ACTA DIDACTICA NAPOCENSIA
Volume 18, Number 2, 2025 – DOI: 10.24193/adn.18.2.7

Developing Spatial Reasoning Test for Pre-Service Primary Teachers

Khin Mimi KYAW, Tibor VIDÁKOVICH

Abstract: Spatial reasoning is a crucial cognitive skill and plays a significant role in STEM fields. Teachers' spatial reasoning abilities are particularly important for it influences students' learning. The purpose of this study was to develop a test to examine the spatial reasoning skills of pre-service primary teachers. The test was aligned with primary mathematics curriculum that enforces problem-solving skills in students. Spatial Operational Capacity Model was used as a framework in designing the test. This study employed quantitative descriptive design. A ten-item multiple-choice test was administered to 292 pre-service primary teachers. Item analysis was conducted using Microsoft Excel 2016 and SPSS 25, and the results indicated that most items were acceptable.Test items' difficulty level ranged from 0.19 to 0.82 and only two items were found to have discrimination value less than 0.30. The mean score and the standard deviation were 6.03 and 1.8 respectively. Pre-service teachers found the cubes comparison task the most difficult, showing their limited spatial experience in manipulating three-dimensional objects. This process of constructing test items with optimal difficulty and discrimination is essential for examining teachers' spatial reasoning skills in future research.

Key words: mathematics, spatial reasoning, test development, test analysis

1. Introduction

Spatial reasoning refers to a collection of cognitive processes involved in perceiving, representing, and mentally manipulating objects and relationships in space. Linn and Petersen (1985) describe spatial reasoning as the ability to represent, transform, generate, and retrieve nonverbal and symbolic information. The concept was identified through intelligence testing. A wide range of closely related terms has been used in the literature to describe individuals' capacity to process spatial information, such as "spatial ability", "spatial reasoning", "spatial thinking" and "spatial visualization". In recent years, growing attention has been given to the development of spatial reasoning skills in children, largely due to the increasing demand for proficiency in science, technology, engineering, and mathematics (STEM) fields (Davis, 2015). Numerous studies have demonstrated a strong relationship between spatial ability and achievement in STEM disciplines, indicating that students with well-developed spatial skills are more likely to engage in and succeed within these areas (Wai, Lubinski, & Benbow, 2009).

Research has also established consistent links between spatial reasoning and mathematical problem-solving, as many complex mathematical tasks rely heavily on spatial thinking (Gunderson, Ramirez, Beilock, & Levine, 2012). Many countries including US, Canada and Australia emphasize spatial reasoning in their curriculum. It is considered foundational and educators encourage schools to foster spatial reasoning skills in children since early childhood (Clements & Battista, 1992; Jones & Tzekaki, 2016). Teachers should also possess strong spatial skills since they are crucial to achieving this.

This study focused on future teachers and developed a spatial reasoning test for pre-service primary teachers. Although there are several spatial tests, navigating and choosing spatial tasks for appropriate use in school education is difficult due to its complexity and diversity. While spatial skills are not typically taught in the school curriculum, spatial instruments are often hard to access and use.

Received February 2025

Cite as: Kyaw, K. M., & Vidákovich, T. (2025). Developing Spatial Reasoning Test for Pre-service Primary Teachers, *Acta Didactica Napocensia,* 18(2), 73-83, https://doi.org/10.24193/adn.18.2.7

Measures of spatial reasoning may include diverse activities such as judging horizontality, performing mental rotations, or identifying simple shapes embedded within more complex figures (Linn & Petersen, 1985). Most of these instruments originated as psychometric tools developed to assess occupational suitability, particularly in military settings, before later being modified for use with high school and university populations (Hegaryt & Waller, 2005). Consequently, many of them focus on measuring discrete spatial abilities identified through factor-analytic research. These tools were largely designed by psychologists for controlled experimental environments, rather than for direct application in everyday classroom settings (Atit et al., 2020; Lowrie et al., 2020). Thus, it is necessary to develop a spatial reasoning test that is appropriate for use in the classroom

In 2014, Myanmar undertook a major reform of its primary education curriculum with support from the Japan International Cooperation Agency (JICA). Through the initiative titled *Project for Curriculum Reform at Primary Level of Basic Education*, new textbooks and teacher guides were developed for ten subjects, including Myanmar, English, Mathematics, Science, and Social studies. One of the key goals of this reform was to move away from traditional memorization-based instruction toward teaching strategies that promote critical thinking, creativity, and effective communication. This educational transformation places greater emphasis on reasoning skills. Key reasoning skill such as logical reasoning involves complex and abstract cognitive processes and has been shown to be more strongly related to spatial ability than to numerical ability (Xie et al., 2020). Many spatial reasoning tests utilize multiple-choice items to assess abilities such as shape manipulation and visualization. Downing (2002) demonstrated that multiple-choice items are capable of assessing all levels of cognitive functioning, including higher-order thinking skills.

