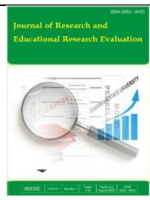
JRERE 13 (1) 2024 95-101

Journal of Research and Educational Research Evaluation



http://journal.unnes.ac.id/sju/index.php/jere

The Validity of Spatial Ability Test Instrument At Primary Level Based on Classical Test Theory

Rina Fardiana[1]✉, Kartono Kartono[2], Saiful Ridlo[3]

[1,2,3]Pascasarjana, Universitas Negeri Semarang

Article Info

History Articles
Received:
18 December 2023
Accepted:
19 January2024
Published:
30 March 2024

Keywords:
Validity, Spatial Ability, Test Instrument, Primary Level.

Abstract

This research developed and proved the validity of an instrument for assessing the spatial ability of phase C students in grade V based on classical test theory. The instrument was validated by 6 raters consisting of 3 assessment expert and 3 mathematics teachers. The results of the 6 raters' assessments were processed using Aiken V. The 20 question items were in accordance with the indicators, material coverage and constructs as measured by a calculated V value of not less than 0.92. The results showed that each question item was proven to be content valid. The instrument was tested on 40 grade V students at SDN 02 Petang Kalideres to prove the construct validity. The test results showed that 14 of the 20 questions were proven to be valid with r value of more than 0.308, 1 out of 6 questions should be eliminated due to a negative r value while the other 5 questions should be revised, and the instrument reliability was 0.659 in the medium category. This research produced 14 valid questions that can be used to measure grade V students' spatial abilities based on Merdeka Curriculum.

✉Correspondence Address :
Department Research and Educational Research Evaluation,
Postgraduate Program, Universitas Negeri Semarang, Indonesia 50229
E-mail : rinafard@students.unnes.ac.id

p-ISSN 2252-6420
e-ISSN 2503-1732

INTRODUCTION

Assessment is a source of information to which educational goals have been achieved and is the basis for further decision making (Herwin et al., 2019). Assessment provides an overview of students' achievements, thus giving them the opportunity to develop their abilities and giving teachers the opportunity to assess their own performance (Ole, 2020, p. 53). The important role of assessment requires teachers to put quality on the assessment instruments so that the information obtained correctly shows and measures the aspects that want to be measured (Weir cite in Hartati & Yogi, 2019, p. 60). In the Merdeka Curriculum, cognitive aspect achievements are described in Learning Outcomes (LO).

Learning outcomes in mathematics are divided into several elements, namely elements of number, measurement, geometry, algebra, and statistics and probability. Geometry learning in elementary schools aims to develop visual, verbal, drawing, logic and application of mathematics skills (Mursalin, 2016). NCTM (*The National Council of Teachers of Mathematics*) recommends that mathematics learning up to grade 12 focus on identification of shapes, spatial relationships, use of visualization, and spatial reasoning in problem solving (Choo et al., 2021). Understanding concepts and the ability to solve geometric problems requires spatial reasoning because students need to study shapes (Aini & Suryowati, 2022). Spatial ability is also an ability that supports various fields of work. Spatial abilities improve success in areas such as spatial planning activities, reading medical X-rays, and visualizing machines (Sudirman & Alghadari, 2020, p. 61). Students studying construction management require high spatial abilities to understand construction concepts (Porter & Glick, 2022, p. 18). In the military field, spatial abilities are also needed, for example to determine flight direction (Sweet et al., 2020, p. 1556).

The need to develop students' spatial abilities is answered by the Merdeka Curriculum. In Surat Keputusan Kepala Badan Standar, Kurikulum number 033/H/KR/2022 geometric elements focus a lot on spatial abilities such as constructing 3D shapes and recognizing spatial visualization. Good spatial abilities tend to provide good mathematics learning achievements (Putri, 2016, p. 346). A valid assessment instrument measuring students' spatial abilities is needed by teachers to assess spatial abilities so that teachers can take effective follow-up steps according to students' needs. Assessment has an impact on improving the quality of student achievement so that if the assessment is not of good quality, it can reduce the skills and quality of student achievement (Villarroel et al., 2020, p. 2). Opportunities to identify students' spatial abilities can be done through the development of visual-spatial ability assessment instruments (Chan, 2010, p. 4).

