DEVELOPMENT OF ASSESSMENT FOR LEARNING (AFL) INSTRUMENTS ON CONGRUENCE AND SIMILARITY SUBJECT FOR GRADE IX JUNIOR HIGH SCHOOL STUDENTS

Nurul Hidayah[1), Ruslan[2), Awi Dassa[3)
[1,2,3]Department of Mathematics Education, Universitas Negeri Makassar
[1,2,3]Jl. Daeng Tata Raya, Parang Tambung, Makassar, Indonesia
Email: nrlhhidayah@gmail.com[1), ruslan_math_unm@yahoo.com[2),
awi.dassa@unm.ac.id[3)

Abstract:
This study aims to (1) know the development process assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students; (2) produce an assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students that are valid and reliable. This present study is development research involving 126 grade IX students of SMP Islam Athirah 2 Makassar. The results showed that: (1) the process of developing assessment for learning (AfL) instruments on congruence and similarity subject for grade IX junior high school students was carried out through the stages of compiling instrument specifications (including determining the purpose of the instrument, compiling the instrument grid, choosing the shape and format of the instrument, determining the length of the instrument, and determining the instrument scales and scoring systems), writing instruments, reviewing instruments (content validity test and instruments improvement), conducting instrument trials, interpreting measurement results, analyzing instruments (construct validity and reliability tests), and assembling instruments; (2) the results of the development of assessment for learning (AfL) instruments on congruence and similarity subject for grade IX junior high school students consisted of a learning process assessment test, a learning outcome assessment test, and a rubric for assessing cognitive aspects, psychomotor aspects, and affective aspects.

Keywords: Assessment for Learning (AfL), Development Procedures, Validity and Reliability, Congruence and Similarity

PENGEMBANGAN INSTRUMEN *ASSESSMENT FOR LEARNING* (AFL) PADA MATERI KEKONGRUENAN DAN KESEBANGUNAN UNTUK PESERTA DIDIK KELAS IX SMP

Abstrak:
Penelitian ini bertujuan untuk (1) mengetahui gambaran pengembangan instrumen assessment for learning (AfL) pada materi kekongruenan dan kesebangunan untuk peserta didik kelas IX SMP; (2) menghasilkan instrumen assessment for learning (AfL) pada materi kekongruenan dan kesebangunan untuk peserta didik kelas IX SMP yang

Nurul Hidayah[1], Ruslan[2], Awi Dassa[3]

valid dan reliabel. Jenis penelitian ini adalah Research and Development yang melibatkan 126 peserta didik kelas IX SMP Islam Athirah 2 Makassar. Hasil penelitian menunjukkan bahwa: (1) gambaran pengembangan instrumen assessment for learning (AfL) pada materi kekongruenan dan kesebangunan untuk peserta didik kelas IX SMP dilaksanakan melalui tahapan menyusun spesifikasi instrumen (mencakup menentukan tujuan, menyusun kisi-kisi, memilih bentuk dan format, menentukan panjang, serta menentukan skala dan sistem penskoran), menulis instrumen, menelaah instrumen (uji validitas isi dan memperbaiki instrumen), melakukan uji coba instrumen, menafsirkan hasil uji coba, menganalisis instrumen (uji validitas konstruk dan reliabilitas), dan merakit instrumen; (2) hasil pengembangan instrumen assessment for learning (AfL) pada materi kekongruenan dan kesebangunan untuk peserta didik kelas IX SMP terdiri dari tes penilaian proses belajar, tes penilaian hasil belajar, dan rubrik penilaian aspek kognitif, keterampilan, dan sikap.

Kata Kunci: Assessment for Learning (AfL), Prosedur Pengembangan, Validitas dan Reliabilitas, Kekongruenan dan Kesebangunan

INTRODUCTION

Assessment is part of the learning process. Assessment is an integral tool that is used as a tool to see and analyze whether students have achieved the expected learning outcomes and find out whether the learning process is following the objectives or still requires development and improvement (Putri, 2022). Assessment is expected to be able to become an instrument of quality assurance, quality control, and quality improvement of the education system (Raharjo, Handayani, Jauhari, Juanita, & Waspodo, 2019).

