Pilots' Spatial Visualization Ability Assessment Based on Virtual Reality

Mengdi Zhang; Meng Wang; Huimin Feng; Xunyuan Liu; Lihong Zhai; Xianrong Xu; Zhanguo Jin

BACKGROUND: The aim of this study was to investigate the effectiveness of the mental rotation test (MRT) based on virtual reality (VR) in predicting pilots' spatial visualization ability (SVA).

METHODS: Based on VR, 118 healthy pilots' SVA were evaluated by MRT. The pilot flight ability evaluation scale was used as the criterion of test validity. According to the scale score, pilots were divided into high, middle, or low spatial ability groups pursuant to the 27% allocation principle. Differences in reaction time (RT), correct rate (CR), and correct number per second (CNPS) of MRT between groups were compared. Correlations between scale scores and MRT scores were analyzed. RT, CR, and CNPS of MRT among different age groups and between genders were also compared.

RESULTS: The RT of the high spatial ability group was remarkably slower than that of the low spatial ability group (363.4 ± 140.2 s, 458.1 ± 151.7 s). The CNPS of the high spatial ability group was dramatically higher than that of the low spatial ability group (0.111 ± 0.045 s, 0.086 ± 0.001 s). There were no significant differences in RT, CR, and CNPS between different genders. Pilots in the 29–35 yr old age group had considerably slower RT than those in the 22–28 yr old age group (330.8 ± 140.3 s, 417.2 ± 132.7 s). Pilots in the 29–35 yr old age group had conspicuously higher CNPS than pilots in the 22–28 yr old age group (0.119 ± 0.040 s, 0.096 ± 0.036 s). All pilots' scale scores were positively correlated with CNPS ($r = 0.254$) and negatively correlated with RT ($r = -0.234$).

DISCUSSION: MRT based on VR has a good discrimination efficacy for SVA of pilots and is a good indicator for the SVA component measurement.

KEYWORDS: visual reality, pilot, visualization, mental rotation, spatial ability.

Zhang M, Wang M, Feng H, Liu X, Zhai L, Xu X, Jin Z. *Pilots' spatial visualization ability assessment based on virtual reality*. Aerosp Med Hum Perform. 2023; 94(6):422–428.

Spatial ability is the individual ability to understand and remember spatial relationships among objects. It is one of the most widely investigated domains of cognitive ability, which is a key element in determining how individuals perceive and interact with their surroundings. Spatial ability enables us to locate targets in space, perceive objects visually, and understand the two-dimensional (2D) and three-dimensional (3D) spatial relationships among objects and the environment.[7] In the course of flight, pilots first need to observe the changes of landmarks through vision, listen to the sound of the engine, and feel the change of acceleration through hearing and proprioception, and then receive, process, integrate, decide, and judge the spatial information to form the perception of the aircraft status. It is of great practical importance to study the basic theory and experimental methods of spatial ability, especially how to predict and train spatial cognitive ability.[11,16] In recent years, one of the key development directions of aviation medicine is how to apply the scientific and technological progress and research achievements of medicine, psychology, and other disciplines to the field of aviation medicine, so as to realize the best model of "human-machine-environment". Some major breakthroughs in the field of cognitive psychology will be applied to pilot selection, training, and medical assessment.

From the Air Force Clinical College, Anhui Medical University, the Fifth School of Clinical Medicine, Anhui Medical University, Hefei, Anhui 230032, China.

This manuscript was received for review in October 2022. It was accepted for publication in February 2023.

Address correspondence to: Zhanguo Jin, M.D., Ph.D., Air Force Clinical College, Anhui Medical University, the Fifth School of Clinical Medicine, Anhui Medical University, Hefei, China, 230032, and Aerospace Balance Medical Center, Air Force Medical Center, PLA, Beijing, 100142; ccjzg@126.com.

DOI: https://doi.org/10.3357/AMHP.6198.2023

The study of spatial ability originated in the late 19th century from Western scholars' research on human intelligence, dating back to the experimental investigation of imagination by Galton's. Since then, scholars have defined spatial abilities from multiple perspectives and developed a great variety of measures of spatial ability. So far, research on human cognitive abilities has proposed several kinds of spatial abilities, with partially overlapping distinctions and no complete agreement. Scholars generally believe that spatial orientation and spatial visualization are the stable elements of spatial ability.[14] Spatial visualization describes a variety of complex mental operations on spatial information, such as imagining the rotation of objects, the folding or unfolding of flat patterns, and the ability to change the relative positions of objects, which was one of the important components of spatial ability. Among the traditional measures of spatial ability, the mental rotation test is regarded as the most representative test content for estimating spatial visualization.[10]

