



Procedural Knowledge Instruments for Grade IV Elementary School

Ni Luh Putu Widiastiningsih^{1*}, Kadek Yudiana², I Wayan Widiana²

^{1,2,3}Program Studi Pendidikan Guru Sekolah Dasar, Jurusan Pendidikan Dasar, Universitas Pendidikan Ganesha, Indonesia

ARTICLE INFO

Article history:

Received May 06, 2021

Revised May 12, 2021

Accepted August 24, 2021

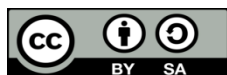
Available online November 25, 2021

Kata Kunci :

Procedural Knowledge,
Development, R&D Model

Keywords:

Pengetahuan Prosedural,
Pengembangan, Model R&D



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2021 by Author. Published
by Universitas Pendidikan Ganesha

ABSTRAK

Proses penilaian tidak jauh dari penggunaan instrumen sebagai alat bantu dalam pengukuran. Beberapa sekolah belum dapat menggunakan instrumen untuk mengukur kemampuan dimensi pengetahuan siswa, salah satunya adalah pengetahuan prosedural. Banyak faktor yang mempengaruhi seperti kurangnya motivasi guru dalam menilai pengetahuan prosedural siswa dan tidak adanya instrumen pengetahuan prosedural. Penelitian ini bertujuan untuk menghasilkan instrumen pengetahuan prosedural yang valid dan reliabel pada Tema 6 kelas IV Sekolah Dasar. Penelitian ini merupakan penelitian pengembangan yang dilakukan dalam sepuluh tahapan, menurut Brog & Gall, yaitu: Meneliti dan mengumpulkan informasi, membaca literatur, melakukan observasi, dan menyusun laporan kebutuhan pembangunan; Merencanakan komponen prototipe yang akan dikembangkan, mendefinisikan, merumuskan tujuan, menentukan urutan kegiatan; Mengembangkan produk awal; Melakukan uji tim ahli; Merevisi inisial berdasarkan hasil uji coba produk; Melakukan uji coba lapangan utama; Merevisi produk berdasarkan hasil uji coba lapangan; Melakukan uji coba lapangan operasional; Melakukan revisi akhir produk uji lapangan, dan Melakukan sosialisasi dan implementasi. Subjek dalam penelitian ini adalah 5 orang dosen ahli materi, 22 mahasiswa dalam kelompok kecil, dan 141 mahasiswa dalam kelompok besar. Dalam penelitian ini pengumpulan data dilakukan dengan wawancara, observasi, dokumentasi, angket, uji validitas, uji reliabilitas, uji diskriminatif, uji level. Kesulitan, dan analisis distraktor. Dari hasil penelitian ini didapatkan bahwa prototipe pengembangan instrumen pengetahuan prosedural dilakukan dalam empat tahap, instrumen pengetahuan prosedural dinyatakan valid dengan kriteria $pbi > r_{tabel} 5\%$, instrumen pengetahuan prosedural dinyatakan reliabel (0,934). Berdasarkan hasil penelitian dapat disimpulkan bahwa instrumen pengetahuan prosedural siswa kelas IV SD telah memenuhi kriteria valid dan reliabel.

ABSTRACT

The assessment process is not far from using the instrument as a tool in measurement. Some schools have not been able to use instruments to measure the ability of students' knowledge dimensions, one of which is procedural knowledge. Many factors influence such as the lack of teacher motivation in assessing students' procedural knowledge and there is no procedural knowledge instrument. This study aims to produce a valid and reliable procedural knowledge instrument on Theme 6 grade IV Elementary School. This research is development research conducted in ten stages, according to Brog & Gall, namely: Researching and collecting information, reading literature, making observations, and preparing reports on development needs; Planning prototype components to be developed, defining, formulating goals, determining the sequence of activities; Develop initial products; Conduct expert team test; Revise the initial based on the results of product trials; Conduct main field trials; Revise the product based on the results of field trials; Conduct operational field trials; Carry out final revision of field test products, and Conduct dissemination and implementation. The subjects in this study were 5 material expert lecturers, 22 students in small groups, and 141 students in large groups. In this study, data collection was carried out by interviews, observation, documentation, questionnaires, validity tests, reliability tests, discriminatory tests, level tests. The difficulty, and distractor analysis. From the results of this study, we found that the prototype of the procedural knowledge instrument development was carried out in four stages, the procedural knowledge instrument was declared valid with the criteria $pbi > r_{table} 5\%$, the procedural knowledge instrument was declared to have reliability (0.934). Based on the results of the study, it can be concluded that the procedural knowledge instrument for fourth-grade elementary school students has met the valid and reliable criteria.

1. INTRODUCTION

Learning is the interaction between students and their environment in which students make progress in achieving specific and purposeful knowledge, skills and attitudes (Hanafy, 2014; Pane & Dasopang, 2017). The learning process cannot be separated from the roping process. Assessment is the

*Corresponding author.