For the teacher's assessment to be effective and meaningful, the teacher needs to (1) formulate clear, firm, and measurable learning targets to be assessed; (2) ensure that the selected assessment technique by each learning target; and (3) ensure that the assessment technique is by the current and future needs of students (Bulan, 2021). Current assessment techniques must have a real contribution to achieving the quality of graduates with critical and creative thinking skills, communication skills, and collaboration skills, as well as materials for consideration and evaluation of learning systems. Therefore,

the assessment must be able to measure the process and results of students' abilities.

In learning in the 2013 curriculum, the assessment not only relies on the assessment of results but also considers assessment in terms of the process so that students' abilities can be measured as a whole and continuously. One type of assessment that has strong relevance to the 2013 curriculum is the assessment for learning (AfL) (Yuwono & Syaifuddin, 2017). The emphasis on this type of assessment lies not on the knowledge that has been mastered but on the process and learning outcomes of students. AfL needs to be applied in the learning processes. The AfL needs to be applied in learning processes because it can be used to guarantee accurate information about the abilities or competencies of students.

The concept of AfL is not a new thing in educational assessment. It is a form of assessment applied in the context of improving the quality of learning. AfL is better planned, directed, and focused. At least this is reflected in the meaning of AfL put forward in the Assessment Reform Group (2002), which says that assessment for learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

The emphasis on the meaning of AfL lies in the process of obtaining and utilizing information. Information is obtained through collaboration between teachers and students and it is used by them (teachers and students) for improvement of the quality of subsequent learning. For teachers, information is used to improve and refine teaching strategies according to the real needs of their students. Meanwhile, for students, it can be used as a basis for changing better learning strategies.

Based on observations of one of the mathematics teachers at the junior high school level, it was found that the application of AfL in schools still needed to be improved. Summative assessment techniques still dominate the teacher's assessment. So far, the assessment has only emphasized mastery of the concepts captured by objective and subjective written tests as measuring tools in the form of learning achievement tests (Thalaat, 2022; Yudha, Masrukan, & Djuniadi, 2014). Current learning assessment still prioritizes final assessment without considering process assessment (Yusron & Sudiyatno, 2021). The final assessment is only limited to the final exam of learning and without any feedback for improving learning. Assessment needs to cover cognitive, affective, and psychomotor aspects thoroughly.

Nurul Hidayah[1), Ruslan[2), Awi Dassa[3)

Based on the background above, the AfL instrument has the potential to be developed. In this research, AfL was developed to measure the junior high school student's ability on congruence and similarity subjects. This study aimed (1) to find out the process of developing assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students and (2) to produce assessment for learning (AfL) instruments on congruence and similarity subject for grade IX junior high school students that are valid and reliable.

METHODS

The type of research used is research and development (R&D) which aims to produce assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students that are valid and reliable. The research was conducted from January to February 2023 at SMP Islam Athirah 2 Makassar. The subjects in the study were students of class IX at SMP Islam Athirah 2 Makassar, consisting of 126 people spread across 5 classes. The research instruments used in this study were a content validity assessment instrument and assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students.

The development procedure carried out in this study refers to the instrument development procedure (test and non-test) proposed by Mardapi (2008). The steps taken are as follows.

1. Determining Instrument Specification

   In preparing instrument specifications, things that must be considered are determining measurement objectives, preparing instrument grids, selecting instrument shapes and formats, determining instrument lengths, and determining instrument scales and scoring systems.

2. Writing Instruments

   Writing the instrument is a step to express the indicators into test items, either questions or statements whose characteristics are by the details on the grid that has been made.

3. Reviewing Instruments

   Instrument review is carried out by two experts or validators. The aspect examined is the suitability of the test items with the indicators detailed on the instrument grid.

4. Testing Instruments

   The implementation of the instrument trial was carried out in class IX of SMP Islam Athirah 2 Makassar which functioned as a means of obtaining empirical data on the feasibility level of the instrument that has been compiled.

5. Interpreting Instrument Testing Results

   Interpretation of measurement results used criteria depending on the scale and scoring system.

6. Analyzing Instruments

   Furthermore, quantitative analysis was carried out to determine the construct validity and reliability of the instrument based on the empirical data from the trial results.

7. Assembling Instruments

   Instruments that have been analyzed were then repaired and then assembled into one complete instrument unit.

   The data analysis carried out in this research is content validity, construct validity and reliability.