With the rapid development of information technology, exploring spatial ability tests with ecology, simulation, and immersion has become an increasing research hotspot for scholars.[1,2,9] The advancement of VR has solved the shortcomings of inferior experience in traditional tests, fewer data analysis indicators, relatively single dimension, and poor visualization, providing a technical feasibility for cognitive psychological science research based on virtual environments.[6,20,25] In this study, based on the Shepard-Metzler Mental Rotation,[21] a spatial ability testing system with an immersive experience was developed by using 3D visualization modeling software and HTC Vive as an interactive experience platform. By comparing the test scores of pilot groups with different spatial abilities, we explored the differences in the mental rotation ability between pilot groups, the differences in spatial ability between different genders and among different age groups of pilots, and the effects of pilots' age and flight time on spatial ability. The results of the present exploratory study show the interest of taking into consideration different spatial abilities, different ages, and different genders in pilots to study the RT, CR, and CNPS based on a VR mental rotation task. In particular, this provides new theories and approaches for predicting pilots' spatial ability, improving the efficiency of pilot selection and training and the quality of medical assessment.

METHODS

Subjects

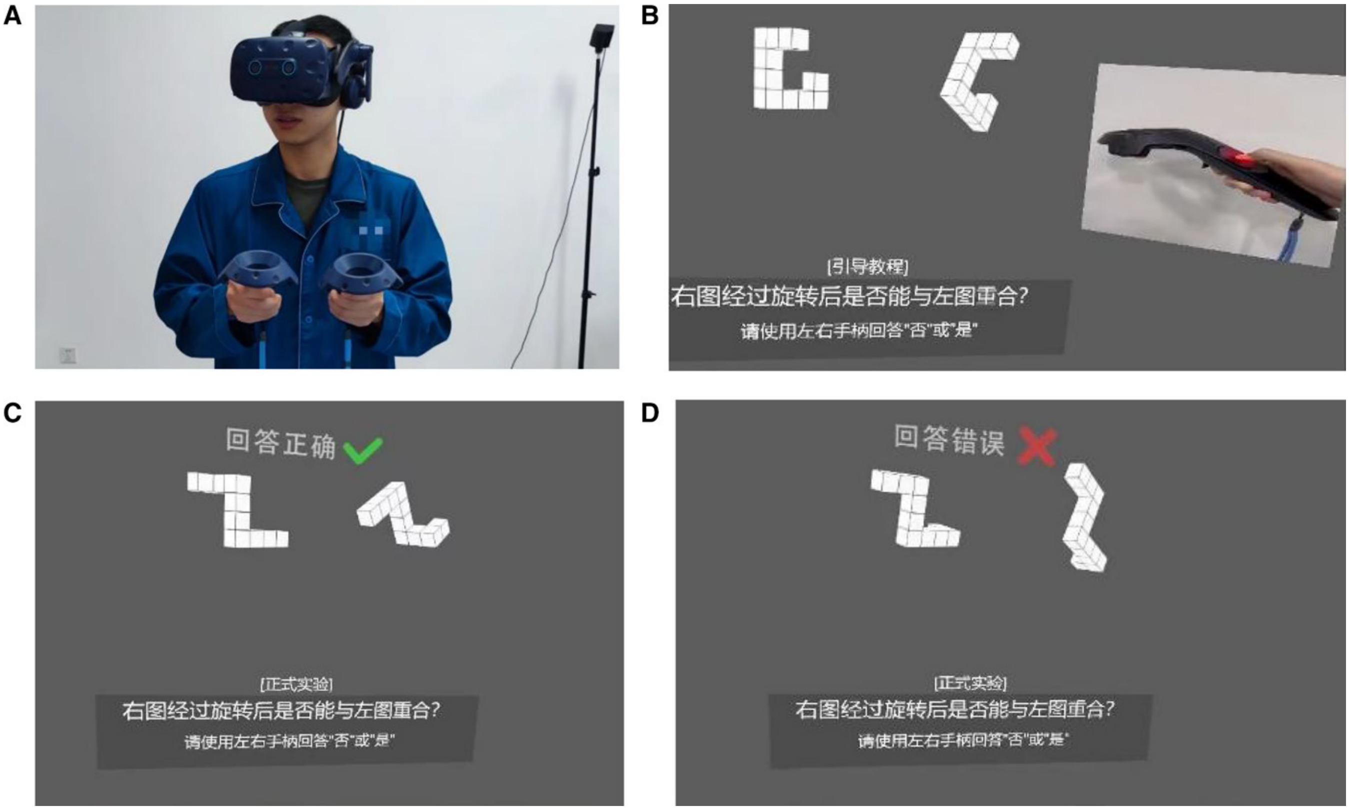
The project was approved by the Ethics Committee of the Air Force Special Medical Center (KT2022-16-SL02), and informed consent was obtained from each subject. The participants were 118 healthy pilots, 22–47 (28.91 ± 7.061) yr of age, with flight time of 260–4500 (1218.69 ± 1060.78) h, including 106 men and 12 women. All participants possessed a bachelor's degree, were right-handed, and had normal vision. All of them could operate the handle skillfully and had no dizziness when wearing the helmet.

