The Development of the Validated Instruments to Assess the Ability of the 9th Graders in Solving Diagnostic Geometry Problem in Two Characteristic Classes

Ovan[1], Yusuf Fuad[2], and Mega Teguh Budiarto[2]

[1] Doctoral Program of Mathematics Education, State University of Surabaya, Indonesia
[2] Mathematics Study Program, State University of Surabaya, Indonesia
ovan.21029@mhs.unesa.ac.id

Abstract. This study aims to develop a diagnostic test for evaluating the ability of the 9th graders junior high school students in solving geometry problems. The instruments utilized include the Initial Ability Test (IAT) and the Geometry Problem Solving Test (GPST). Before being given to the participants, the research instruments were validated by three senior lecturers holding doctoral degrees in mathematics education. Furthermore, the readability test was conducted by two math teachers and two students from a public junior high school in Surabaya. After that, 45 students of both grades mentioned above were asked to work on both tests in two consecutive days. The participants were grouped into high math ability groups, medium math ability groups, and the low math ability group. Based on these results, IAT and GPST was found to be valid, reliable, significantly easy (IAT) and moderate significant (GPST), and very good discriminating power. More specifically, with t-test and significance level 휘 = 0.05, the results on IAT confirm that there is no significant difference between the scores of the two classes, which have significantly different characteristics. Whereas in GPST there is a significant difference between the scores of the two classes. Thus, from the assessment of the validator and the analysis of the test items, both instruments are categorized as valid and can be used to evaluate the ability of the 9th graders junior high school students in solving geometric problems. Instruments that have been validated will be tested for a larger population with expanded materials.

Keywords: Validated Instrument, Problem Solving Ability, Geometry, Characteristic Classes.

1 Introduction

Problem-solving is a key component in the current curriculum of mathematics. However, many students experience problems in problem-solving, especially geometry material [1,2,3]. Problem solving is a mental activity consisting of various skills and actions intended to get the right solution [4,5]. Geometry problems might make students struggle in understanding, analyzing, and finding answers [6]. Moreover, geometry problem solving is a process which is formulated in the form of non routine questions

A. Mustofa et al. (eds.), *Proceedings of the International Joint Conference on Arts and Humanities 2023 (IJCAH 2023)*, Advances in Social Science, Education and Humanities Research 785,
https://doi.org/10.2991/978-2-38476-152-4_196

which are designed for students to not easily find answers or solve these questions [1]. Geometry problem-solving is an ongoing process of solving a problem for students until it no longer becomes a problem [7,8,9]. Therefore, geometry problem-solving is a process carried out by students in completing a task with unknown concepts and procedures on how to solve it.

Geometry needs to be understood, learned, and interpreted, as it is one of the intellectual competencies that students must have. It also can be used to solve problems in students' everyday lives, for example calculating the area of a pyramid-shaped house, the area of a kite, the volume of water in a cube-shaped aquarium, the volume of water in a cube-shaped bathtub, and the area of pizza that is cut into small triangular pieces. Geometry problem-solving contains more abstract concepts than other areas of mathematics, especially three-dimensional objects whose content requires students to think comprehensively using their imagination [10]. Therefore, to be able to study geometry well, students are required to master the basic skills of geometry, skills in proving, and skills in making basic geometric shapes. Teaching geometry can also train students to think critically, creatively, and structurally [11].

The problem-solving phase consists of five stages: analysis, exploration, planning, implementation, and verification [9]. Analysis is understanding the problem seriously. Exploration is an activity to find certain concepts in solving mathematical problems. Planning is devising a strategy to solve the problem. Implementation is implementing a strategy to solve the problem. Lastly, verification is re-checking the solution to the problem that was carried out.