1. Content Validity

   According to Susetyo (2015), content validity is validity that will check compatibility between items made with indicators, material, or learning objectives set. Hamzah (2014) states that the content validity of an instrument is concerned with the extent to which an instrument measures the level of mastery of the contents of a particular material that should be mastered by teaching objectives.

   Content validity testing is done by examining the items and the suitability of the contents based on the instrument grid. Therefore, the content validity of an instrument does not have a statistically calculated quantity, but the validity of a test is based on a review against the instrument grid (Sappaile, 2007). To determine the content validity of both test and non-test instruments, the most common way is through expert judgment. The pattern used in this analysis uses the Gregory pattern (in Ruslan, 2009), as shown in the following figure.

		Validators 1	
		The weak relevance of the item is scored 1 or 2	The relevance of grain strength is rated at 3 or 4
Validators 2	The weak relevance of the item is scored 1 or 2	A	B
	The relevance of grain strength is rated at 3 or 4	C	D

Figure 1. Agreement between Validators

The formula used is as follows.

$$koefisien\ konsistensi\ internal = \frac{D}{A + B + C + D} \tag{1}$$

A : The number of items that have weak relevance according to the two validators

B : The number of items that have weak relevance according to one of the validators and strong relevance according to other validators

C : The number of items that have weak relevance according to one of the validators and strong relevance according to other validators

D : The number of items that have strong relevance according to the two validators

If the validity coefficient is high ($\geq$ 0.75) then it can be said that the measurement or intervention is valid (Gregory in Ruslan, 2009).

2. Construct Validity

Construct validity refers to the extent to which an instrument can measure the meanings contained in the material to be measured (Uno & Satria, 2012). Meanwhile, Yusuf (2017) states that construct validity questions whether what is being asked is an important part of a concept or is part of an instrument that has been compiled. Furthermore, Hamzah (2014) stated that construct validity is validity which concerns how far the items are able to measure what is really to be measured by a specific concept or predetermined conceptual definition. Ruslan (2006) also stated that construct validity indicates the extent to which an instrument is able to measure certain properties or theoretical constructs.

Testing the instrument on construct validity can be done through factor analysis (Hintze in Ruslan, 2006). There are two approaches to factor analysis, namely, the exploratory factor analysis using the principal component analysis (PCA) method and the confirmatory factor analysis

using the maximum likelihood (ML) analysis method (Ruslan, 2006). Sappaile (2005) states that to determine homogeneous items, attention is paid to factor loading That is greater than or equal to 0.30.

Construct validity testing was carried out after the instruments were tried out in class IX at Athirah 2 Islamic Middle School Makassar. This implementation aims to obtain information and prove the construct validity of the instrument empirically. It aims to show the extent to which the instrument reveals a particular ability or construct to be measured and to determine the suitability between theoretical concepts and empirical data. Construct validity testing of the assessment for learning instrument was analyzed using the Confirmatory Factor Analysis (CFA) approach through the maximum likelihood (ML) analysis method.

3. Reliability

The word reliability is derived from the word reliable which means trustworthy (Setiyawan & Wijayanti, 2020). Sappaile (2005) states that the measurement results are declared reliable if they are consistent and result from several times the measurement. The reliability test on the instrument was carried out using the Cronbach alpha and KR-20. The reliability test on the AfL instrument is in the form of a learning process assessment test using the Cronbach alpha technique. Reliability testing with the Cronbach alpha technique was carried out on instruments that are scored polytomies, so they are often used for tests in the form of descriptions or questionnaires (Susetyo, 2015). The following is the Cronbach alpha reliability formula.

$$\alpha = \left[\frac{k}{(k-1)}\right]\left[1 - \frac{\sum s_i^2}{s_i^2}\right] \tag{2}$$

α : Cronbach alpha reliability coefficient

k : number of items

$\sum s_i^2$: the total number of variants of the item

s_i^2 : variance of the total score

The reliability test on the AfL instrument was in the form of an assessment of learning outcomes, assessment of attitude aspects, and skills aspects using the Kuder Richardson 20 (KR-20) technique. The following is the formula for calculating the reliability of the items using the Kuder Richardson 20 (KR-20) formula (Susetyo, 2015).