Measures

Apparatus. This task was the modification of the original task by Shepard and Metzler.[21] This study used 3D visualization modeling software (3DMAX) to create 3D mental rotation stimuli consisting of 10 cubes, while using the HTC Vive as the interactive experience platform, combined with the Lighthouse base station and Unity3D, to complete the participants' immersive experience in the virtual scene (Air Force Special Medical Center, China) (**Fig. 1**). For subjects' comfort, we used adjustable headbands to fit the head-mounted display during the testing. A pair of operating handles were used to rotate the stimuli and answer a question during the testing as well. By using the handle and head-mounted display, participants could observe and operate the cubes as a first-person viewer. The VR testing was administered on a desktop workstation running Microsoft Windows 10 and equipped with a high-end graphics card. The testing was observed by a researcher through a 24-in light-emitting diode flat monitor, which showed the progress of the program.

Each group of stimuli consisted of eight original stimuli. A total of 96 different morphological stimuli were produced by its original or mirror-reversed version, which was rotated around the Z-axis respectively (30°, 60°, 90°, 120°, 150°, and 180°) clockwise or counter-clockwise from its vertical upright position (**Fig. 2**). One stimulus was randomly selected from the eight original stimuli via the Application Program Interface (API). The original mental rotation stimulus was placed on the left side of the virtual scene space, and the test stimulus was placed on the right side. First, if the right stimulus was the same as the left stimulus (just rotated at different angles), then they were identical. And if the subject could determine that the pair of cubes were identical, the response was correct. If the subject thought the pair of cubes were a mirror image, the response was incorrect. Second, if the right stimulus was formed by the mirror image of the left stimulus, then they were different. And if the subject could judge that the pair of cubes had a mirror-reversed relationship, the response was correct. If the subject thought the pair of cubes were the same (just rotated at different angles), the response was incorrect. There was a total of 48 questions in a test (composed of 24 identical versions and 24 mirror-reversed versions rotated around the Z-axis, respectively). Participants were required to judge, as quickly and accurately as possible, whether a pair of cubes had a mirror-reversed relationship or were identical (just rotated at different angles). The method was to press the VR handle trigger key with the right hand (indicating a "yes" answer, namely coincidence), or press the VR handle trigger key with the left hand (indicating a "no" answer, namely mirror-reversed relationship).

The principle of the questioning adhered to the following logic: four identical versions and four mirror-reversed versions were randomly selected from six angles (30°, 60°, 90°, 120°, 150°, and 180°) to ask questions. Of those questions, 24 contained a cube pair with an identical relationship and 24 contained a pair with a mirror-reversed relationship. Both the order of angles and the order of identical versions or mirror-reversed versions were random.



[引导教程]
右图经过旋转后是否能与左图重合?
请使用左右手柄回答"否"或"是"

回答正确
[正式实验]
右图经过旋转后是否能与左图重合?
请使用左右手柄回答"否"或"是"

回答错误
[正式实验]
右图经过旋转后是否能与左图重合?
请使用左右手柄回答"否"或"是"

Fig. 1. Practical demonstration of the mental rotation test. A) An example of using handle and head-mounted display for testing. B) Screenshot of the tutorial for mental rotation tasks. C) An example of a correct answer (as shown, the two test stimuli are mirrored, and if the subject makes a correct judgement, the answer is correct). D) An example of an incorrect answer (as shown, the two test stimuli coincide, and if the subject cannot make a correct judgement, the answer is incorrect).

VR-based mental rotation test. The subjects were tested individually in a quiet room. Before the experimental session, the participants were informed about the aim of the study, the procedure, their rights, and the possibility of stopping the experiment at any time they chose. Afterward, the participants signed a written informed consent form.