2. Theoretical framework

2. 1. Spatial Operational Capacity (SOC) model

The Spatial Operational Capacity (SOC) model suggests that spatial reasoning skills can be strengthened through the use of forms of representation to act on various physical and mental objects and transformations. This model emphasizes the importance of engaging learners in diverse spatial tasks that involve both concrete (physical) and abstract (mental) manipulations, enabling them to develop the necessary skills to solve complex spatial problems. The Spatial Operational Capacity (SOC) model explains the complexity of visual representations by categorizing four key variables: perception, dimensionality, transformation, and mobility. These components contribute to how individuals process spatial information and solve spatial reasoning tasks (Sack and Vazquez, 2016).

- Perception – The stimulus with which visual information is presented, can include full-scale 3-D figures; conventional graphic 2D pictures that resemble the 3-D figures; semiotic representations such as front, top and side views that do not obviously resemble the 3-D figures; and verbal descriptions. The form of representation influence how individuals recognize and interpret spatial relationships.
- Dimensionality – The objects, which are presented via the visual information that the individual perceives, processes or acts on, may be one-dimensional, such as points and lines; two-dimensional, including shapes like triangles and quadrilaterals; or three-dimensional, such as prisms and pyramids. The objects may constitute either part of the stimulus or the entire visual representation.
- Transformation – A key cognitive process required during spatial reasoning, namely the ability to recognize and mentally process changes applied to objects. These changes can include positional (rotations, translations) or structural (scaling, deformation) changes, or a combination of both.
- Mobility – The nature of the visual image per se, that can be static or dynamic in nature. Visual stimuli may exist on a continuum ranging from fixed, printed images to dynamic, digitally animated representations that change over time (Sack and Vazquez, 2016).

In line with this framework, test items that require the manipulation of two- and three-dimensional objects are particularly effective indicators of spatial reasoning. Furthermore, research has shown that

proficiency in tasks involving moving or transforming objects—such as folding, bending, rotating, scaling, cross-sectioning, or comparing two-dimensional and three-dimensional views—is positively associated with mathematical achievement (Newcombe et al., 2019; Uttal et al., 2013). These tasks demand mental manipulation of spatial stimuli and are therefore strong measures of spatial ability.

2. 2. Item Difficulty

Item difficulty and item discrimination are essential metrics in test construction, particularly in educational and psychological fields. These two indices are frequently used to select items for the test development. Item difficulty indicates the overall proportion of examinees that accurately answered an item. The proportion is usually denoted as p and called the difficulty index (Crocker & Algina, 1986). The higher the difficulty index, the easier the item is understood to be (Quaigrain & Arhin, 2017). Test items with moderate levels of difficulty are generally preferred over those that are excessively easy or overly challenging. Most studies use a method known as the 27% rule to divide the participants into top 27% and bottom 27% to determine item difficulty (Boopathiraj & Chellamani, 2013). This approach is commonly used in classical test theory (CTT) by comparing the performance of high and low scorers (Kelley, 1939). Item difficulty is calculated using the formula:

DL= RU+RL/NU+NL, where RU and RL represent the number of correct responses in the upper and lower groups, respectively, and NU and NL denote the number of students in each group.