There have been a number of studies exploring on junior high school students' ability in geometry. The results of a research done at a junior high school in the city of Bindura Africa with a total of 100 students show that students can develop their abilities when studying geometry [2]. Furthermore, students' attitudes toward the use of geometry were positive and many of the students believed that geometry was a valuable and necessary topic for them in their future careers. On the contrary, 80% of students in the research do not like solving geometry problems. Another research at a junior high school in Malang shows that junior high school students' problem-solving abilities were still low because students were not able to distinguish, identify, and categorize geometric problems properly [12]. In addition, a research on 63 junior high school students in Kirsehir Turkey found that students could solve geometry problems, especially on triangle material [6]. Relating to gender, boy students at SMP Al-Hikmah 2 Benda can solve heuristic problems well, especially in number patterns [13].

Several studies measuring the validity of geometric instruments have also been carried out, both the validity of multiple-choice questions, essays, and Likert scale questionnaire [14,15,16]. Furthermore, a research has focused on the validity and reliability of STEM instruments to evaluate the abilities of grade 5, 6, 7, and 8 secondary school students from fifteen different schools in ten provinces in Turkey [17]. A valid instrument is a test set that meets the validity criteria and can be used to measure students' abilities. The benefit of having a valid instrument is that it can measure students' abilities precisely and measurably. Therefore, this study focuses on conducting instrument validity to evaluate the ability of ninth grade students in solving geometric problems.

2 Methods

This research was conducted on the 9^{th} graders students in a private junior high school in Mojokerto, East Java. The instrument consists of the Initial Ability Test (IAT) and the Geometry Problem Solving Test (GPST). IAT covers algebra, arithmetic, and geometry, while GPST covers triangles, rectangles, trapezoids, cubes, and pyramids. The research instrument was in the form of an essay and was first validated by three senior lecturers with doctoral degrees in mathematics education with criteria adapted from [18]. By using a Likert scale (maximum scale of 5) and content and feasibility validations, the three senior lecturers stated that both instruments were valid and suitable for use with minor revisions. The validation results are presented in Table 1 below.

Table 1. Results of the IAT and TPMG Content Validation and Feasibility from Three Validators

No	Validators	Validation Score (Maximum 5)		Category	Suggestion
		IAT	GPST		
1	Validators 1	4.40	4.43	Valid & Eligible	Minor revisions for the questions
2	Validators 2	4.40	4.46	Valid & Eligible	Minor revisions for the questions
3	Validators 3	4.90	4.90	Valid & Eligible	Minor revisions for the questions
	Average	4.57	4.60	Valid & Eligible	Valid if score> 3.50

After repairing and perfecting the two instruments, by accommodating suggestions and input from the three validators, the readability test was carried out by two mathematics teachers and two students from a state junior high school in Surabaya who were not included in the research sample class. The results of the readability test recommended that both instruments be declared feasible and with sentences that are easily understood by the 9^{th} graders students.

Both instruments were tested on two classes that were purposively selected from eight classes of the 9^{th} graders, namely the 9^{th} graders A and the 9^{th} graders E. 45 students from both classes, under direct supervision by the researchers and mathematics teachers, were asked to work on both tests in two consecutive days with a duration of 80 minutes each. Based on data from both instruments, the results are grouped into groups of mathematical abilities [19] which is presented in Table 2 below.

Table 2. Category of Mathematics Ability of the 9^{th} Graders Junior High School Students

No	Score Range	IAT	GPST	Category
1	Score ≥ 80	21	7	High
2	$65 \leq$ Score ≤ 79	10	18	Medium

3	Score < 65	14	20	Low
	Amount	45	45	

IAT and GPST empirical validation tests using product moment Pearson correlation, Cronbach alpha, difficulty index, and discrimination index [20] are categorized in Table 3 below.

Table 3. Categories of Empirical Validation of the 9th Graders Junior High School Students

No.	Aspects of Empirical Validation	Category
1	Validity test	Score > 0,29
2	Reliability Test	Score > 0,60
3	Difficulty Level Test	Score > 0,70
4	Discriminating Power Test	Score > 0,39

Furthermore, to test the two instruments from two different classes, a t-test was carried out to find out significant differences between the two tested classes with a significance level of $\alpha = 0.05$ [21].