$$r_i = \frac{k}{(k-1)}\left\{\frac{s_t^2 - \sum p_i q_i}{s_t^2}\right\} \tag{3}$$

r_i : instrument internal reliability

k : the number of items in the instrument

p_i : the proportion of the number of subjects who answered each item

q_i $= 1 - p_i$

s_t^2 : total variance

s_t^2 $= \frac{\sum(x-\bar{x})^2}{n}$, where x is the value of each question and n is the number of respondents

RESULTS AND DISCUSSION

1. Description of the Assessment for Learning (AfL) Instrument Development Process on congruence and similarity subject for Grade IX Junior High School Students

a.   Determining Instrument Specification

The first step in developing an assessment for learning (AfL) instrument is compiling instrument specifications. The preparation of instrument specification includes the activities of determining measurement objectives, preparing instrument grids, selecting instrument formats, determining instrument lengths, and determining instrument scales and scoring systems.

The AfL instrument aims to test and evaluate students' learning processes and outcomes in the cognitive aspects, skills, and attitudes. The results of this instrument play a role in providing information on the achievement of learning objectives and providing feedback for teachers and students to improve the quality of the learning process. The items presented in the instrument represent the characteristics of the congruence and similarity subject that is being taught by the teacher.

The preparation of the AfL instrument grid serves as a guide for the preparation of the instrument items. Following the grid is to ensure that all aspects of assessment and subject are proportionally covered. The AfL instrument grid on the cognitive aspect is prepared based on the topic of Congruence and Similarity in Mathematics class IX even semester of the 2013 curriculum, which includes topics: (1) congruence of plane figures, (2) congruence of triangles, (3) similarity of plane figures, and (4) similarity in triangles with process assessment using problem-solving criteria.

The AfL instrument grid on the skill aspect is prepared based on criteria that include aspects of (1) preparation, (2) process, (3) measurement, (4) pictures, (5) results, and (6) presentation. The AfL instrument grid on the attitude aspect is prepared based on attitude criteria, which include aspects of

(1) active discussion, (2) seriousness/concern, (3) cooperation, and (4) courtesy. The instrument grids are presented in matrix form.

The AfL instrument consists of test and non-test instruments. The instrument on the cognitive aspect is in the form of a learning process assessment test. The test consisted of essay question items accompanied by the assessment rubric. The learning outcome assessment test instrument is in the form of multiple choice. The AfL instrument on the skills and attitude aspects is described as an assessment rubric. The state of the AfL instrument was chosen based on an agreement between the researcher and the teacher.

b.   Writing Instruments

The writing of the instrument items is a step in translating the indicators into questions/criteria that are adjusted to the details on the grid that has been prepared. After compiling the instruments based on the indicators on the AfL instrument grid, the researcher then combined the instruments that had been compiled to form the AfL instrument on congruence and similarity subjects.
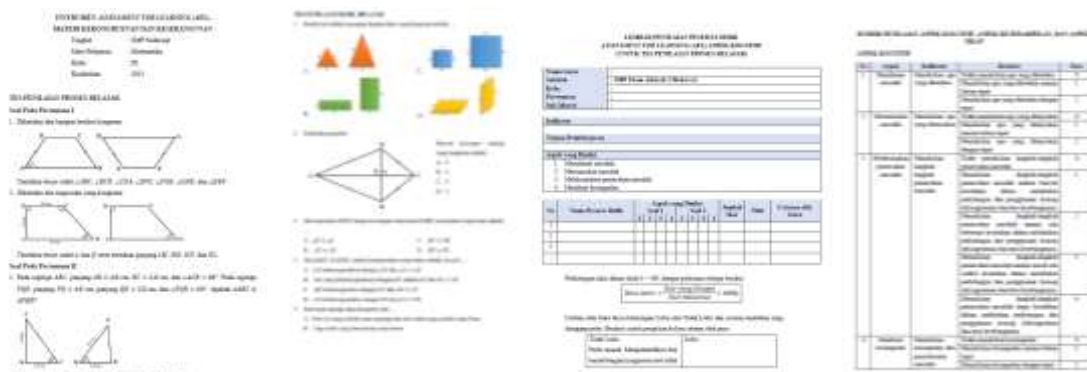

Figure 2. Design of AfL Instrument

c.   Reviewing Instruments

After the instrument was made, it underwent a review process. This review was carried out by testing the content validity. The test aimed at determining the suitability of the items/criteria with the indicators. The researcher used the content validity test of the Gregory validity testing model by taking into account the judgment of two experts/validators in testing the content validity of the instrument. After being validated by the experts, it was then analyzed using the Gregory validity test. After being validated, the instrument was then revised based on the notes from the two validators.