Research into the development and the test of the validity of spatial ability assessment instruments has been carried out. Ramful et al., (2017) developed instrument to measure spatial ability consisting of spatial orientation, mental rotation, and spatial visualization related to mathematics and STEM subjects in Australia. His research produced 30 valid and reliable multiple choice questions to measure spatial abilities based on factor and Rasch analysis. Research on the development of spatial ability and critical thinking instruments was also carried out by Kusumawardhana et al. (2020) but the instrument has not been tested on students due to Covid-19 pandemic. Putri et al. (2022) also conducted develop spatial ability test instrumen. The instrument consists of 6 descriptive questions which have been proven valid from the results of small-scale tests. However, the instrument developed refers to the 2013 Curriculum. In contrast to the previous studies, this research developed an instrument for assessing spatial ability to build space for phase C grade V students in accordance with the Merdeka Curriculum learning outcomes.

The spatial ability assessment instrument developed in this research has been proven to have content and construct validity. Content validity test means proving based on test content such as test specification, question forms, instructions, language used connected to the theoretical basis and the construct being measured (Sumintono & Widhiarso, 2015, p. 8). Construct validity is validity that proves a test or non-test instrument can measure the aspects to be measured based on the theoretical basis used when compiling the instrument (Iskandar, 2017). Analysis using classical test theory is still used to identify the quality of test items at the beginning (Allen & Yen, 1979). This research aims to develop and prove the validity of the content and construct of a spatial ability assessment instrument based on classical test theory. It is hoped that the results of this research can be a resource when assessing spatial abilities for primary students.

METHODS

This research is development research which refers to Borg & Gall (Sugiyono, 2021, p. 764). In this research, the procedure was carried out up to the third step, namely preliminary studies, planning and initial product development. In the preliminary study, the research conducted interviews with 5 grade V homeroom teachers to determine the needs and scope of material for 3D shapes of phase C grade V. A needs analysis needs to be carried out to obtain information that provides an overview of what kind of progress can be offered (Tumangger et al., 2022). The research continued with planning, namely preparing the test spesification consist of learning outcomes, materials, constructs and indicators of spatial ability.

The next stage is to prepare an instrument for assessing the spatial ability of phase C grade V students according to the test spesification and verifying the validity of the content. Proof of content validity was carried out by 6 raters consisting of 3 assessment expert and 3 grade V mathematics teachers. The assessment experts assessed the suitability of the question items to the constructs of spatial ability, cognitive level, learning outcomes and indicators. The teacher assesses the language, clarity of images, level of difficulty, and coverage of the material from the questions presented.

The validation sheet is presented using a 1-5 Likert scale. Teachers and assessment experts got different validation sheet according to their experties. The results of the rater's assessment are processed using Aiken V with the help of Microsoft Excel. One way of processing data to prove content validity is using Aiken V (Wulandari & Oktaviani, 2021). Aiken V can provide an overview of expert assessments of measuring instruments developed for both cognitive and non-cognitive measuring instruments (Wulandari & Oktaviani, 2021).

The item is valid if the validity coefficient is not less than V table (Aiken, 1985 cited in Jatiningtyas et al., 2022). The V value that must be met with the assessment of 3 raters is 0.92. Question items are proven valid if the V value of each item is more than or equal to 0.92.

Construct validity was proven through testing the use of instruments on 40 class V students at SDN 02 Petang, Kalideres, West Jakarta. Construct validity was proven using Classical Test Theory based on calculated r values, level of difficulty, and reliability estimates based on KR-20 with the help of Microsoft Excel.

