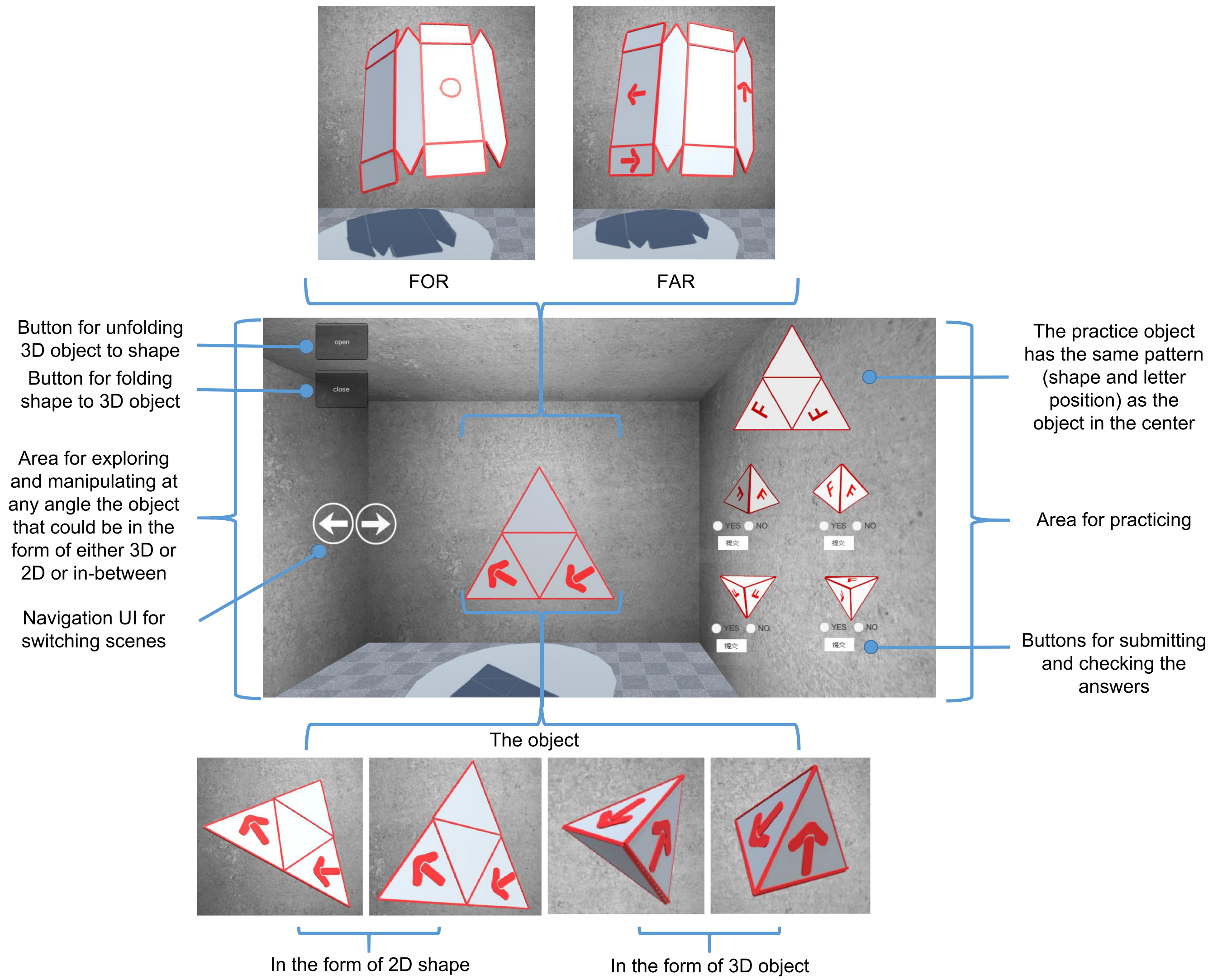


Fig. 5. Interactive space of one scene in FORSpatial.

testing phases were based on a within-subject design using the pretest/post-test comparison. In this experiment, 24 participants were assigned to the experimental group and 25 participants to the control group (see Fig. 6).

Participants in the experimental group completed tasks of three phases: the pretest, training, and post-test phases. In the pretest phase, participants filled out the spatial ability task test, at most 30 min, trying to answer correctly at the fastest. We also collected the demographic information, pretest, matrix questionnaire, and short interview data. After about one week on average, participants continued to participate in the training and post-test phase. In the training phase, participants learned with FORSpatial in a desktop VR environment, where they interacted with 3-D objects and practiced with designed test items (see Fig. 7). We used a designed combination of the letter R and shapes in the training application. To investigate whether the orientation of the letter on the 3-D model of the learning area affects the training, we dynamically assigned the participants to either FOR training group or FAR group, keeping the balance of the gender and performance. Each group had 12 participants. Both FOR and FAR were FORSpatial applications, the tasks in which were the same. The only difference was the letters used in the learning area. In the FOR group, we used the letter O, offering nonorientation. In the FAR group, we used the arrow offering orientation (see Section III-A3 for the interpretation of FOR and FAR). The training stage took about 30 min. After completing the training tasks, we instructed participants to perform in the post-test phase. We collected the data of post-test, usability questionnaire, matrix questionnaire, and a short interview.