3 Result

The results of the empirical validation test of the two instruments can be presented in Table 4 below.

Table 4. Results of the Empirical Validation Test from IAT and GPST

No.	Aspects of Empirical Validation	IAT	GPST
1	The average test value of the validity of the questions	0.66	0.80
2	The value of the reliability test questions	0.80	0.71
3	The average test value of the difficulty level of the questions	0.72	0.62
4	The average value of the test of discriminating power of questions	0.51	0.53

Referring to Table 4, IAT is categorized as valid because from the validity test using the product moment Pearson correlation, the obtained average value is 0.65 (>0.29). The instrument is categorized as reliable because the reliability test using Alpha Cronbach gets a value of 0.80 (>0.60). The instrument is in the easy significant category, because from the difficulty level test using the index of difficulty, the value is 0.72 (>0.70). The instrument also has very good discriminating power because the discriminating power test using the discrimination index obtains a value 0.51 (>0.39). When given to the subjects, questions number 2,3 & 4 have a difficulty level value that is less than the limit value 0.70. On the other hand, questions number 1 & 8 have a discriminating power value that is less than the limit value 0.40.

For the GPST instrument, it was also found that the instrument can be categorized as valid because the validity test using the product moment Pearson correlation obtained a value of 0.80 (>0.29). The instrument is categorized as reliable because the reliability test using Alpha Cronbach gets a value of 0.71(>0.60). The instrument is categorized as moderately significant because from the difficulty level test using the difficulty index 0.62 (<0.70), the values of the three items on the instrument show that the difficulty level value is 0.70. Furthermore, the instrument has a very good discriminating power because its discrimination index has the value of 0.53 (>0.39). Thus, the IAT instrument is in the category of valid, highly reliable, easy, and has a very good discriminating power, while the GPST instrument is in the category of valid, moderately reliable, quite easy, and also has a very good discriminating power because the two instruments have met the criteria of content validity and feasibility, item validity test, item reliability test, item difficulty level test, and item discrimination test.

More specifically, to see if there was a difference, a t-test was carried out to find out a significant difference between the two test classes with the significance level of $\alpha = 0.05$. Based on statistical tests, the IAT instrument confirmed that there was no significant difference between the scores of the two classes, which had significantly different characteristics, with $\bar{x}_A = 53.52$; $s_A = 20.80$; $\bar{x}_E = 62$; $s_E = 12.31$. On the other hand, there was a significant difference between the scores of the two classes with $\bar{x}_A = 20.78$; $s_A = 5.66$; $\bar{x}_E = 16.36$; $s_E = 6.66$ in GPST.

4 Discussion

Data collection was carried out in two different classes which were divided based on the students' gender. It means boy students (the 9th graders A) and girl students (the 9th graders E) were separated since the private junior high school being researched is a religious-based school that requires separate classes for boy and girl students. However, data collection using IAT and GPST was carried out simultaneously. Furthermore, this research did not touch on the gender aspects of students as in some previous research [22,23,24,25]. It is more focused on the validity of the research instrument to evaluate the ability of class IX students in solving geometric problems.

The IAT instrument was adapted from junior high school students' National Examination questions which included material on arithmetic, algebra, and geometry. Initially the instrument was composed of ten questions consisting of arithmetic (3 questions), algebra (4 questions), geometry (2 questions), and statistics (1 question). Three senior lecturers with a doctorate degree in mathematics provided advice and input regarding the two instruments. The first lecturer recommended that the geometry questions that were still at C3 and KD3 levels should be revised. The second expert gave input regarding the time allocation that should be matched according to the level of complexity of the questions to be worked on and regarding question number 10 that needs to be revised because it contains an element of ambiguity (height or weight). Lastly the advice from the third senior lecturer was about the number of IAT questions which needed to be adapted to the abilities of students. Therefore, the instrument was improved based on the suggestions and input of the three senior lecturers.