The basic process included four stages: 1) Preparation Stage: firstly, the notes related to this experiment were introduced,

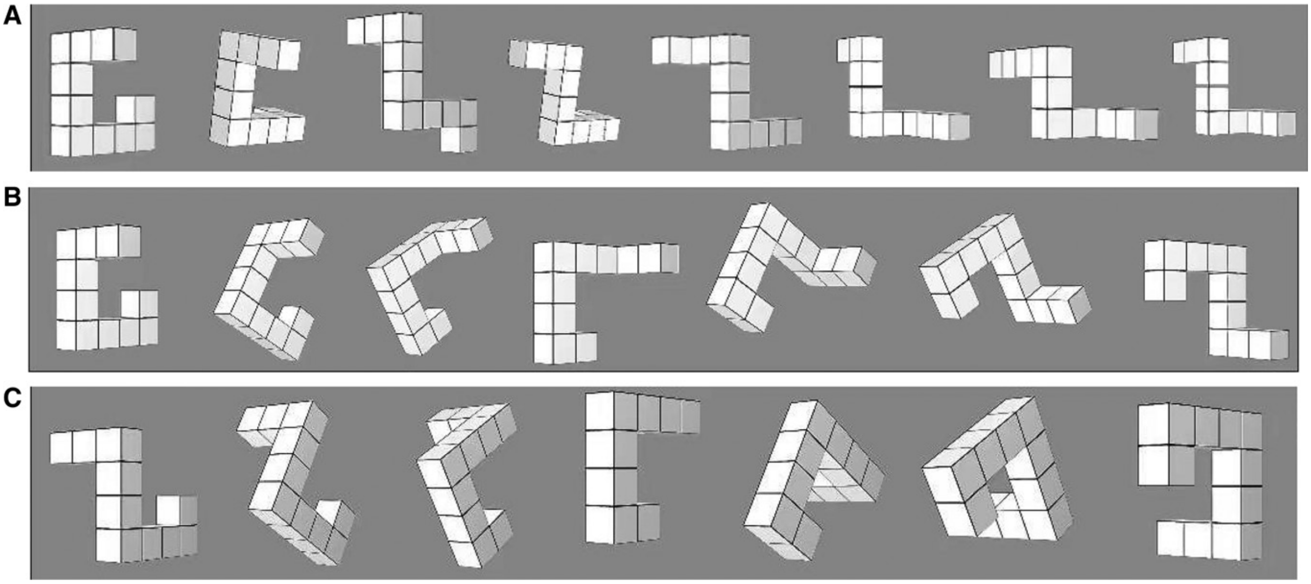


Fig. 2. Presentation of the stimulus of mental rotation test. A) Eight original stimuli. B) One of the eight original stimuli with different rotation angles (0°, 30°, 60°, 90°, 120°, 150°, and 180°). C) Mirroring the stimulus of Fig. 2B with different rotation angles (0°, 30°, 60°, 90°, 120°, 150°, and 180°).

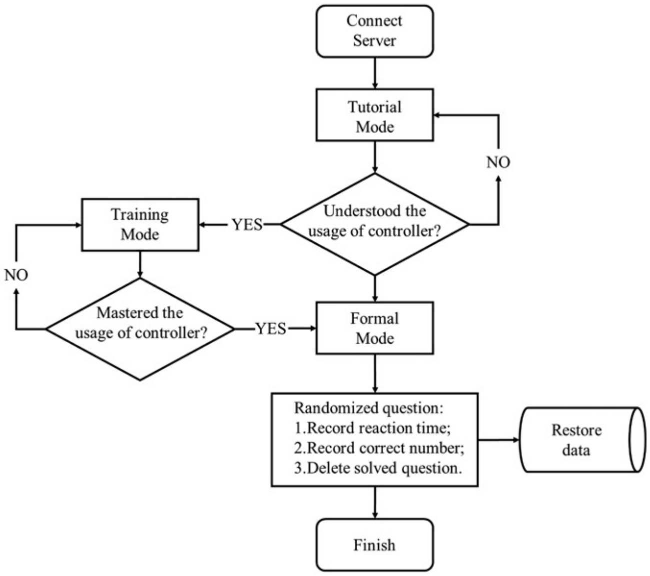


Fig. 3. Flow chart of the experiment.

then the participants read the paper instructions to clarify the task of this experiment; 2) Learning Stage: participants wore VR helmets to watch the guide tutorials, followed the instructions to complete the corresponding keystroke operations, and chose to enter the next stage or learned the guide tutorial stage again by operating the handle; 3) Practice Stage: after completing 10 exercises and thinking that they had mastered the operation keys, the participants entered the formal experiment; 4) Formal Experiment Stage: there were 48 questions, which required participants to respond as quickly and accurately as possible to judge whether the pair of cubes had a mirror-reversed relationship or were identical (just rotated at different angles). The reaction time and the correct number were stored in the local database simultaneously (**Fig. 3**).