Table 1. *Classification of difficulty level*

Value of difficulty index (p)	Classification
$P < 0.20$	Very difficult
$0.20 \leq P \leq 0.40$	difficult
$0.40 \leq p \leq 0.60$	Moderately difficult
$0.60 \leq p \leq 0.80$	Easy
$P > 0.80$	Very easy

Source: Hotiu (2006)

2. 3. Item Discrimination

Item discrimination refers to an item's effectiveness in distinguishing between high- and low-performing examinees. The discrimination index (D) indicates how well an item differentiates stronger students from weaker ones, with higher values suggesting that high scorers are more likely to answer correctly while low scorers are not (Angoff, 1971, pp. 1–13). High discrimination is essential for enhancing test reliability. To compute the discrimination index, examinees are ranked according to their test scores, and the upper and lower 27% groups are compared. Zubairi and Kassim (2006) note that this proportion maximizes group differences in a normal distribution while retaining an adequate sample size for analysis. Discrimination power (DP) is calculated using the formula: DP = (RU − RL) / NU or NL.

Table 2. *Classification of Discriminating Power (DP)*

Discriminating Power (DP)	Classification
$D \leq 0.19$	Poor item (to be excluded/changed)
$0.20 \leq D \leq 0.29$	Marginal item (to be revised)
$0.30 \leq D \leq 0.39$	Good item
$D \geq 0.40$	Satisfactory item

Source: Ebel (1979)

3. Rationale of the study and research questions

Currently, there are no spatial ability tests specifically designed to assess Myanmar teachers, and the importance of spatial reasoning is not yet widely recognized. Moreover, the recent changes in the primary mathematics curriculum have introduced solid geometry as a new topic, highlighting the need

for teachers to possess strong spatial reasoning skills to effectively teach these concepts. Therefore, a spatial reasoning test tailored to mathematics education was developed for pre-service teachers in Myanmar. The primary aim of this study was to construct a teacher-made spatial reasoning assessment and to evaluate the difficulty levels and discrimination indices of its items. The study was guided by the following research questions:

RQ1.How well did pre-service primary teachers perform in the spatial reasoning test items ?

RQ2.Is there a gender difference in the performance of spatial reasoning test?

RQ3. What are the difficulty and discrimination indices of the spatial reasoning test items?

4. Methodology

4. 1. Participants

The study involved graduate students aged 25-35 who were enrolled in the Certificate in Primary Teacher Education program at Yankin Education Degree College. Purposive sampling method was adopted. The sample constituted both male and female student teachers as shown in Table 3.

Table 3. *Distrbution of Students by gender*

Gender	Number	Percentage (%)
Male	62	22
Female	230	78
Total	292	100

4. 2. Instrument

Data for this study were collected using a spatial reasoning assessment developed by the researchers. Test items were created based on an extensive review of online aptitude assessments and psychometric test preparation resources to ensure a wide variety of spatial reasoning tasks. The construction of the instrument was guided by the Spatial Operational Capacity (SOC) model as well as Myanmar's revised primary mathematics curriculum. Although the updated curriculum and textbooks include numerous new topics, curriculum developers suggest that teachers require additional practice and experience to effectively engage with these materials (Itoh et al., 2022). The revised primary mathematics textbooks emphasize problem-solving in solid geometry, incorporating tasks related to two- and three-dimensional shapes, part–whole relationships, transformations, and symmetry. In response, the spatial reasoning test was designed to capture a comprehensive range of spatial abilities involving both 2D and 3D manipulation (see Figure 1). The assessment included items focused on mirror reflection, rotation of two- and three-dimensional figures, matching and combining planar shapes, cube comparison, mental folding of cubes, and perspective-taking. The test consisted of ten multiple-choice items, each offering four response options, and was administered in the participants' native language (Burmese). Instrument validity was established through expert evaluation (Cohen et al., 2007). To ensure content and face validity, three teacher educators from the departments of Mathematics Education and Educational Psychology reviewed the test. Based on their feedback, minor revisions were made to item wording, while all original items were retained.

4. 3. Data Collection and Analysis

The spatial reasoning test was administered to students in a classroom setting. As an effective means of incorporating technology into learning, teaching and assessment, Google Forms, a powerful online technology tool was utilized in assessing students' spatial skills. Students were provided an overview of the test questions before the test began. Students completed the test on their devices within a 20-

minute time limit and submitted their responses online. The data analysis was conducted using Microsoft Excel 2016 and SPSS 25.