RESULTS AND DISCUSSION

The results of interviews with five grade V teachers showed that the scope of material for 3D shapes of phase C students included the characteristics of prisms, prism components and nets. Geometry element learning outcomes based on Merdeka Curriculum shows the construct of phase C students' spatial abilities consisting of spatial perception, spatial orientation, and spatial visualization. Spatial perception is the ability to determine horizontal and vertical dimensions, where this ability appears when children are 9 years old (Oslon, 1975 cited in Mohler, 2008, p. 23). In a spatial orientation ability, students imagine how a shape looks from others' view and determine whether an object is located on the right or left side using one's own body as a reference (Ramful et al., 2017, p. 712). In the Purdue Spatial Visualization Test, spatial visualization is commonly measured by showing the ability to combine, separate, change, or rotate a shape (Lin & Chen, 2016). Spatial visualization is a form of manipulating the shape of 3-dimensional objects from the nets presented or imagining the shape of a space when sliced horizontally (Patahuddin et al., 2022). Table 1 presents brief specifications of the instruments developed.

Tabel 1. Test Specifications

Spatial Ability	Indicators
Spatial perception	Students are able to differentiate the length, height, width of cube/cuboid.
Spatial Orientation	Students are able to determine the position of cube/cuboid's components.
	Students are able to analyze the visualization of 3D shapes from various viewing directions.
Spatial Visualization	Students are able to construct and decompose cube/cuboid.

The spatial ability instrument test developed consists of 20 multiple choice questions with four answer choices. The results of the 3 teachers' assessments show that 19 numbers have Vcount not less than 0.92 unless number 4 has a Vcount = 0.75. Question number 4 has Vcount < Vtable so it is invalid. The image in question number 4 is considered unclear and needs to be changed. Question number 4 was revised according to feedback and submitted again to the rater. The Vcount for question number 4 is calculated again and the result is 1 where Vcount > Vtable so the item is proven to be valid. The assessment results from the assessment experts show that all numbers have Vcount not less than 0.92. All items are proven to be valid based on teacher and lecturer assessments with Vcount no less than Vtable.

The instrument consisting of 20 multiple choice questions was tested on 40 phase C grade V students at SDN 02 Petang Kalideres, West Jakarta. The value of r_{tabel} = 0.308 with a significance level of 5% and df = 38. The instrument validation results are shown in the table 2.

Tabel 2. Validation Result

Number	r count	Conclusion
1	0.52	Valid
2	0.48	Valid
3	0.46	Valid
4	0.25	Invalid
5	0.48	Valid
6	0.42	Valid
7	0.42	Valid
8	0.46	Valid
9	0.51	Valid
10	0.17	Invalid
11	0.57	Valid
12	-0.03	Invalid
13	0.03	Invalid
14	0.3	Invalid
15	0.48	Valid
16	0.42	Valid
17	0.41	Valid
18	0.39	Valid
19	0.11	Invalid
20	0.45	Valid

The biserial coefficient is a product moment correlation where the score of each item which is dichotomous is correlated with the total score which is continuum (Surapranata, 2009, p. 61). The question is proven valid if the correlation coefficient (Syofian, 2015 cited in Erfan et al., 2020). The results of data processing showed that 14 of the 20 questions were proven to be valid with value $r_{xy} > r_{tabel}$, while the other 6 questions, namely numbers 4, 10, 12, 13, 14, and 19, were proven to be invalid with value less than r table. Item number 12 needs to be eliminated because it has a negative r value, while the other 5 items need to be reviewed or revised because the r value is less than 0.308.

The level of difficulty is the ratio of participants who answered correctly to the total number of participants, which in classical theory can be expressed as the proportion of correct answers (Surapranata, 2009). The proportion of correct answers (p) is the most commonly used level of difficulty.

$$p = \frac{\sum x}{S_m N}$$

p = level of difficulty

$\sum$x= the number of test takers who answered correctly

Sm = skor maximum

N = the number of test takers

(Surapranata, 2009)

Difficulty level categories are created based on the p value of each question item as in the table 3.

Tabel 3. Level of Difficulty Category

Value of p	Criteria
$p < 0.3$	Hard
$0.3 \leq p \leq 0.7$	Middle
$0.7 < p$	Easy

(Bagiyono, 2017, p. 5)

The results of calculating the level of difficulty of the questions can be seen in table 4.