Criterion test. The pilot flight ability evaluation scale was used (**Table I**),[12] and, according to the scale scores, the participants were divided into groups based on a 27% allocation principle. The upper and lower 27% rule was commonly used in item analysis based on Kelley's derivation.[8] He stated that 27% should be selected at each extreme to yield upper and lower groups that were most indubitably different with respect to the trait in question. First, the total scale scores were calculated and ranked; then, the first 27% of the total scores ranked from highest to lowest were used as the high spatial ability group (32 cases), the last 27% as the low spatial ability group (32 cases), and the rest as the middle spatial ability group (54 cases).

TABLE I. Pilot Flight Ability Evaluation Scale.

NO.	QUESTIONNAIRE OPTION
Q1:	How satisfied are you with the subjects you have completed during your actual flight training?
Q2:	How accurate do you think you are in judging various situations during the actual flight?
Q3:	In general, what is the speed and quality when you deal with in-air events?
Q4:	Basically, how well do you think you have accomplished for the flight subjects in all weather?
Q5:	How well do you think you did for your performance in completing the complex, advanced, and special flight subjects?
Q6:	What do you think about your flight training performance in a long time?
Q7:	How do you evaluate your overall reaction ability and reaction speed to the several instruments' information when compared to other pilots in terms of the relevant flight subjects?
Q8:	How well do you think about your ability to control the conditions of the aircraft while completing various flying subjects?
Q9:	How well do you orient the aircraft status by instrument in complex weather conditions?
Q10:	How do you think about your performance on the landing?
Q11:	How often have you experienced the flight illusion?
Q12:	Based on all the factors and your flying performance, what level do you think your flying skills are within the whole regiment?
Q13:	How well do you handle problems in complex situations?
Q14:	How well can you adapt to low altitude and high-speed flight conditions?
Q15:	How well are you able to overcome the illusion of flight once it has occurred?
Q16:	What is your emotion when completing difficult technical movements?
Q17:	How do you think about your speed of refitting and adaption when flying new types of aircraft compared to other pilots?
Q18:	Do you think you are suitable for flying in higher performance aircraft at your current skill level?
Q19:	Do you perform well in aerial target shooting compared to other colleagues?
Q20:	How do you feel about your comprehension skills compared to other colleagues when learning a new flying discipline or maneuver?
Q21:	How do you perform in imitating the new flight maneuvers?

Statistical Analysis

Data analyses were done with Statistical Package for the Social Sciences (SPSS) 23 package software. Data with a normal distribution were presented as mean ±SD, whereas the data with a nonnormal distribution were described with median, upper, and lower quartiles. The Freeman-Halton Test was used to analyze the categorical variables. One-way ANOVA was used when a quantitative data followed a normal distribution. The Mann-Whitney *U*-test and Kruskal-Wallis test were used when the quantitative data followed nonnormal distribution. The Spearman test was used for the correlation analysis. The criterion for significance was set at 0.05.

RESULTS

Among the 118 pilots, 32 cases were in the high spatial ability group, 54 cases in the middle spatial ability group, and 32 cases in the low spatial ability group. There were no significant differences in age, gender, or flight time among the three groups [$F(2, 115) = 2.894$, $P = 0.059$; $\chi^2 = 1.553$, $P = 0.535$; $F(2, 115) = 1.950$, $P = 0.147$].

There were significant differences in reaction time (RT) and correct number per second (CNPS) among the high, middle, and low spatial ability groups [458.1 ± 151.7,

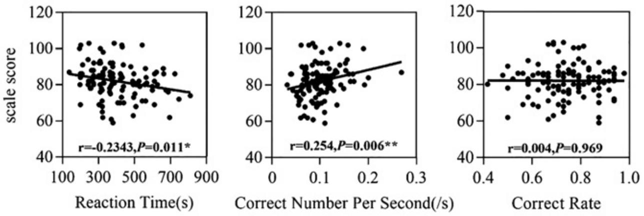


Fig. 4. Correlation analysis of the result of the mental rotation test and calibration test; ** $P < 0.01$; *$P < 0.05$.

386.2 ± 125.6, 363.4 ± 140.2, P = 0.029; 0.086 ± 0.028, 0.102 ± 0.034, 0.111 ± 0.045, P = 0.028, respectively; Kruskal-Wallis test]. RT of the high spatial ability group was remarkably slower than that of the low spatial ability group [363.4 ± 140.2 and 458.1 ± 151.7, respectively; Kruskal-Wallis test, P = 0.030]. CNPS of the high spatial ability group was notably higher than that of the low spatial ability group [0.111 ± 0.045 and 0.086 ± 0.028, Kruskal-Wallis test, P = 0.042].

In this study, there were 106 male pilots and 12 female pilots. There was no distinguishable difference in RT, CR, or CNPS between male and female pilots [394.8 ± 141.0 and 441.7 ± 137.5, P = 0.183; 0.748 ± 0.116 and 0.736 ± 0.161, P = 0.989; 0.102 ± 0.037 and 0.088 ± 0.032, P = 0.247; respectively; Mann-Whitney test].