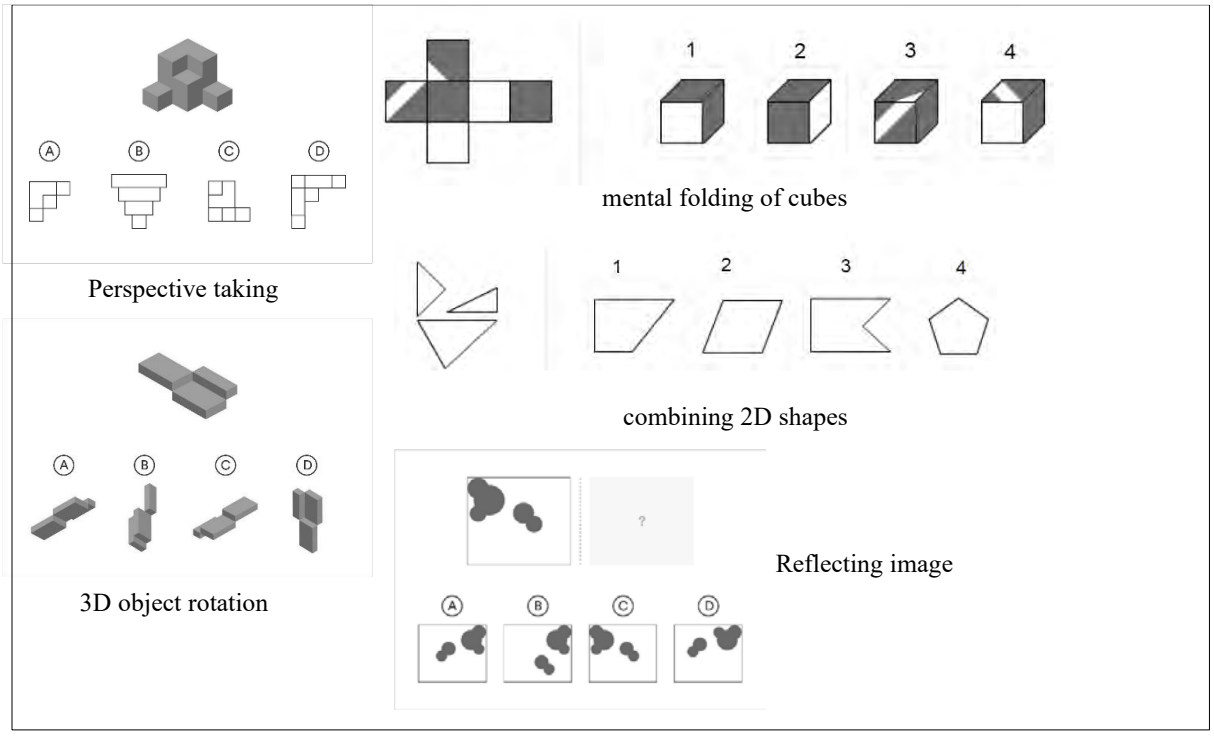

Figure 1. *Sample spatial reasoning test items (Source: Practice Aptitude Tests)*

5. Results

The assessment comprised ten items, resulting in a maximum possible score of 10, with all items weighted equally. Responses were scored dichotomously, assigning one point for each correct answer and zero points for incorrect responses. A total of 292 completed responses were obtained. To evaluate the internal consistency of the instrument, a reliability analysis was performed, yielding a Cronbach's alpha coefficient of 0.62. This value indicates an acceptable level of reliability for a teacher-developed assessment.

Student performance was evaluated by summing individual item scores. The distribution of total test scores is illustrated in Figure 2. Most participants achieved scores ranging from 4 to 8, with the highest frequencies occurring at scores of 5 (58 students), 6 (56 students), and 7 (53 students). These results suggest that the majority of students demonstrated average to moderately above-average performance, reflecting a moderate level of spatial reasoning ability within the sample. Additionally, extreme low scores (e.g., total score of 1) are rare, implying that most students could cope with spatial reasoning tasks to some extent.

To examine potential gender differences in performance, normal Q–Q plots and boxplots were reviewed, showing that score distributions for both male and female participants approximated normality. For male students, skewness was −0.347 (SE = 0.304) and kurtosis was −0.099 (SE = 0.599), while female students exhibited a skewness of −0.125 (SE = 0.160) and a kurtosis of −0.441 (SE = 0.320), meeting accepted criteria for normal distribution assumptions (Crammer & Howitt, 2004, pp. 157–159).