Participants of the control group were only involved in the pretest and post-test phases. This group performed the same pretest and post-test as the experimental group.

We first determined whether there was a difference in spatial ability among two groups. Since scores in the pretest were not normally distributed, we did a Mann–Whitney test. The preliminary result showed that pretest scores in the experimental group (Mdn=38.5) did not differ significantly from the control group (Mdn=36), $p > 0.05$. Therefore, there were no pre-existing differences among two groups in performing spatial tasks.

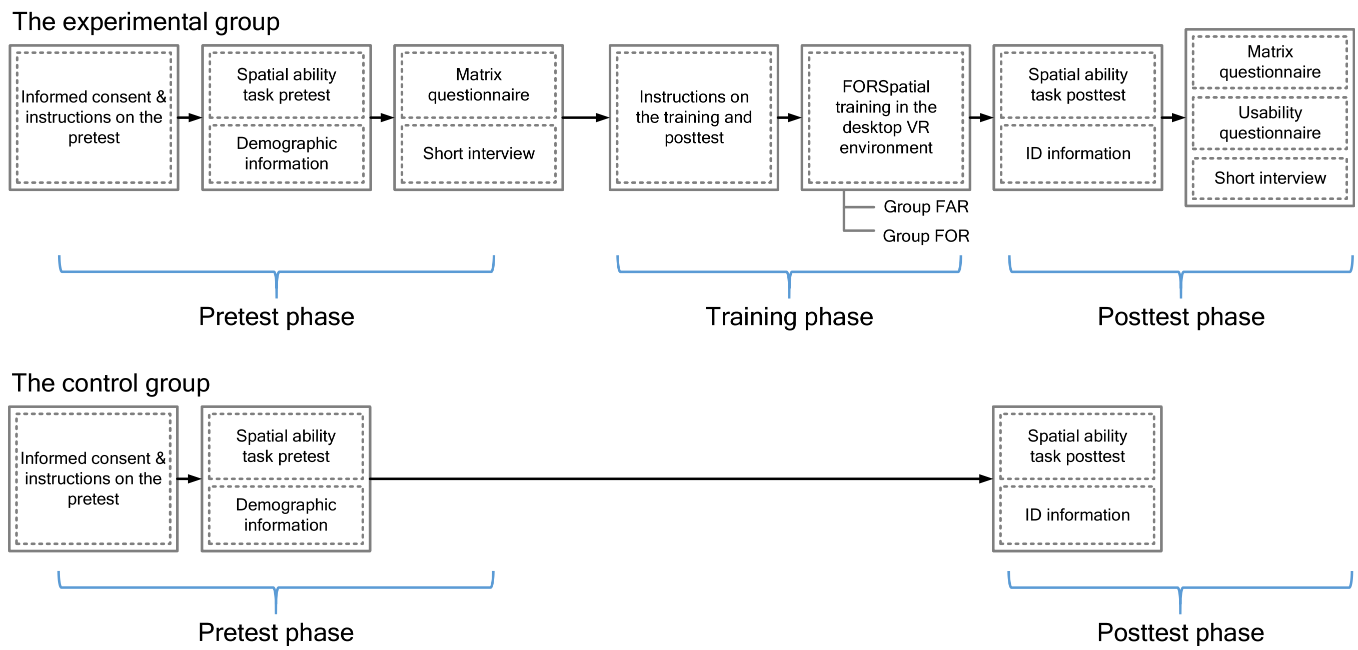


Fig. 6. Procedure of the experiment. The experimental group participated in three phases, while the control group did not learn with FORSpatial.

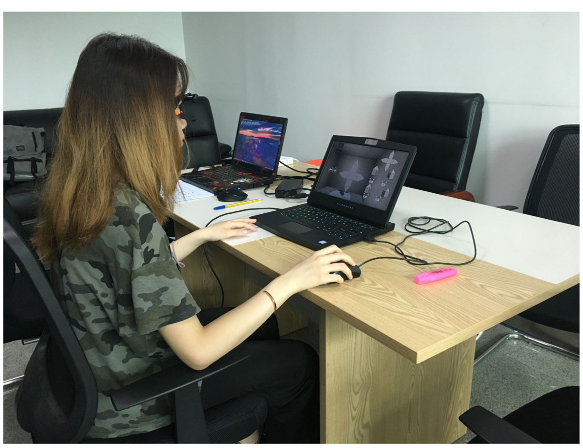


Fig. 7. Illustration of the experimental setup. One participant sat before the desk and interacted with FORSpatial via a mouse.