Tabel 4. Level of Difficulty

Value of p	Criteria	Number
$p < 0.3$	Hard	3.19. 20
$0.3 \leq p \leq 0.7$	Middle	1. 2. 4. 5. 7. 11. 12. 13. 14. 15. 16.17
$0.7 < p$	Easy	6. 8. 9. 10. 18

Table 4 shows that there are 3 difficult questions, 12 medium questions, and 5 easy questions. Reliability estimates for dichotomous items were carried out using dichotomous alpha calculations or KR-20 (Kuder & Richardson, 1937 cited in Azwar, 2022, p. 73).

Tabel 5. Correlation Coefficient Criteria

Correlation Coefficient	Category
$0.9 \leq r \leq 1$	Very high
$0.7 \leq r < 0.9$	High
$0.4 \leq r < 0.7$	Medium
$0.2 \leq r < 0.4$	Low
$r < 0.2$	Very low

(Erfan et al., 2020)

The results of reliability calculations with the help of Microsoft Excel in the test results show the value $r_{11} = 0.659$ then the consistency or reliability value of the instrument is in the medium category.

Fourteen questions of the instrument developed measure three constructs of spatial ability namely spatial perception, spatial orientation, dan spatial visualization. There are several constructs that shape spatial abilities. There is spatial ability constructs that is not included in the instrument, such as mental rotation. Despite this, the instrument can still be used to measure the spatial ability of fifth grade students. Differences in constructs in developing spatial ability assessment instruments may occur. Spatial abilities are divided into mental rotation, spatial perception, and spatial visualization (Linn & Petersen cited in Lin & Chen, 2016, p. 24). The spatial assessment instrumen developed by Ramful et al., (2017) measure spatial ability consisting of spatial orientation, mental rotation, and spatial visualization related to mathematics and STEM subjects in Australia. Hence the instrument developed in this research can be used to measure the abilities of class V students based on Merdeka Curriculum.

CONCLUSION

The instrument for assessing the spatial ability of 3D shapes for phase C grade V students was proven to be content valid, consisting of 20 multiple choice questions. All items have a Vcount value of no less than Vtable. The instrument measures 3 constructs, namely spatial perception, spatial orientation, and spatial visualization.

Based on the item difficulty category, there are 3 difficult questions, 12 medium questions, and 5 easy questions. The results of construct validity verification showed that 14 of the 20 items were proven to be valid with r count of greater than r table. One item needs to be eliminated because it has a negative r value, while the other 5 items need to be reviewed or revised because the r value is less than 0.308. The reliability of the instrument based on the value r=0.659 is in the medium category. Suggestions for further research are that it would be better to increase the number of rater and include spatial ability experts also to assess the content validity of the instrument.

REFERENCES

Aini, N., & Suryowati, E. (2022). Mengeksplor penalaran spasial siswa dalam menyelesaikan soal geometri berdasarkan gender. *Mosharafa: Jurnal Pendidikan Matematika*, *11*(1), 61–72.

Allen & Yen. (1979). *Introduction to measurement theory* (pp. 1–320). Brooks/Cole Publishing Company.

Azwar, S. (2022). *Reliabilitas dan validitas* (IV). Pustaka Pelajar.

Bagiyono, B. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1 The Analysis of Difficulty Level and Discrimination Power of Test Items of Radiography Level 1 Examination. *Widyanuklida*, *16*(1), 1–12.

Chan, D. W. (2010). Developing the Impossible Figures Task to Assess Visual-Spatial Talents Among Chinese Students : A Rasch Measurement Model Analysis. *Gifted Child Quarterly*, *54*(1), 59–71.

Choo, S., Park, S., & Nelson, N. J. (2021). Evaluating spatial thinking ability using item response theory: Differential item functioning across math learning disabilities and geometry instructions. *Learning Disability Quarterly*, *44*(2), 68–81.

Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Analisis kualitas soal kemampuan membedakan rangkaian seri dan paralel melalui teori tes klasik dan model rasch. *Indonesian Journal of Educational Research and Review*, *3*(1), 11–19.

Hartati, N., & Yogi, H. P. S. (2019). Item analysis for a better quality test. *English Language in Focus*, *2*(1), 59–70.