An independent-samples t-test was then used to compare spatial reasoning scores by gender. The analysis revealed no statistically significant difference between male students (M = 6.00, SD = 1.71) and female students (M = 6.02, SD = 1.82), $t(290) = -0.085$, $p > .05$.

Further tabulation was made to compare students' performance across items, "see Table 4". It was observed that students' strongest performance area in the test was 2D shapes matching. Students

achieved the highest success rate of 83.6%, indicating their strong ability to recognize and compare identical shapes.

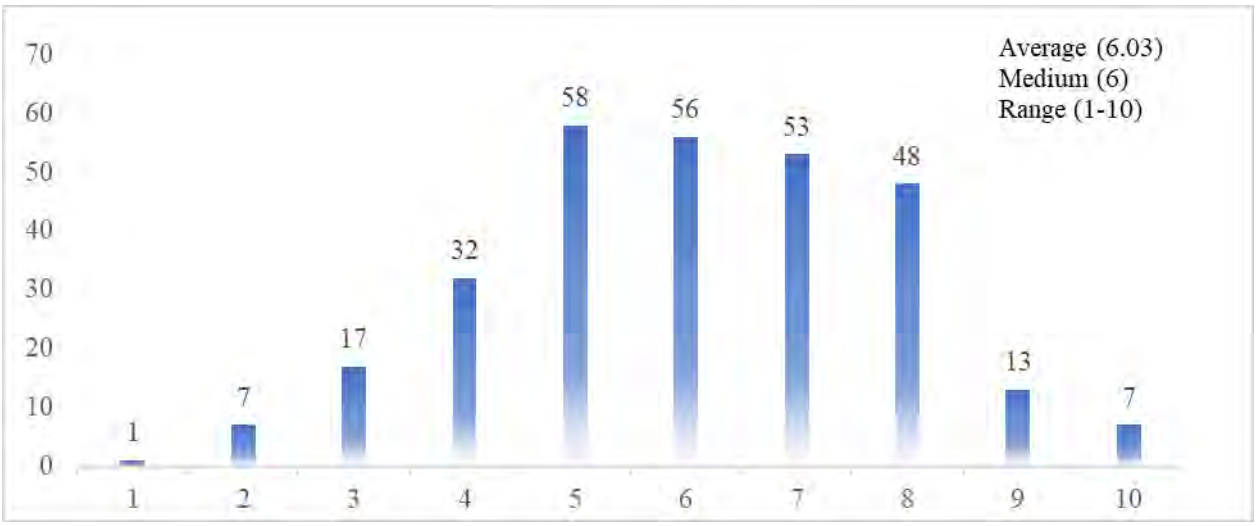

Figure 2. *Average, Median, and Range*

Additionally, high success rates were recorded in 3D rotation and group rotation of 2D shapes, suggesting their ability to manipulate and understand rotational transformation in both two- and three-dimensional shapes.

Students' moderate performance was found in the mirror reflection task (45.9%) where students encountered difficulties in visualizing mirrored versions of objects. They struggled to capture detailed features accurately and had challenges in assessing the relative positions of objects after reflection. In addition, the perspective-taking task (64.0%), which required students to visualize objects from different viewpoints, demonstrated a reasonable success rate. However, some students faced challenges in accurately translating 3D objects into 2D representations.

Table 4. *Test Items and the Number of Correct Responses*

No.	Test Items	Correct Responses	Success Rate
1.	Which of the given shapes is the correct mirror image? (reflecting)	134/292	45.9%
2.	What would the 3D object look like from above? (perspective-taking)	187/292	64.0%
3.	Which figure is a rotation of the object? (rotating 3D object)	222/292	76.0%
4.	Which of the answer figures is a rotation of the question figure? (group rotation of 2D shape)	220/292	75.3%
5.	Which of the answer figures is a rotation of the question figure? (group rotation of 2D shape)	215/292	73.6%
6.	Which of the complete shapes can be made from the components shown? (combining 2D shapes)	125/292	42.8%
7.	Which cube can be made based on the unfolded cube? (mental folding of cubes)	84/292	28.8%
8.	Which two pictures are identical? (2D shapes matching)	244/292	83.6%
9.	Which of the four possible options represent the cube shown from a different perspective? (3D cubes comparison)	45/292	15.4%
10.	Which of the complete shapes can be made from the components shown? (combining 2D shapes)	226/292	77.4%