B. *Pretest, Post-Test, and Practice in the Training Application*

To explore whether pretest and post-test tasks are relevant to spatial ability, we conducted an online survey. We recruited 100 participants from two universities in varying majors. These participants were not the same in the experimental group of the experiment. The data of 96 participants (36 males and 60 females) aged 18 to 26 years (Mean=20.27, SD=2.07) were valid. We asked participants to perform a spatial visualization and rotation test online. They should finish our proposed 48-item pretest and the 20-item Purdue visualization of rotations (ROT) test [29]. ROT test has been used widely and proved as a valid measure of spatial ability. We did a correlation analysis by running Pearson's correlation to measure the relationship between FORSpatial task scores and ROT test scores. The result showed a significant relationship between the pretest and ROT test scores, r=0.66, $p < 0.05$. Therefore, the tasks in the pretest and post-test can reflect spatial ability.

The pretest and post-test have similar questions, based on shapes with the letter R. Both the pretest and post-test contain 48 items. The practice in the training application includes similar test items as in the pretest and post-test but based on shapes with letter F. Principles of the question and distractors used in pretest, post-test, and practice of training application were consistent. We did a reliability analysis to measure the consistency of our proposed test, with a coefficient Cronbach's α [42] of 0.841, indicating good overall reliability of this test.

C. *Measures and Subjective Opinions*

We collected and analyzed the test performance, matrix questionnaire, interaction logging, usability questionnaire, and the interview for the experimental group. The measures include score, test completion time (TCT), mental representation vote, interaction frequency, and duration.

1) Score measures the number of correct answers in the pretest or post-test, obtained through pretest and post-test.
2) TCT refers to the amount of the time (in seconds) that a participant spends to complete a pretest or post-test, collected through interaction logging.
3) Mental representation vote measures whether a participant can represent the shape and the letter in mind, obtained through matrix questionnaire.
4) Interaction frequency measures the number of click and drag actions, reflecting the frequency of rotating the 3-D object, unfolding the 3-D object to the 2-D flat shape, and folding 2-D to 3-D object. We obtained interaction frequency via interaction logging.

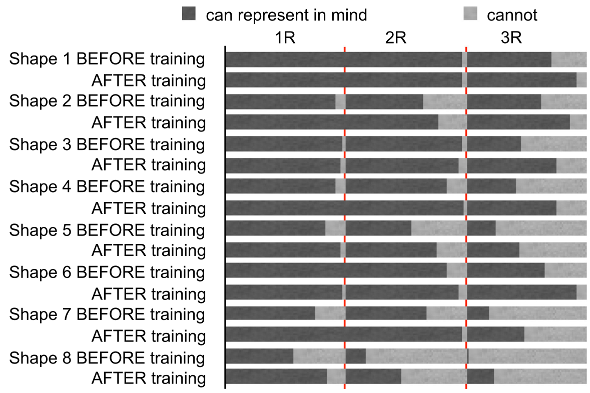


Fig. 8. Mental representation votes gathered from the matrix questionnaires in pretest and post-test phases.

5) Duration measures the amount of time (in seconds) that a participant uses, obtained through interaction logging.

We interviewed participants and asked them to describe their thinking process in the tests and reflect how they perform 3-D generation and rotation in mind and solve the problem. We also asked about what strategies they would adopt when failing to generate 3-D or the numbers of the letter increasing.

V. Results

A. Learning Effect

We tested the learning effect of training application from two aspects: objective and subjective aspects. Thus, results on learning effect include the following measures. We used these results to analyze whether FORSpatial can improve spatial ability:

1) scores and TCT in pretest/post-test;
2) participant's votes on mental representation in pretest/post-test;
3) scores and TCT between the high and low group in pretest/post-test.

Kolmogorov–Smirnov (K–S) test of observed values (differences between scores) and visual inspections of their histograms, normal Q-Q plots, and box plots showed that scores in pretest and post-test were not normally distributed. Therefore, we used the Wilcoxon signed ranks test to analyze data. For the control group, scores were not significantly different between the post-test (Mdn=35) and the pretest (Mdn=36), $z=-2.676$, $p > 0.05$, $r=-0.54$. However, for the experimental group, scores were significantly higher in the post-test (Mdn=42.5) than in the pretest (Mdn=38.5), $z=-4.07$, $p < 0.05$, $r=-0.59$. Since differences between TCT in the pretest and post-test were normally distributed, we used paired t-test. On average, participants spent significantly less TCT in the post-test (M=813.58 s, SD=221.57) than in the pretest (M=1491.08, SD=503.30), $t(23)= 6.40$, $p < 0.05$, $r=0.80$. Results of scores and TCT showed a significant pretest/post-test improvement in mental representation and spatial reasoning using the FORSpatial training application.