Nurul Hidayah[1], Ruslan[2], Awi Dassa[3]

Table 1. Results of the Content Validity Test of the Assessment for Learning (AfL) Instrument on congruence and similarity subject for Grade IX Junior High School Students

No	Instrument	Internal Consistency Coefficient Value
1	Learning process assessment test	1
2	Learning outcomes assessment test	0.95
3	Rubrics for assessing cognitive aspects, skills aspects, and attitude aspects.	1

Based on the results of the content validity test of the AfL instrument on congruence and similarity subjects, the internal consistency coefficient value was calculated. If the validity coefficient is greater than or equal to 0.75, it can be stated that the results of measuring content validity through the judgment of two experts are valid (Ruslan, 2009). Therefore, the AfL instrument that was developed is considered valid and feasible to be used for students.

d. Trialing Instruments

The instrument trial was conducted in class IX at SMP Islam Athirah 2 Makassar involving 126 students.

e. Interpreting Instrument Testing Results

After the instrument was tested, the measurement results were analyzed and interpreted to determine the extent to which students have achieved the learning objectives.

f. Analyzing Instruments

Furthermore, a quantitative analysis was carried out to determine the construct validity and reliability of the instrument that had been developed based on empirical data from the trial results. The instrument analysis was carried out using SPSS and Microsoft Excel. After the AfL instrument was tested empirically, construct validity testing could be performed.

The developed instrument was tested involving 126 class IX students of SMP Islam Athirah 2 Makassar. The construct validity of the instruments in this study was analyzed using factor analysis with the Confirmatory Factor Analysis (CFA) approach through the maximum likelihood (ML) analysis method using SPSS. The reliability test on the assessment for learning instrument was carried out using SPSS and Microsoft Excel employing the Cronbach alpha and Kuder-Richardson 20 (KR-20) techniques.

The empirical analysis of the AfL instrument on congruence and similarity subjects was conducted through a construct validity test employing

factor analysis with a Confirmatory Factor Analysis (CFA) approach. Based on the results of the construct validity test, the assessment of cognitive aspects in the learning process assessment test obtained 8 items that formed a factor, and all valid items were found in the 4 indicators developed. In the learning outcome assessment test, 17 items were obtained that formed a factor, and all valid items were found in the 7 indicators developed. In the skill aspect assessment rubric, 6 items of criteria were obtained that formed a factor, and valid items were found in the 6 indicators developed. In the attitude aspect assessment rubric, 4 criteria were obtained that formed a factor, and valid items were found in the 4 indicators developed. The results of processing the construct validity test revealed that 8 learning process assessment items, 17 learning outcomes assessment items, 6 skills aspect assessment rubric items, and 4 attitude aspect assessment items were declared valid with a factor loading value greater than or equal to 0.30. However, according to Agung (2016), an item does not need to be deleted or eliminated based on the results of validity tests conducted using sample data. So, in the learning outcomes assessment test, there were still 20 question items spread on the 7 indicators developed which were declared valid.

The reliability of the Afl instrument in the form of a learning process test instrument had a coefficient value of 0.74, the learning outcomes assessment test instrument had a coefficient value of 0.75, the non-test instrument in the form of a skill aspect assessment rubric had a coefficient value of 0.72, and the non-test instrument in the form of the attitude aspect assessment rubric had a coefficient value of 0.73. Thus, the developed AfL instrument was reliable. According to Azwar (2014), the concept of reliability is the extent to which the results of a measurement can be trusted. The standard prescribed in Mardapi (2008) stated that the magnitude of the accepted reliability index is at least 0.70.

g. Assembling Instruments

After all the instruments were analyzed and repaired, the final step was to assemble the instruments into a complete package of instruments. Products from the AfL instrument on congruence and similarity subjects are attached in Appendix A4.