Poor performance of the students was found in tasks requiring complex three-dimensional visualization and mental manipulation. The 3D cubes comparison task (15.4%) had the lowest success rate, highlighting significant difficulty in visualizing how 3D objects appear from different

perspectives. Similarly, the mental folding of cubes task (28.8%) proved challenging, indicating that students struggled with mentally converting 2D nets into 3D structures. Both tasks require the execution of multi-step transformations and advanced spatial skills, areas in which students demonstrated weaknesses. Additionally, combining 2D shapes task (42.8%) exhibited variability in performance, indicating inconsistencies in students' ability to mentally assemble components into complete figures. These areas highlight the need for targeted interventions and practice to strengthen students' spatial reasoning abilities.

The current test showed varying difficulty indices from 0.19 to 0.82, "see Table 5". The distribution of items based on difficulty level classification, as presented in Table 1, is illustrated in a pie chart. It was found that 50% of the items were in easy category, "see Figure 3". The item discrimination indices for each item are presented in Table 6, with their classifications based on the discrimination index provided in Table 2.

Table 5. *Difficulty Indices of spatial reasoning test items*

Item No.	Difficulty index (p)	Item No.	Difficulty index (p)
1.	0.53	6.	0.44
2.	0.61	7.	0.32
3.	0.72	8.	0.82
4.	0.71	9.	0.19
5.	0.72	10.	0.79

According to item discrimination index, 70% of items in this spatial reasoning test fall in the satisfactory category. Two items were marginal that need to be revised, with a discriminative index of 0.25. However, the current test showed no poor item nor item with negative or zero discrimination.

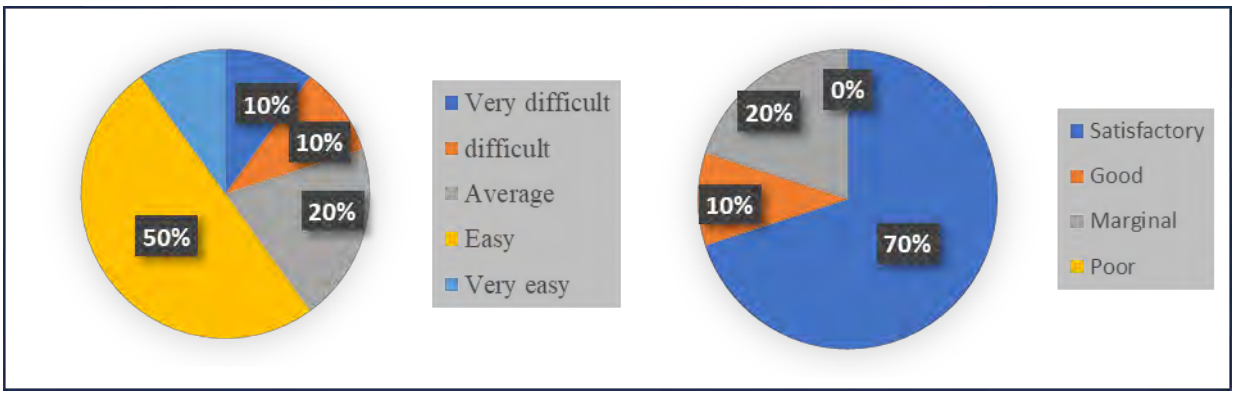

Figure 3. *Distribution of Items based on Difficulty and Discrimination Indices*

Table 6. *Discrimination Indices of spatial reasoning test items*

Item No.	Discriminating Power (DP)	Item No.	Discriminating Power (DP)
1.	0.52	6.	0.48
2.	0.61	7.	0.54
3.	0.46	8.	0.25
4.	0.46	9.	0.25
5.	0.44	10.	0.34

6. Discussion

The discussion was made based on the research question framed for this study. This study aimed to assess the students' performance in a teacher-made spatial reasoning test and determine the item difficulty level and discrimination power of the individual test items. 292 responses were collected from pre-service primary teachers using Google Forms. The mean score of the pre-service primary teachers is 6.03, and range is 1-10. About 40% of the students were high performing students, i.e., their scores were within 7-10 points. The easiest task was 2D shapes matching question (item 8) with 83.6 % of correct answers, while the most difficult was 3D cubes comparison (item 9) with only 15.4% of correct answers. Cubes comparison test measures mental rotation which is the core factor of spatial skills and students need exercises to understand the orientation of objects with respect to viewing angles.