We explored how the training application of FORSpatial scaffolded 3-D representation in mind for participants. As shown in Fig. 8, light gray bars were getting shorter in the post-test than in the pretest for each shape. Therefore, by comparing votes whether participants could represent shapes with letters in mind (see Fig. 8) before and after the training, it showed that participants could mentally imagine more shapes after the training with FORSpatial. Besides, when compared the votes of different letters R, we found that as the number of letters R on the shape increases, it becomes more difficult for participants to represent the shape mentally (light gray bars were getting longer).

We also tested the performance improvement difference between high score and low score groups. We divided participants based on the median of the pretest (Mdn = 38.5). Participants with scores above 38.5 were assigned to the high score group, and participants below 38.5 were assigned to the low score group. K–S test showed that observed values were normally distributed. Thus, we used an independent t-test to explore the difference. High score group learners improved in the post-test (M=44.67, SD=2.10) compared with the pretest (M=42.56, SD=2.91). Low score group learners also performed better in the post-test (M=40.00, SD=1.91) compared with the pretest (M=32.00, SD=4.41). On average, the low score group achieved more (M=8.00, SE=1.11) than the high score (M=2.08, SE=0.48) group in terms of performance improvement. This difference was significant $t(22)=-4.87$, $p < 0.05$. Concerning TCT, high score group learners spent less time in the post-test (M=784.50, SD=232.15) compared with the pretest (M=1681.25, SD=532.96). Low score group learners were also faster in the post-test (M=842.67, SD=216.59) than in the pretest (M=1300.92, SD=408.30). On average, the TCT time difference between the pretest and post-test in the high score group (M=896.75, SE=139.38) was more significant than that of the low group (M=458.25, SE=136.98). This difference was significant $t(22)=2.24$, $p < 0.05$. Therefore, the low score group benefited more than the high score group after the training, which was consistent with the results obtained in [11].

B. Usability

We tested the usability of the training application via collecting users' attitudes toward the guided tour, 3-D interaction, layout, navigation, and learning. The questions asked were listed in the lower part of Fig. 9. We used these results to analyze what learners think about FORSpatial. The median of each category is as shown in Fig. 9. All median were above 4 except the mouse interaction, indicating that FORSpatial had good usability.

We asked participants to express their opinions. A total of 17 of 24 participants reported that learning was a little complicated due to the limited mouse interaction. They expressed that the mouse interaction was not very flexible, but they could rotate in the direction they wanted. For example, No. 11 expressed that *I need to click several times to rotate the object to the desired direction.* No. 16 reported that *I could achieve my goal when I used the mouse to rotate, but it is not that flexible.* Therefore, we found that FORSpatial had good usability, and the learning was useful. However, there was still room to improve mouse interaction.

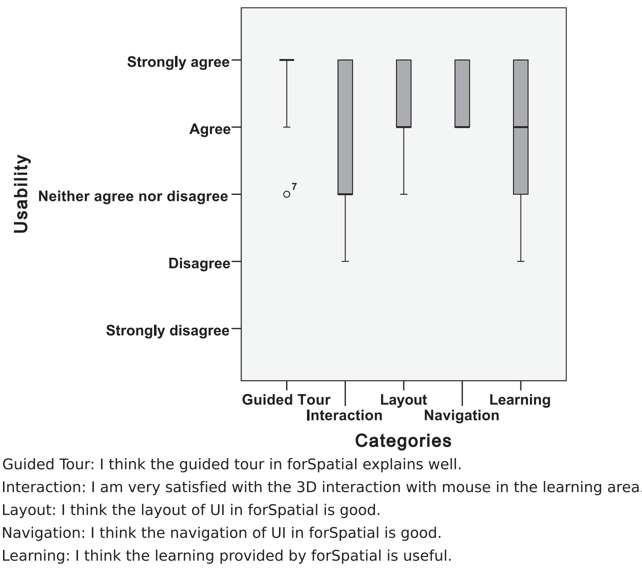


Fig. 9. Usability boxplot.

By coding comments on expressing the 3-D learning experience, we found that the 3-D interaction in the learning area also promoted spatial imagination. Ten participants reported that their spatial imagination had been augmented, including perceiving more 3-D shapes and positions, imaging more details, and enhanced rotating skills in the post-test. For example, No. 1 expressed that *After learning I found that when rotating the 3D object in mind, the letter started to appear on the surface that I cannot succeed to imagine in the pretest.* No. 17 commented that *I can rotate the object with letters in mind in many directions, which was impossible in the pretest.* No. 10 reported that *I can find the details in mind of shape 8, but I cannot do it in the pretest.*

C. Individual Learning Differences

We also collected data on learner individual differences and obtained behavior analysis results. We explored the following questions.