2. Discussion

This study aims to produce a valid and reliable assessment for learning (AfL) instrument for grade IX students of junior high school. From the results

Nurul Hidayah[1], Ruslan[2], Awi Dassa[3]

of this study, the process of developing instruments and the quality of the assessment for learning (AfL) instruments for grade IX students of junior high school will be explained. The development products in this study were assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students in the form of learning process assessment tests, learning outcome assessment tests, and assessment rubrics on cognitive, psychomotor, and affective aspects. The development product is in the form of an AfL instrument that has passed two stages of assessment, namely, the results of content validation by experts/validators and trials on students. The content validation involved two mathematicians from Universitas Negeri Makassar and the implementation of the instrument trial involved 126 class IX students of SMP Islam Athirah 2 Makassar.

The procedures carried out in this development research were preparing instrument specifications (including determining instrument objectives, preparing instrument grids, instrument shapes and formats, determining instrument length, and determining instrument scales and scoring systems), writing instruments, analyzing instruments (content validity test and repairing the instrument), conducting instrument trials, interpreting the results of instrument trials, analyzing the instruments (construct validity and reliability tests), and assembling the instruments. The product of these procedures was the final package of the AfL instrument on congruence and similarity subjects that is suitable for use.

The AfL instrument on congruence and similarity subjects was developed through a grain study process with a content validity test. The results of the content validity of the AfL instrument revealed that, obtained an internal consistency coefficient of 1 out of 8 items on the learning process assessment test, 0.95 out of 20 items on the assessment of learning outcomes, and 1 of 34 rubrics for assessing cognitive aspects, skills, and attitudes have been arranged. This shows that the results of the assessment of the two validators in this validity test have strong relevance. In line with the standard stated in Ruslan (2009), the coefficient of content validity is high if it is greater than or equal to 0.75, and it can be stated that the results of the measurements or interventions carried out are valid.

After the instrument was stated to be theoretically valid, it was trialed to obtain empirical data. As for the empirical analysis to determine the construct validity and reliability of the instruments that have been developed. The empirical analysis of the AfL instrument for the topic of congruence and

congruence was conducted through a construct validity test employing factor analysis with Confirmatory Factor Analysis (CFA) approach. Based on the results of the construct validity test, the assessment of cognitive aspects in the learning process assessment test obtained 8 items that formed a factor, and all valid items were found in the 4 indicators developed. In the learning outcome assessment test, 17 items were obtained that formed a factor, and all valid items were found in the 7 indicators developed. In the skill aspect assessment rubric, 6 items of criteria were obtained that formed a factor, and valid items were found in the 6 indicators developed. In the attitude aspect assessment rubric, 4 criteria were obtained that formed a factor, and valid items were found in the 4 indicators developed. The results of processing the construct validity test revealed that 8 learning process assessment items, 17 learning outcomes assessment items, 6 skills aspect assessment rubric items, and 4 attitude aspect assessment items were declared valid with a factor loading value greater than or equal to 0.30. However, according to Agung (2016), an item does not need to be deleted or eliminated based on the results of validity tests conducted using sample data. So, in the learning outcomes assessment test, there were still 20 question items spread on the 7 indicators developed which were declared valid.

The reliability of the Afl instrument in the form of a learning process test instrument had a coefficient value of 0.74, the learning outcomes assessment test instrument had a coefficient value of 0.75, the non-test instrument in the form of a skill aspect assessment rubric had a coefficient value of 0.72, and the non-test instrument in the form of the attitude aspect assessment rubric had a coefficient value of 0.73. Thus, the developed AfL instrument was reliable. According to Azwar (2014), the concept of reliability is the extent to which the results of a measurement can be trusted. The standard prescribed in Mardapi (2008) stated that the magnitude of the accepted reliability index is at least 0.70.

Based on the results of the research that has been put forward, it shows that the assessment for learning (AfL) instrument on congruence and similarity subjects for grade IX students of junior high school that was developed was valid and reliable. Thus, it can be used in the assessment for learning mathematics on the topic of congruence and similarity for grade IX in junior high school.

Nurul Hidayah[1], Ruslan[2], Awi Dassa[3]

CONCLUSION

The process of developing assessment for learning (AfL) instruments on the topic of congruence and similarity for grade IX junior high school students was carried out through the stages of compiling instrument specifications (including determining the purpose of the instrument, compiling instrument grids, choosing the shape and format of the instrument, determining the length of the instrument, and determining instrument scales and scoring systems), writing instruments, reviewing instruments (content validity test and instruments improvement), conducting instrument trials, interpreting measurement results, analyzing instruments (construct validity and reliability tests), and assembling instruments.

The assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students consists of a learning process assessment test, a learning outcome assessment test, and a rubric for assessing cognitive, skill, and attitude aspects. The assessment for learning (AfL) instruments on congruence and similarity subjects for grade IX junior high school students satisfied the aspects of validity (content validity and construct validity) and reliability.

REFERENCES

Agung, I. G. N. (2016). Tantangan, keberhasilan, dan kebanggaan statistikawan. *Retirement Speech of Prof. Drs. I Gusti Ngurah Agung, M.Sc., M.St., Ph.D*. Makassar, Indonesia: Universitas Negeri Makassar.

Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Norwich: DfES Publications.

Azwar, S. (2014). *Pustaka pelajar*. Yogyakarta: Pustaka Pelajar.

Bulan, S. (2021). *Pengembangan asesmen berbasis proyek dalam pembelajaran matematika peserta didik kelas IX SMP*. Universitas Negeri Makassar.

Hamzah, A. (2014). *Evaluasi pembelajaran matematika*. Jakarta: Rajagrafindo Persada.

Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia Press.

Putri, I. K. A. (2022). *Asesmen dalam pembelajaran*. UIN Raden Intan Lampung.

Raharjo, S. B., Handayani, M., Jauhari, M. R. A., Juanita, F., & Waspodo, R. M. (2019). *Sistem penjaminan mutu pendidikan*. Jakarta: Pusat Penelitian Kebijakan Pendidikan dan Kebudayaan, Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan.

Ruslan. (2006). Aplikasi analisis faktor dalam uji-validitas instrumen penelitian. *Seminar Nasional Statistika IKAPSTAT Institut Teknologi Surabaya Dan Jurusan Matematika FMIPA UNM*. Makassar: FMIPA UNM.

Ruslan. (2009). Validitas isi. *Bulein Pa' Biretta*. Makassar: Media Informasi dan Komunikasi LPMP Sulawesi Selatan.

Sappaile, B. I. (2005). Validitas dan reliabilitas tes yang memuat butir dikotomi dan politomi. *Jurnal Ilmu Pendidikan (Parameter) Lembaga Penelitian UNJ*, *22*(24), 99–107.

Sappaile, B. I. (2007). Konsep instrumen penelitian pendidikan. *Jurnal Pendidikan Dan Kebudayaan*, *066*, 379–391.

Setiyawan, R. A., & Wijayanti, P. S. (2020). Analisis kualitas instrumen untuk mengukur kemampuan pemecahan masalah siswa selama pembelajaran daring di masa pandemi. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Sains*, *1*(2), 130–139. https://doi.org/10.46306/lb.v1i2.26.

Susetyo, B. (2015). *Prosedur penyusunan dan analisis tes*. Bandung: Rafika Aditama.

Thalaat, M. (2022). Pengaruh CTL dan asesmen autentik terhadap hasil belajar IPA siswa kelas IV SDN gugus selatan kec. Lenek tahun ajaran 2021/2022. *COLLAPSE (Creative of Learning Students Elementary Education)*, *5*(2), 315–325. https://doi.org/10.22460/collase.v5i2.10420.

Uno, H. B., & Satria, K. (2012). *Asesmen pembelajaran*. Jakarta: Bumi Akasara.

Yudha, R. P., Masrukan, M., & Djuniadi, D. (2014). Pengembangan instrumen asesmen otentik unjuk kerja materi bangun ruang di sekolah dasar. *Journal of Research and Educational Research Evaluation*, *3*(2), 62–67. Retrieved from https://journal.unnes.ac.id/sju/index.php/jere/article/view/4397.

Yusron, E., & Sudiyatno, S. (2021). How is the impact of assessment for learning (AfL) on mathematics learning in elementary schools? *Jurnal*

Prima Edukasia, 9(1), 75–84.

Yusuf, A. M. (2017). *Asesmen dan evaluasi pendidikan*. Jakarta: Prenada Media.

Yuwono, M. R., & Syaifuddin, M. W. (2017). Pengembangan problem based learning dengan assessment for learning berbantuan smartphone dalam pembelajaran matematika. *Beta Jurnal Tadris Matematika*, 10(2), 184–202.