No gender difference was found and this result coincides with some previous studies. While gender differences in cognitive abilities are generally small, research indicates that males tend to systematically outperform females in tests measuring spatial ability or mathematical ability. This trend has been observed in various studies examining different aspects of spatial cognition. However, it is important to note that findings on gender differences in spatial ability are not always consistent. For example, Turğut and Yılmaz (2012) reported no significant gender differences in spatial orientation and visualization among pre-service teachers.

Furthermore, studies by Tasova (2012), Yılmaz (2012), and Sevimli (2011) indicate that pre-service teachers generally exhibit low levels of spatial visualization and representation transition abilities. These findings highlight a concerning gap in the preparedness of future educators, as spatial skills are crucial for teaching subjects such as mathematics and science, where spatial reasoning plays a vital role in understanding geometric concepts, interpreting diagrams, and solving real-world problems. The inability to perform well in spatial tasks may hinder their effectiveness in the classroom and limit their ability to foster students' spatial reasoning skills. In contrast to these findings, the current study observed relatively above-average success in spatial test items among participants. Nevertheless, these results should be interpreted with caution due to differences in research design and assessment instruments, as well as the limited number of test items used in the current study. Many earlier investigations employed more demanding standardized measures, such as the Purdue Spatial Visualization Test (PSVT), which are psychometrically complex and typically administered under strict time limitations.

An effective assessment tool should comprise items with a range of difficulty levels, typically beginning with more challenging questions, followed by items of moderate difficulty (approximately 40%–60%), and concluding with easier items (Boopathiraj & Chellamani, 2013). The spatial reasoning test developed in this study was designed in accordance with this recommended structure. In general, items with difficulty indices between 20% and 80% are considered suitable for inclusion in an assessment. In the present test, item difficulty values ranged from 0.19 to 0.82, indicating that most items fell within an acceptable range.

Regarding item discrimination, questions with discrimination indices of 0.40 or higher are regarded as highly effective in distinguishing between high- and low-performing examinees. In the current assessment, seven items demonstrated strong discriminatory power, while two items were identified as marginal and therefore require revision. Reliability analysis yielded a Cronbach's alpha coefficient of 0.62, suggesting a moderate level of internal consistency. Reliability coefficients between 0.50 and 0.60 are generally considered adequate for exploratory, teacher-developed instruments, although they may be insufficient for high-stakes testing contexts (Quaigrain & Arhin, 2017).

Future refinements of the spatial reasoning test will focus on improving the marginal items to enhance overall reliability. Based on the findings of this study, it can be concluded that the test demonstrates acceptable levels of validity and reliability for its intended purpose.Overall, while the current findings are encouraging, they should be interpreted with caution. Further studies using standardized, time-controlled, and psychometrically validated spatial tests are needed to accurately assess teacher candidates' spatial abilities. Additionally, longitudinal studies examining the impact of spatial training throughout teacher education programs could provide valuable insights into how these skills develop

over time and how they can be effectively nurtured to enhance teaching competence in spatially intensive subjects.

7. Conclusion

In conclusion, examining spatial reasoning abilities among pre-service teachers is vital to ensuring that they are well prepared to teach geometry and other spatially intensive topics, thereby supporting effective mathematics instruction and student learning. Findings from the study will help shape more effective teacher training programs, ultimately contributing to the development of a more spatially literate generation of students capable of tackling real-world challenges. This study also provides well-designed test items with an appropriate difficulty level and strong discrimination power, making them suitable for use with both students and teachers in classroom settings. The multiple-choice items used in this study can serve as a valuable addition to an item bank for future assessment of spatial reasoning skills.

References

Angoff, W. H. (1971). The admissions testing program. In The College Board admissions testing program: A technical report on research and development activities relating to the Scholastic Aptitude Test and Achievement Tests (pp. 1–13). College Entrance Examination Board.

Atit, K., Uttal, D. H., & Stieff, M. (2020). Situating space: Using a discipline-focused lens to examine spatial thinking skills. Cognitive Research: Principles and Implications, 5, Article 19. https://doi.org/10.1186/s41235-020-00210-z

Boopathiraj, C., & Chellamani, K. (2013). Analysis of test items on difficulty level and discrimination index in the test for research in education. International Journal of Social Science & Interdisciplinary Research, 2(2), 189–193.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics*, 420-464.