1) Whether the pattern on the 3-D model of the learning area affects the learning?
2) What is the relationship between the number of letters and learning difficulty?
3) What strategies do participants adopt to solve spatial problems?

To investigate whether the pattern, namely, the orientation of the letter, on the 3-D model of the learning area affects the learning, we tested differences between FOR and FAR in supporting learning. We used the Mann–Whitney test since the K–S test showed that observed score values were not normally distributed. Median is more appropriate than the mean for nonparametric tests. Therefore, we reported the median here. The increase of the score in the post-test in the FOR learning group (Mdn = 3.5) did not differ significantly from the one in the FAR learning group (Mdn = 3.5), U=69.00, $z=-0.18$, r=−0.04, $p > 0.05$. The K–S test showed that observed TCT values were normally distributed. Thus, we used the independent t-test. The TCT time between the pretest and post-test in the FOR learning group (M = 646.83, SE = 149.52) did not differ significantly from the one in the FAR learning group (M = 708.17, SE = 156.21), $t(22) = -0.284$, $p > 0.05$. Therefore, there was no significant difference between FOR and FAR on performance improvement and TCT reduction.

However, participants' attitudes toward O and Arrow's role in scaffolding learning with FOR and FAR were different. Half participants in group O (six persons) indicated that the letter O is horizontally and vertically symmetrical, making the learning difficult. For example, No. 20 expressed that *O has no direction and I advice change it to the letter like R.*

We investigated the relationship between the number of letters and learning difficulty. We calculated the interaction frequency and the duration with one F (1F), two F (2F), and three F (3F) for each participant. The K–S test showed that observed values of interaction frequency were normally distributed. We used one-way repeated-measures ANOVA to test differences between three groups: 1F (M=36.46, SD=29.17), 2F (M=56.00, SD=44.08), and 3F (M=52.58, SD=40.12). Results showed that the interaction frequency in learning was significantly affected by the number of letters, $F(1.89, 43.50) = 7.73$, $p < 0.05$. The interaction frequency significantly changed when participants were learning in conditions 1F, 2F, and 3F. Observed values of duration were normally distributed. We used one-way repeated-measures ANOVA. On average, it cost the longest time to learn with 3F (M=707.38, SD=284.10), followed by 2F (M=641.67, SD=301.50), and it cost the least time to learn with 1F (M=435.58, SD=201.50). Results showed that the duration of learning was significantly affected by the number of letters, $F(1.56, 36.0) = 44.68$, $p < 0.05$. Therefore, there were significant differences in interaction frequency and time in conditions 1F, 2F, and 3F. An increase in the number of letters raises the difficulty of the task.

Finally, we explored learners' behavioral differences in using strategies for solving spatial problems. Based on participants' responses on strategies, results showed that the participants could be categorized into mental imagery, spatial analytic, and poor spatial analytic users.

1) Mental imagery user (high performance; mainly relying on imaging strategies to solve problems). For example, No. 1 expressed that *Except shape two with 3R, the rest can be imaged in my mind. But I can only rotate part of the objects, and I can rotate the object in the left-right direction (horizontal direction) but cannot rotate in the up-down direction (vertical direction).*No. 23 expressed that *The shape with letters can be easily generated in my mind naturally and I then can find R on the object.* No. 5 expressed that *I can directly image objects.*
2) Spatial analytic user (moderate performance; using fewer imagination strategies, mainly relying on spatial cognitive-analytic strategies; strategies are used adequately). For example, No. 10 expressed that *I can image folded shapes, but need to think the details on the shapes. It's hard to image these details.* No. 21 expressed that *I can image only simple objects, but for slightly more difficult ones, in fact, I mainly looked at the relative position*

of the letter R and the edge of the corresponding figure, repeatedly folding and unfolding in the brain, and compared them.

3) Poor spatial analytic user (low performance; using fewer imagination strategies, mainly relying on spatial analysis strategies; Inadequate use of strategies). For example, No. 12 expressed that *I can image the simple figures with just one R. When the number of R increases, I fixed one R and checked another's position, and compared the R direction. Sometimes, I cannot make it clear what the correct direction of three R on the surface.* No. 17 expressed that *"I cannot imagine the 3D generation when R increases. When I failed, I started to use the position of R and the edge or guess the answer.*

Participants used spatial analytic strategies differently. The strategies include checking the direction of R (the direction of the top, bottom, and front of R), the relative position of several R, the relationship of R and shape, the position between R and edge, the position between R and angle, and whether reversed R appears.