Remarkable differences in RT and CNPS existed between pilots of different age groups [417.2 ± 132.7, 330.8 ± 140.3, 397.3 ± 151.5, P = 0.017 and 0.096 ± 0.036, 0.119 ± 0.049, 0.100 ± 0.036, P = 0.048, respectively; Kruskal-Wallis test]. Pilots in the age group of 29–35 yr had considerably slower RT than those in the age group of 22–28 yr [330.8 ± 140.3 and 417.2 ± 132.7, respectively; Kruskal-Wallis test, P = 0.013]. Pilots in the age group of 29–35 yr had a dramatically higher CNPS than pilots in the age group of 22–28 yr [0.119 ± 0.397 and 0.096 ± 0.035, respectively; Kruskal-Wallis test, P = 0.042].

Correlation analysis between the scale scores and mental rotation test scores showed that the flight ability of 118 pilots in this study was positively correlated with CNPS (r = 0.254, P = 0.006) and negatively correlated with RT (r = -0.234, P = 0.011) (**Fig. 4**).

DISCUSSION

In the study, using flight skills of pilots in flight practice as a validator, we examined the predictiveness of spatial visualization ability on flight skills with a newly developed immersive spatial mental rotation test and the possible effects of age and gender on test results in this population. The validity scale is closely related to the actual working practice and operating environment, and we found the flight skill level of pilots was significantly positively correlated with CNPS in the virtual mental rotation test, and negatively correlated with RT, which indicates that the spatial visualization test with virtual reality can offer good differentiation and prediction for individuals with different flight abilities. Many researchers have shown that there is a close relationship between operational tasks and cognitive abilities.[3,17] Especially in the aerospace field, strong spatial ability is the basis to ensure the correct flight status and spatial orientation of the pilots.[5] A pilot's workload requires him or her to react to spatial information and process it accurately in as short a time as possible. Therefore, a superior pilot must have excellent spatial ability. It has been demonstrated that a pilot's spatial ability has some predictive value for their flight ability.[4] Superb pilots are more capable in perceptual quick conversion, as well as verbal and graphic memory, and they have superior mental rotation, efficient thinking, and excellent short-term working memory.

In the present study, three groups of pilots with different spatial abilities were required to judge, as quickly and accurately as possible, whether the two stimulus models in the virtual scenario coincided. The results showed significant differences in RT, CR, and CNPS among the high, middle, and low spatial ability groups. The results revealed that the pilots in the high spatial ability group had remarkably slower RT and higher CNPS compared to the pilots in the low spatial ability group. The significant gap between groups illustrated the correlation between pilots' spatial ability and mental rotation ability. This was consistent with the results of previous studies. In a recent study, Roach et al.[15] used a timed mental rotation test to explore differences in the performance of individuals in different spatial ability groups when solving mental rotation tasks, and they found that individuals with high spatial ability showed a strong advantage over individuals with low spatial ability in terms of the time required to complete the test and the accuracy of their answers. Many studies have identified that when the speed of visuospatial tasks is accelerated, the burden on working memory increases and the speed of mental processing becomes a key factor in performance.[26] Thus, the number of incorrect answers may not be directly attributable to a deficiency in ability or the lack of knowledge tested, but instead to their ability to process information with great speed. Too much emphasis on answer speed may mask the ability of individuals with low spatial ability to perform on tests requiring spatial reasoning, and the RT may directly affect concentration during problem-solving and reduce the ability to accurately perform MRTs.