The GPST instrument was also adapted from junior high school students' National Exam questions which included triangles, rectangles, trapezoids, cubes, and pyramids. Initially, the instrument consisted of six questions consisting of cubes (task number 1), beams (task number 2), beams and pyramids (task number 3), prisms (task number 4), prisms and squares (task number 5), and cones (task number 6). The three senior lecturers also provided input and suggestions for improving this instrument. The first senior lecturer suggested to add commands to direct students to KD4 and C4 levels. The second lecturer's input was to fix some errors in writing mathematical symbols and notation. Lastly, the third senior lecturer did not provide input and suggestions. Taking into account the suggestions and input from the three senior lecturers, the number of GPST questions was reduced to three questions, but there was a change in materials in some questions. Based on the validator's suggestions and input, the three questions were revised so that the GPST materials tested on students were triangles (task number 1), trapezoids and squares (task number 2), and cubes and pyramids (task number 3).

This research is in line with previous research, where research tools in the form of instruments to evaluate students' mathematical abilities need to be validated [14,15,26,27]. Furthermore, IAT, which emphasizes that there is no significant difference between the scores of the two classes that have significantly different characteristics, needs to be explored further, for example with regard to the curriculum taught in schools and the mathematical ability of each student. On the other hand, GPST, in which there is a significant difference between the scores from both classes, can be developed on other geometry materials.

There is no end to talk about solving math problems, because almost all countries in the world place problem-solving as the main element in the mathematics learning curriculum. The results of this study contributed to obtaining a valid instrument. Thus, from the assessment of three senior lecturers and the analysis of the test items, both instruments are categorized as valid and can be used to evaluate the ability of class IX students in solving geometric problems. Instruments that have been validated will be tested for a larger population with some material expansion.

5 Conclusion

The results of the validation of the three senior lecturers in mathematics education showed that both instruments were valid and suitable for use with minor revisions. The results of the test item analysis of the two instruments show that both instruments are categorized as valid. Thus, the tested IAT and GPST instruments can be used to evaluate the ability of the 9th graders students in solving geometric problems.

Recommended Future Study

The instruments in this study that have been validated can be tested for a larger population with some extension of the material. Other researchers can use these two instruments on similar populations and samples and the same study focus.

Acknowledgements

The first author would like to express their gratitude to the Higher Education Funding Center (BPPT) of the Ministry of Education, Culture, Research and Technology and the Indonesia Endowment Fund for Education (LPDP) of the Ministry of Finance of the Republic of Indonesia for supporting this doctoral research process.