From comments on the training process, we found that their 3-D learning experience was similar. For simple shapes, they directly imagined the test area's question and worked it out in mind. For complex shapes, they first tried to imagine the learning area's O or Arrow as F. Then, they fold the flat shape into a 3-D object and thought about the questions in the test area, repeatedly rotating, unfolding, and folding this object. For example, participant No.7 reported that *I first folded 2D shape and made it into a 3D object (in the learning area), rotated it to the same position as the test item with F in the right test area, then I observed, imaged and thought about it*. Participant No. 5 stated that *I matched F to the position of O and remembered this position, then folded it up and rotated it*.

VI. Discussions and Limitations

A. Can FORSpatial Improve Spatial Ability?

The answer is yes. Results from the comparison between experimental and control groups showed a significant improvement in training. Via training, participants gained mental representation, rotation, and spatial reasoning skills. From comments on the 3-D learning experience, we found that the 3-D interaction in the learning area promoted mental representation. Moreover, concerning the experimental group, the low score group achieved more than the high score group on performance on average. TCT time difference between the pretest and post-test in the high score group was greater than that of the low group. All these showed that the low score group benefited more than the high score group through FORSpatial training.

B. What Do Learners Think About FORSpatial?

From the learning effects and usability results, we found that FORSpatial can be used to train spatial ability, and its current version has good usability. The score of usability and comments from participants indicated FORSpatial application is useful and can support spatial skills learning. However, there is still room for improvement in rotating 3-D objects with mouse interaction.

C. How to Improve the Design of the FORSpatial Training Application?

There is no significant difference between FOR and FAR on scores and TCT. But from comments on O and Arrow's role of scaffolding learning, we found that some participants in group FOR complained O and advised a letter similar to R in training. In the current version, participants were assigned to the FOR group or the FAR group for learning, and there might be cases where the participant did not adapt to the current group. FORSpatial should provide an option for users to choose the preferred pattern to practice and play.

Based on participants' interaction logging results, we found that the difficulty of 1F, 2F, and 3F increases in order. This finding indicates that the testing and training difficulty level can be adjusted through a customized combination of shapes and letters.

Furthermore, we found that users used increased strategies when patterns became more complex by coding interview comments. Responses showed that FORSpatial can provide a scaffold for spatial thinking, supporting visualization and rotation operation outside the mind to foster internal reflection. However, it is inadequate to use 3-D visualization and exercises with feedback for users. FORSpatial should be embedded with dynamic interactive visual demonstrations of the reasoning process and question explanations.

VII. Conclusion and Future Work

In this work, we propose FORSpatial to improve spatial visualization and mental rotation for users. The main contributions are as follows. First, the scheme and principles can be used for creating tasks to train two primary spatial skills, and the difficulty of tasks can be customized and adjusted. FORSpatial is downloadable and free to use, and its components are made publicly available on the Unity asset store. Therefore, other researchers are encouraged to use components to create their training tools to enhance spatial ability. Second, the FORSpatial system considers both the interactivity and pedagogical aspects. At present, pedagogical consideration in designing and developing VR products is still lacking. Very few studies have explored the rationale of design and learning experience. This study offers critical insights into integrating a pedagogical approach into VR to create a practical educational application to reach the desired learning outcomes. Third, results from the user study verified the usability and learning effects of FORSpatial, and the qualitative data revealed how FORSpatial supports spatial thinking. Findings of this work have contributed to the potential of the virtual environment in spatial skills training.

However, current FORSpatial did not consider individual differences in spatial abilities and provide personalized learning paths. In our future work, we will investigate and propose a novel intelligent spatial ability training system, providing fine-grained interaction, personalized training support, and automatic analysis.

ACKNOWLEDGMENT

The authors would like to thank Y. Wang, Y. Peng, and L. Deng for their help and support in making this work possible.

REFERENCES

[1] D. H. Uttal, D. I. Miller, and N. S. Newcombe, "Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics?," *Curr. Directions Psychol. Sci.*, vol. 22, no. 5, pp. 367–373, 2013, doi: 10.1177/0963721413484756.

[2] M. Harle and M. Towns, "A review of spatial ability literature, its connection to chemistry, and implications for instruction," *J. Chem. Educ.*, vol. 88, no. 3, pp. 351–360, 2011, doi: 10.1021/ed900003n.

[3] M. Hegarty and D. A. Waller, "Individual differences in spatial abilities," in *The Cambridge Handbook of Visuospatial Thinking*. New York, NY, USA: Cambridge Univ. Press, 2005, pp. 121–169, doi: 10.1017/CBO9780511610448.005.

[4] J. Wai, D. Lubinski, and C. P. Benbow, "Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance," *J. Educ. Psychol.*, vol. 101, no. 4, pp. 817–835, 2009, doi: 10.1037/a0016127.