Researchers have demonstrated that adults' cognitive abilities decline gradually with age after reaching optimal levels. Salthouse et al.[18] compared the spatial visualization ability of male subjects in two age groups and found that there was a decreasing trend in spatial visualization test scores with increasing age. Using seven paper-and-pencil tests reflecting basic cognitive ability, Wu et al.[27] explored the differences in the cognitive test scores and the flight ability of pilots in different age groups, and they discovered that the optimal ages of basic cognitive abilities for pilots were between ≥26 and <29 yr of age, with a decline trend after 29 yr. In another study, they also revealed that the short-term memory ability of pilots decreased significantly after 35 yr of age, and the acceleration algorithm and dual operation ability of pilots decreased specifically after 29 yr.[28] This has similarity with the outcomes of our study. There were significant differences in RT and CNPS on the MRT among pilots in various age groups in our study. The pilots in the age group of 29–35 yr had the fastest RT and the highest CNPS in the mental rotation test. In addition, adult cognitive ability generally tends to decline after reaching optimal levels; however, on the MRT in this study, pilots over 35 yr had the higher CNPS, the higher CR, and the slower RT than the younger pilots (<28 yr). There are a few possible reasons. Firstly, mental rotation ability reflects the basic ability to encode and transform spatial graphics. More experienced pilots, who have accepted a greater amount of highly intelligent and informative training, manifest superior spatial cognitive ability and are able to construct mental images of objects in their minds more speedily after rotation. Secondly, it is closely related to the actual working practice and operating environment of pilots. As young pilots have shorter flight time, insufficient flight working experience, and smaller spatial working memory capacity, they fail to adapt to the need to search, process, integrate, and decide on various information decisively and correctly in a shorter period of time, and therefore they have longer reaction latency and reaction time than senior pilots in mental rotation tasks. As a matter of fact, our findings showed a significant difference in the flight time of pilots in different age groups [388.0 ± 125.4, 1682.0 ± 455.2, 2564.8 ± 636.9, $F(2115)=373.1$, $P = 0.000$]. The flight time of the older group was significantly longer than the other two groups. This may be a verification of our assumption. However, we did not investigate the effect of subjects' gaming or VR experience on the experiential criteria of this test, and future studies are needed to explore the impact of prior gaming or VR experience on testing.

A large amount of literature also reports a relationship between space performance and gender.[24] Earlier studies indicated that females lagged in performance in tasks with spatial factors, and the variance between males and females increased with age. However, it has also been indicated that gender difference may be limited and even absent, depending on the stimuli and tasks.[22] Some studies have also pointed out that gender has less effects on cognitive spatial processing speed in the pilots with aviation experience.[13] Gender effects may be associated with lower spatial cognitive abilities, whereas in pilots who are considered to have higher spatial cognitive abilities, there may be an experience factor that leads to no gender differences.[23] In our study, the male and female pilots in our study did not exhibit any noticeable differences in RT, CR, or CNPS in the MRT. Schnable et al.[19] pointed out that the ability to get immediate feedback by interacting with objects in a 3D virtual environment contributed to a better understanding of spatial issues, whereas display effects and technical drawings presented in 2D form generally made it difficult for subjects to intuitively understand space. The VR environment used in our study can realistically simulate the structure of the stimulus after rotation, which can reduce the cognitive load, thus helping to observe and evaluate spatial information more intuitively, improving the processing of spatial perceptual information and compensating for the stress of weaker spatial abilities in females. Besides, our findings showed insignificant differences in RT, CR, and CNPS in consideration of gender in different ages and flight times ($P = 0.454$, $P = 0.207$ and $P = 0.468$). However, a limitation of this study was that there were less female pilot subjects, and possibly different consequences may emerge with subsequent expansion of the sample size.

With the development of VR, it is believed that the experimental research of spatial visualization ability will focus more on the aspect of ecological scenarios, the spatial ability test based on VR will be combined with comprehensive dynamic flight actual tasks, and the test scene will be closer to real flight scenarios, providing a new technical tool for the prediction and training of pilots' spatial cognitive ability.

ACKNOWLEDGMENTS

The views expressed in the article are those of the authors, and we would like to thank the participants for their effort and commitment. The authors also would like to thank Dr. Entong Wang for reviewing and revising the article.

Financial Disclosure Statement: This work was supported by an Army-Wide Logistic Key Project (BKJ19J020) and an Air Force Medical University Science and Technology Research Project (2019ZTC02). The authors have no competing interests to declare.

Authors and Affiliations: Mengdi Zhang, M.D., and Zhanguo Jin, M.D., Ph.D., Air Force Clinical College, Anhui Medical University, the Fifth School of Clinical Medicine, Anhui Medical University, Hefei, Anhui 230032, China, and Aerospace Balance Medical Center, Air Force Medical Center, PLA, Beijing, 100142; Meng Wang, M.D., Xunyuan Liu, M.D., Ph.D., Lihong Zhai, M.D., and Xianrong Xu, M.D., Ph.D., Aerospace Balance Medical Center, Air Force Medical Center, PLA, Beijing, 100142, China; and Huimin Feng, M.D; Emergency Department, Beijing Sijiqing Hospital, Beijing, 100142, China.