References

1. Gridos, P., Avgerinos, E., Mamona-Downs, J., Vlachou, R.: Geometrical Figure Apprehension, Construction of Auxiliary Lines, and Multiple Solutions in Problem Solving: Aspects of Mathematical Creativity in School Geometry. International Journal of Science and Mathematics Education 20(3), 619–636 (2022)
2. Sunzuma, G., Masocha, M., Zezekwa, N.: Secondary School Students' Attitudes towards their Learning of Geometry: A Survey of Bindura Urban Secondary Schools. Greener Journal of Educational Research. 3(8), 402–410 (2013)
3. Syam, H., Sutawidjaja, A., Sa'dijah, C., Abadyo: Junior high students' critical thinking in geometry problem solving. Universal Journal of Educational Research 8(11), 5880–5887 (2020)
4. DeJarnette, A. F., González, G.: Building Students' Reasoning Skills by Promoting Student-Led Discussions in an Algebra II Class. The Mathematics Educator 23(1), 3–23 (2013)
5. Foshay, R., Kirkley, J.: Principles for Teaching Problem Solving. PLATO Learning, Minnesota (1998)
6. Özçakır, B., Aytekin, C., Altunkaya, B., Doruk, B. K.: Effects of Using Dynamic Geometry Activities on Eighth Grade Students' Achievement Levels and Estimation Performances in Triangles. Participatory Educational Research 2(3), 43–54 (2015)
7. Mason, J., Burton, W. L., & Stacey, K.: Thinking mathematically. Pearson, Ney Jersey (2010)
8. Polya, G.: How T o Solve It: A New Aspect of Mathematical Method (2nd ed.). Princeton University Press, New Jersey (1957)
9. Schoenfeld, A. H.: Mathematical problem solving. In Academic Press, Inc (1st ed.). Academic Press, United States (1895)
10. Bülbül, B.Ö.: Factors affecting prospective mathematics teachers' beliefs about geometric habits of mind. Journal of Pedagogical Research 5(2), 36–48 (2021)
11. Ovan, Budiarto, M. T., Fuad, Y.: Berpikir Kritis dan Pemecahan Masalah Matematika (1st ed.). Eureka Media Aksara, Purbalingga (2023)
12. Basri, H., Purwanto, As'ari, A. R., Sisworo: Investigating critical thinking skill of junior high school in solving mathematical problem. International Journal of Instruction 12(3), 745–758 (2019)
13. Ruliani, I. D., Fuad, Y., Ekawati, R.: Heuristik siswa perempuan smp dalam memecahkan masalah pola bilangan. 5(3), 625–638 (2022)
14. Akhter, N., Usmani, A. A., Iqbal, S.: Development and Validation of Multiple-Choice Test of the Geometry Part of Mathematics for Secondary Class. Global Social Sciences Review 4(2), 203–210 (2019)
15. Alhunaini, S., Osman, K., Abdurab, N.: The Development and Validation of Mathematical Thinking Beliefs (MTB) Instrument. Eurasia Journal of Mathematics, Science and Technology Education 17(11), 1–13 (2021)

16. Sudihartinih, E., Prabawanto, S.: Test instrument validation in plane geometry using rasch model. Mathematics Education Journals 4(2), 102–115 (2020)
17. Benek, I., Akcay, B.: Development of STEM attitude scale for secondary school students: Validity and reliability study. International Journal of Education in Mathematics, Science and Technology 7(1), 32–52 (2019)
18. Retnawati, H.: Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, dan Psikometrian (1st ed.). Parama Publishing, Dhaka (2016)
19. Azwar, S.: Reliabilitas dan Validitas (4th ed.). Pustaka Pelajar, Yogyakarta (2012)
20. Masriyah, Rahaju, E. B., Rosyidi, A. H., & Hidayat, D.: Asesmen Untuk Pembelajaran Matematika (2nd ed.). Unesa University Press, Surabaya (2022)
21. Sukestiyarno, Y. L.: Olah Data Penelitian Berbantuan SPSS. Unnes, Semarang (2013)
22. Casey, B. M., Ganley, C. M.: An examination of gender differences in spatial skills and math attitudes in relation to mathematics success: A bio-psycho-social model. Developmental Review 60(October 2020), 100963 (2021)
23. Matteucci, M., Mignani, S.: Investigating gender differences in mathematics by performance levels in the Italian school system. Studies in Educational Evaluation 70(June 2020), 101022 (2021)
24. Oppermann, E., Vinni-Laakso, J., Juuti, K., Loukomies, A., Salmela-Aro, K.: Elementary school students' motivational profiles across Finnish language, mathematics and science: Longitudinal trajectories, gender differences and STEM aspirations. Contemporary Educational Psychology 64(November 2020), 101927 (2021)
25. Rahe, M., Quaiser-Pohl, C.: Can (perceived) mental-rotation performance mediate gender differences in math anxiety in adolescents and young adults? Mathematics Education Research Journal 35(1), 255–279 (2023)
26. Annisah, S., Zulela, Boeriswati, E., Wildaniati, Y., Supriatin, A.: Test instrument development of mathematical problem solving skills. International Journal of Advanced Science and Technology 29(6), 1483–1492 (2020)
27. Zainuddin, M., Subali, B., Jailani: Construct Validity of Mathematical Creativity Instrument: First- Order and Second-Order Confirmatory Factor Analysis Source details. International Journal of Instruction 12(3), 595–614 (2019)