[5] J. M. Lakin and J. Wai, "Spatially gifted, academically inconvenienced: Spatially talented students experience less academic engagement and more behavioural issues than other talented students," *Brit. J. Educ. Psychol.*, vol. 90, no. 4, pp. 1015–1038, 2020, doi: 10.1111/bjep.12343.

[6] J. B. Carroll, *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York, NY, USA: Cambridge Univ. Press, 1993, doi: 10.1017/CBO9780511571312.

[7] I. D. Cherney, "Mom, let me play more computer games: They improve my mental rotation skills," *Sex Roles*, vol. 59, no. 11, pp. 776–786, 2008, doi: 10.1007/s11199-008-9498-z.

[8] M. Yu, Y.-J. Liu, G. Zhao, C. Yu, and Y. Shi, "Spatial ability improvement by tangible interaction: A case study with EasySRrobot," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. LBW013:1–LBW013:6, doi: 10.1145/3170427.3188545.

[9] H.-C. Wang, C.-Y. Chang, and T.-Y. Li, "The comparative efficacy of 2D- versus 3D-based media design for influencing spatial visualization skills," *Comput. Hum. Behav.*, vol. 23, no. 4, pp. 1943–1957, 2007, doi: 10.1016/j.chb.2006.02.004.

[10] J. S.-K. Chang *et al.*, "Evaluating the effect of tangible virtual reality on spatial perspective taking ability," in *Proc. 5th Symp. Spatial User Interact.*, 2017, pp. 68–77, doi: 10.1145/3131277.3132171.

[11] E. A.-L. Lee and K. W. Wong, "Learning with desktop virtual reality: Low spatial ability learners are more positively affected," *Comput. Educ.*, vol. 79, pp. 49–58, 2014, doi: 10.1016/j.compedu.2014.07.010.

[12] S. Oberdörfer, D. Heidrich, and M. E. Latoschik, "Usability of gamified knowledge learning in VR and desktop-3D," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–13, doi: 10.1145/3290605.3300405.

[13] C. A. Cohen and M. Hegarty, "Visualizing cross sections: Training spatial thinking using interactive animations and virtual objects," *Learn. Individual Differences*, vol. 33, pp. 63–71, 2014, doi: 10.1016/j.lindif.2014.04.002.

[14] T. A. Hays, "Spatial abilities and the effects of computer animation on short-term and long-term comprehension," *J. Educ. Comput. Res.*, vol. 14, no. 2, pp. 139–155, 1996, doi: 10.2190/60Y9-BQG9-80HX-UEML.

[15] A. Rafi, K. Samsudin, and A. Ismail, "On improving spatial ability through computer-mediated engineering drawing instruction," *Educ. Technol. Soc.*, vol. 9, no. 3, pp. 149–159, 2006.

[16] M. J. Contreras, R. Escrig, G. Prieto, and M. R. Elosúa, "Spatial visualization ability improves with and without studying technical drawing," *Cogn. Process.*, vol. 19, no. 3, pp. 387–397, 2018, doi: 10.1007/s10339-018-0859-4.

[17] S. Jang, J. M. Vitale, R. W. Jyung, and J. B. Black, "Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment," *Comput. Educ.*, vol. 106, pp. 150–165, 2017, doi: 10.1016/j.compedu.2016.12.009.

[18] D. Liu, K. K. Bhagat, Y. Gao, T.-W. Chang, and R. Huang, "The potentials and trends of virtual reality in education," in *Virtual, Augmented, and Mixed Realities in Education*. Singapore: Springer, 2017, pp. 105–130, doi: 10.1007/978-981-10-5490-7_7.

[19] A. Dünser, K. Steinbügl, H. Kaufmann, and J. Glück, "Virtual and augmented reality as spatial ability training tools," in *Proc. 7th ACM SIGCHI New Zealand Chapter's Int. Conf. Comput.-Hum. Interact., Des. Centered HCI*, 2006, pp. 125–132, doi: 10.1145/1152760.1152776.

[20] H.-H. Liou, S. J. H. Yang, S. Y. Chen, and W. Tarng, "The influences of the 2D image-based augmented reality and virtual reality on student learning," *J. Educ. Technol. Soc.*, vol. 20, no. 3, pp. 110–121, 2017.

[21] J. S.-K. Chang, A. Doucette, G. Yeboah, T. Welsh, M. Nitsche, and A. Mazalek, "Keep the ball rolling: Designing game-based tangible vr for spatial penetrative thinking ability," in *Proc. Des. Interactive Syst. Conf.*, 2019, pp. 215–226, doi: 10.1145/3322276.3322280.

[22] C. J. Chen, S. C. Toh, and W. M. Fauzy, "The theoretical framework for designing desktop virtual reality-based learning environments," *J. Interactive Learn. Res.*, vol. 15, no. 2, pp. 147–167, 2004.