Herwin, H., Tenriawaru, A., & Fane, A. (2019). Math elementary school exam analysis based on the Rasch model. *Jurnal Prima Edukasia*, *7*(2), 106–113.

Iskandar, A. (2017). *Teknik analisis validitas konstruk dan reliabilitas instrumen test dan non test dengan software Lisrel*. 77–154.

Jatiningtyas, P. D., Kartono, K., & Mindyarto, B. N. (2022). Test Instruments to Measure Non-Routine Mathematics Problem Solving Ability Grade IV Elementary School Students. *Journal of Educational Research and Evaluation*, *6*(3), 407–414.

Kusumawardhana, I. N., Arnyana, I. B. ., & Dantes, N. (2020). Pengembangan instrumen pengukuran kemampuan spatial sense (kss) dan kemampuan berpikir kritis pada materi bangun ruang siswa kelas v sd. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, *10*(2), 51–61.

Lin, C. H., & Chen, C. M. (2016). Developing spatial visualization and mental rotation with a digital puzzle game at primary school level. *Computers in Human Behavior*, *57*, 23–30.

Mohler, J. L. (2008). A review of spatial ability research. *Engineering Design Graphics Journal*, *72*(2), 19–30.

Mursalin. (2016). Pembelajaran geometri bidang datar di sekolah dasar berorientasi teori belajar. *Jurnal Dikma*, *4*(2), 250–258.

Ole, F. C. (2020). Development and Validation of Teachers' Practices on Formative Assessment Scale (TPFAS): A Measure Using Feedback Loop Model. *International Journal of Education*, *13*(1), 53–62.

Patahuddin, S. M., Ramful, A., Lowrie, T., & Bholoa, A. (2022). Subtleties in spatial visualization maneuvers: Insights from numerical solutions. *Journal of Mathematical Behavior*, *67*(June 2021).

Porter, D., & Glick, S. (2022). Spatial visualization in construction management education: A review and validation of the literature using professionals and related practitioners. *International Journal of Construction Education and Research*, *18*(1), 17–31.

Putri, H. E., Nuraeni, F., Haftani, D. A., Rosalia, E., & Pratama, N. (2022). Spatial sense ability instrument for primary school students. *Metode Didaktik Jurnal Kependidikan Ke-SD-An*, *18*(1), 53–64.

Putri, R. O. E. (2016). Peran kemampuan spasial siswa dalam menyelesaikan masalah matematika yang berkaitan dengan geometri. *Prosiding Konferensi Nasional Penelitian Matematika Dan Pembelajarannya (KNPMP I)*, *knpmp I*, 345–352.

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment*, *35*(7), 709–727.

Sudirman, S., & Alghadari, F. (2020). Bagaimana Mengembangkan Kemampuan Spasial dalam Pembelajaran Matematika di Sekolah?: Suatu Tinjauan Literatur. *Journal of Instructional Mathematics*, *1*(2), 60–72.

Sugiyono. (2021). *Metode Penelitian Pendidikan*. Alfabeta.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan* (B. Trim & D. Binikna (eds.)). Trim Komunikata.

Surapranata, S. (2009). *Analisis, validitas, reliabilitas dan hasil interpretasi tes implementasi kurikulum 2004*. Remaja Rosdakarya Offset.

Surat Keputusan Kepala Badan Standar, Kurikulum, dan A. P. (n.d.). *Surat Keputusan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan nomor 033/H/KR/2022*. 132–142.

Sweet, J., Coyne, J., Herdener, N., & Ries, A. J. (2020). Does a New Version of the Direction Orientation Task Measure Spatial Ability? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 1556–1560.

Tumangger, M. H., Kartono, & Ridlo, S. (2022). Analisis kebutuhan pengembangan instrumen penilaian kemampuan berpikir kreatif pada mata pelajaran matematika untuk siswa sekolah dasar. *Jurnal Basicedu*, *6*(5), 9170–9179.

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, *57*(1), 38–49.

Wulandari, I., & Oktaviani, N. M. (2021). Validitas bahan ajar kurikulum pembelajaran untuk pendidikan guru sekolah dasar. *Jurnal Cakrawala Pendas*, *7*(1), 90–98.