REFERENCES

1. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, et al. The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. Anat Sci Educ. 2020; 13(5):558–567.
2. Commins S, Duffin J, Chaves K, Leahy D, Corcoran K, et al. NavWell: A simplified virtual-reality platform for spatial navigation and memory experiments. Behav Res Methods. 2020; 52(3):1189–1207.
3. Foti F, Sorrentino P, Menghini D, Montuori S, Pesoli M, et al. Peripersonal visuospatial abilities in Williams syndrome analyzed by a table radial arm maze task. Front Hum Neurosci. 2020; 14:254.
4. Giancola M, Verde P, Cacciapuoti L, Angelino G, Piccardi L, et al. Do advanced spatial strategies depend on the number of flight hours? The case of military pilots. Brain Sci. 2021; 11(7):851–865.
5. Glasauer S, Mittelstaedt H. Perception of spatial orientation in microgravity. Brain Res Brain Res Rev. 1998; 28(1–2):185–193.
6. Ho S, Liu P, Palombo DJ, Handy TC, Krebs C. The role of spatial ability in mixed reality learning with the HoloLens. Anat Sci Educ. 2022; 15:1074–1085.
7. Hodgkiss A, Gilligan KA, Tolmie AK, Thomas MSC, Farran EK. Spatial cognition and science achievement: the contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. Br J Educ Psychol. 2018; 88(4):675–697.
8. Kelley TL. The selection of upper and lower groups for the validation of test items. J Educ Psychol. 1939; 30(1):17–24.
9. Kurul R, Ögün MN, Neriman Narin A, Avci Ş, Yazgan B. An alternative method for anatomy training: immersive virtual reality. Anat Sci Educ. 2020; 13(5):648–656.
10. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities training in anatomy education: a systematic review. Anat Sci Educ. 2020; 13(1):71–79.

11. Liu S, Wei W, Chen Y, Hugo P, Zhao J. Visual-spatial ability predicts academic achievement through arithmetic and reading abilities. Front Psychol. 2021; 11:591308.
12. Miao D, Hu W, You X, Zhang Q, Bai Y. The multiple criteria evaluative system for flight ability and validity analysis of Sternberg dual task test. [Article in Chinese]. Chinese Journal of Aerospace Medicine. 1997; 8(3):149–153.
13. Munion AK, Stefanucci JK, Rovira E, Squire P, Hendricks M. Gender differences in spatial navigation: characterizing wayfinding behaviors. Psychon Bull Rev. 2019; 26(6):1933–1940.
14. Pak R, Rogers WA, Fisk AD. Spatial ability subfactors and their influences on a computer-based information search task. Hum Factors. 2006; 48(1):154–165.
15. Roach VA, Fraser GM, Kryklywy JH, Mitchell DGV, Wilson TD. Different perspectives: spatial ability influences where individuals look on a timed spatial test. Anat Sci Educ. 2017; 10(3):224–234.
16. Roach VA, Mi M, Mussell J, Van Nuland SE, Lufler RS, et al. correlating spatial ability with anatomy assessment performance: a meta-analysis. Anat Sci Educ. 2021; 14(3):317–329.
17. Rodes W, Gugerty L. Effects of electronic map displays and individual differences in ability on navigation performance. Hum Factors. 2012; 54(4):589–599.
18. Salthouse TA, Babcock RL, Skovronek E. Age and experience effects in spatial visualization. Dev Psychol. 1990; 26(1):128–136.
19. Schnabel MA, Kvan T. Spatial understanding in immersive virtual environments. Int J Archit Comput. 2003; 1(4):435–448.
20. Sezer B, Sezer TA, Elcin M. Exploring spatial ability in healthcare students and the relationship to training with virtual and actual objects. Eur J Dent Educ. 2022; 26(2):310–316.
21. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. Science. 1971; 171(3972):701–703.
22. Toth AJ, Campbell MJ. Investigating sex differences, cognitive effort, strategy, and performance on a computerised version of the mental rotations test via eye tracking. Sci Rep. 2019; 9(1):19430.
23. Verde P, Piccardi L, Bianchini F, Trivelloni P, Guariglia C, Tomao E. Gender effects on mental rotation in pilots vs. nonpilots. Aviat Space Environ Med. 2013; 84(7):726–729.
24. Voyer D, Rodgers MA, McCormick PA. Timing conditions and the magnitude of gender differences on the Mental Rotations Test. Mem Cognit. 2004; 32(1):72–82.
25. Wanzel KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet. 2002; 359(9302):230–231.
26. Wilhelm O, Schulze R. The relation of speeded and unspeeded reasoning with mental speed. Intelligence. 2002; 30(6):537–554.
27. Wu G, Bi H, Aqing Y, Deng X, Tian G, et al. Influence of aging on short-term memory, high-speed calculative skill and dual task performances in fighter pilots. [Article in Chinese]. Chinese Journal of Aerospace Medicine. 1998; 9(4):229–231.
28. Wu G, Li Z, Tian G, Ma X, Deng X, et al. Age difference in basic cognitive ability of fighter pilots and its influence upon flying. [Article in Chinese]. Chinese Journal of Aerospace Medicine. 1998; 9(3):133–135.