Cohen, L., Manion, L., & Morrison, K. (2007). Research methods in education (6th ed.). Routledge.

Crammer, D., & Howitt, D. (2004). The SAGE dictionary of statistics: A practical resource for students in the social sciences. SAGE Publications.

Crocker, L., & Algina, J. (1986). Item analysis. In Introduction to classical and modern test theory (pp. 311–329). Holt, Rinehart and Winston.

Davis, B. (2015). *Spatial reasoning in the early years: Principles, assertions, and speculations*. Routledge. https://doi.org/10.4324/9781315762371

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), Handbook of test development (pp. 3–25). Lawrence Erlbaum Associates.

Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. Developmental Psychology, 48(5), 1229–1241. https://doi.org/10.1037/a0027433

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), The Cambridge handbook of visuospatial thinking (pp. 121–169). Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.005

Hotiu, A. (2006). *The relationship between item difficulty and discrimination indices in multiple-choice tests in a Physical science course* (MSc thesis). Florida Atlantic University, Boca Raton, FL. Retrieved from www.physics.Au/eAdu/research/education/A.Hotiu_thesis.pdf

Itoh, T., Imahori, I., & Takahashi, K. (2022). The development of mathematics textbooks in Myanmar [Conference presentation]. ICDME–Tsukuba Conference, Tokyo, Japan.

Jones, K., & Tzekaki, M. (2016). Research on the teaching and learning of geometry. *The second handbook of research on the psychology of mathematics education: The journey continues*, 109-149.

Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 30(1), 17–24. https://doi.org/10.1037/h0057123

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development, 56(6), 1479–1498. https://doi.org/10.2307/1130467

Lowrie, T., Resnick, I., Harris, D., & Logan, T. (2020). In search of the mechanisms that enable transfer from spatial reasoning to mathematics understanding. Mathematics Education Research Journal, 32, 175–188. https://doi.org/10.1007/s13394-020-00336-9

Newcombe, N. S., Booth, J. L., & Gunderson, E. A. (2019). Spatial skills, reasoning, and mathematics. *The Cambridge handbook of cognition and education*, 100-123, DOI: 10.1017/9781108235631

Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. Cogent Education, 4(1), 1301013. https://doi.org/10.1080/2331186X.2017.1301013

Sack, J., & Vazquez, I. (2016). A 3D visualization teaching-learning trajectory for elementary grades children. Springer. https://doi.org/10.1007/978-3-319-29799-6

Sevimli, E., & Delice, A. (2011). The influence of teacher candidates' spatial visualization ability on the use of multiple representations in problem solving of definite integrals: A qualitative analysis. Research in Mathematics Education, 13(1), 93–94. https://doi.org/10.1080/14794802.2011.550750

Tasova, H. I., & Delice, A. (2012). An analysis of pre-service mathematics teachers' performance in modelling tasks in terms of spatial visualisation ability. Research in Mathematics Education, 14(3), 297–298. https://doi.org/10.1080/14794802.2012.734994

Turğut, M., & Yılmaz, S. (2012). Relationships among preservice primary mathematics teachers' gender, academic success and spatial ability. International Journal of Instruction, 5(2), 5–20.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*(2), 352–402, doi.org/10.1037/a0028446

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817-835. https://doi.org/10.1037/a0016127

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is spatial ability related to mathematical ability? A meta-analysis. Educational Psychology Review, 32, 113–155. https://doi.org/10.1007/s10648-019-09496-y

Zubairi, A. M., & Kassim, N. L. A. (2006). Classical and Rasch analysis of dichotomously scored reading comprehension test items. Malaysian Journal of ELT Research, 2, 1–20.

Authors

Khin Mimi KYAW, University of Szeged, Hungary. E-mail: khin.mi.mi.kyaw@edu.u-szeged.hu

Tibor VIDÁKOVICH, Institute of Education, University of Szeged, Hungary. E-mail: tibor.vidakovich@gmail.com

Acknowledgement

The authors would like to thank the school principal for permitting data collection and acknowledge the students and teachers for their valuable participation in this study.