[23] H.-M. Huang, U. Rauch, and S.-S. Liaw, "Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach," *Comput. Educ.*, vol. 55, no. 3, pp. 1171–1182, 2010, doi: 10.1016/j.compedu.2010.05.014.

[24] M. Hegarty, "Components of spatial intelligence," in *The Psychology of Learning and Motivation*, vol. 52. San Diego, CA, USA: Academic, 2010, pp. 265–297, doi: 10.1016/S0079-7421(10) 52007–3.

[25] M. C. Linn and A. C. Petersen, "Emergence and characterization of sex differences in spatial ability: A meta-analysis," *Child Develop.*, vol. 56, no. 6, pp. 1479–1498, 1985.

[26] R. Ekstrom and H. Harman, *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ, USA: Educ. Testing Service, 1976.

[27] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, no. 3972, pp. 701–703, 1971, doi: 10.1126/science.171.3972.701.

[28] S. G. Vandenberg and A. R. Kuse, "Mental rotations, a group test of three-dimensional spatial visualization," *Perceptual Motor Skills*, vol. 47, no. 2, pp. 599–604, 1978, doi: 10.2466/pms.1978.47.2.599.

[29] G. M. Bodner and R. B. Guay, "The Purdue visualization of rotations test," *Chem. Educator*, vol. 2, no. 4, pp. 1–17, 1997, doi: 10.1007/s00897970138a.

[30] M. C. Harrington, "Empirical evidence of priming, transfer, reinforcement, and learning in the real and virtual trillium trails," *IEEE Trans. Learn. Technol.*, vol. 4, no. 2, pp. 175–186, Apr.–Jun. 2011, doi: 10.1109/TLT.2010.20.

[31] Y. Zhou, S. Ji, T. Xu, and Z. Wang, "Promoting knowledge construction: A model for using virtual reality interaction to enhance learning," *Procedia Comput. Sci.*, vol. 130, pp. 239–246, 2018, doi: 10.1016/j.procs.2018.04.035.

[32] E. Gazit, Y. Yair, and D. Chen, "The gain and pain in taking the pilot seat: Learning dynamics in a non immersive virtual solar system," *Virtual Reality*, vol. 10, no. 3, pp. 271–282, 2006, doi: 10.1007/s10055-006-0053-3.

[33] Arloon geometry, Accessed: Aug. 2, 2021. [Online]. Available: http://www.arloon.com/apps/arloon-geometry/

[34] Complete anatomy, Accessed: Aug. 2, 2021. [Online]. Available: https://3d4medical.com/

[35] W.-Y. Hwang and S.-S. Hu, "Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving," *Comput. Educ.*, vol. 62, pp. 308–319, 2013, doi: 10.1016/j.compedu.2012.10.005.

[36] R. C. Clark and R. E. Mayer, *e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*, 3rd ed. San Francisco CA, USA: Pfeiffer, 2011.

[37] J. Hattie, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London, U.K.: Routledge, 2008.

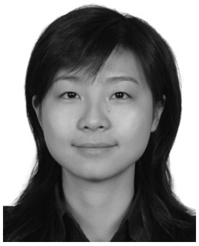
[38] J. Barrett and G. Williams, *Test Your Own Aptitude*. London, U.K.: Kogan Page, 2003.

[39] "AUTODESK Software 3Ds Max," Accessed: Aug. 2, 2021. [Online]. Available: https://www.autodesk.com/education/free-software/3ds-max

[40] "Unity developer tools," Accessed: Aug. 2, 2021. [Online]. Available: https://unity.com/developer-tools

[41] Unity Assetstore, Accessed: Aug. 2, 2021. [Online]. Available: https://assetstore.unity.com/

[42] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 3rd ed. London, U.K.: Sage, 2013.



Yun Zhou (Member, IEEE) received the Ph.D. degree in computer science from Ecole Centrale de Lyon, Écully, France, in 2012.

She is currently a Full Professor in education with Shaanxi Normal University, Xi'an, China. Her research interests include educational technology and human–computer interaction, including virtual reality, brain–computer interfaces, and innovative interaction for education.

Han Yang received the B.E. degree in education from the Xi'an University of Architecture and Technology, Xi'an, China, in 2018. She is currently working toward the master's degree in education with Shaanxi Normal University, Xi'an.

Her research interests include educational technology and human–computer interaction.

Tao Xu (Member, IEEE) received the Ph.D. degree in computer science from Ecole Centrale de Lyon, Écully, France, in 2013.

He is currently an Associate Professor with Northwestern Polytechnical University, Xi'an, China. His research interests include ubiquitous computing and human–computer interaction.

Shiqian Li received the B.A. degree in education from Shaanxi Normal University (SNNU), Xi'an, China, in 2018, where he is currently working toward the master's degree in education.

His research interests include educational technology and human–computer interaction.