E-mail addresses: Widiastiningsih03@gmail.com (Ni Luh Putu Widiastiningsih)

process of obtaining information in any form that can be used as a basis for making decisions about students (Suryanto, 2012; Zahro, 2015). A good assessment uses the principles of assessment, which are comprehensive; objective; clear goals; appropriate, valid, and reliable instruments; fair; open; educate; accountable; and sustainable (Carrington et al., 2020; Salamah, 2018). One of the principles of assessment is the instrument. The instrument is a tool that meets academic requirements, so it can be used as a measuring tool for a measuring object or collecting data from a variable (Preston & Colman, 2000; Yusup, 2018). In line with this, instruments in education or learning are usually in the form of tests or non-tests. Test is an instrument or tool in measurement (Subarman, 2018). The test must be developed properly, to suit students' needs in procedural knowledge in a given material. This test is said to be feasible if it meets the test requirements, namely, valid, reliable, objective, and practical (Mujianto, 2017). In a test usually contains the dimensions of knowledge. The knowledge dimension consists of four types, namely: (1) factual knowledge; (2) conceptual knowledge; (3) procedural knowledge; and (4) metacognitive knowledge (Novferma, 2016; Nurfarida et al., 2017). In procedural knowledge, it is more about the process, 'how' something happens and how something can be solved by using the form/steps in solving it. Procedural knowledge can be defined as a process of scientific activity that requires methods, steps to solve a problem from a phenomenon that occurs to obtain meaningful conclusions at the However, in this study there are problems obtained based on interviews, observations and study documentation obtained are (1) Lack of use of questions regarding procedural knowledge by teachers, (2) Dimensions of students' procedural knowledge are still low, (3) Test instruments given by teachers still use tests the same as the previous year, (4) students were given more questions about factual and conceptual knowledge, (5) the implementation of procedural knowledge by teachers was still low. In line with the problems found are similar to the research namely the existence of LKS can help improve students' procedural knowledge (Junike, Yusrizal, 2016). In addition, research stated that students have not mastered the concepts related to the system of linear equations of two variables and students have not been able to answer the questions given with the right arguments and steps (Khamidah, 2017). Based on the results of interviews with fourth grade teachers, it is stated that the questions that have been created and given to students can only measure learning outcomes (cognitive) and assess students' understanding only rather than mastery of concepts and procedural knowledge.

Solutions that have been carried out to increase procedural knowledge and develop test instruments include research shows that measuring procedural knowledge using the type of essay test instrument (description) with an inquiry approach carried out by practicum activities using student worksheets (LKS) can increase students' procedural knowledge (Junike, Yusrizal, 2016). Meanwhile, in the research stated that students have not mastered the concepts related to the system of linear equations of two variables and students have not been able to answer the questions given with the right arguments and steps (Khamidah, 2017). Previous research shows that the developed mathematics learning outcomes test must have test validity, test instrument reliability, item difficulty index and good item discriminatory index and questions oriented to High Order Thinking Skill for fifth grade elementary school students (Ndiung & Jediut, 2020). Research shows that the science assessment instrument in the form of a multiple-choice test developed is feasible and has met the requirements to map critical thinking and practical skills of students with validity obtained in the range of 0.8 to 1.00 and the reliability of the science assessment instrument in the form of multiple choice tests has met conditions, including high with a reliability coefficient of more than 0.90 (Dewi & Prasetyo, 2016). Research shows that the concept understanding of students who receive learning through the application of problem-based learning (experimental class) is better than students who receive ordinary learning (control class) (Siregar et al., 2011).

Based on the problems described above, the development of procedural knowledge instruments is very much needed in the application of learning in elementary schools. Understanding, memorizing, and working on problems that are only on the basis of knowledge of facts and concepts are very lacking to solve a problem and make decisions that students will face later. These student competencies can develop if learning is directed to knowing and implementing step by step the activities carried out. The development of this instrument is devoted to measuring students' procedural knowledge, standing alone without the use of models, measuring students' cognitive domains from C2-C6, as well as the material contained in theme 6 (Cita-citaku) on science, social studies, Indonesian language, PPKn, and SBdP. It is hoped that with the procedural knowledge instrument it can measure the extent of students' procedural knowledge. The purpose of carrying out this research is to obtain valid and reliable procedural knowledge instruments.

2. METHODS

The type of research used is descriptive research with quantitative and qualitative data analysis techniques. Quantitative descriptive research is a type of research whose data is systematically arranged in

the form of numbers or percentages, and is related to the object under study, then qualitative descriptive research is research that presents data by describing all data obtained using sentences or words according to categories and existing context (Agung, 2014; Tohirin, 2012). The research was conducted using a 10-stage model (R&D), according to Brog & Gall stated that the 10 stages are, (1) Researching and gathering information, reading literature, making observations, and preparing reports on development needs; (2) Planning prototype components to be developed, defining, formulating goals, determining the sequence of activities; (3) Develop the initial product; (4) Conduct expert team test; (5) Initial revision based on the results of product trials; (6) Conducting main field trials; (7) Revise the product based on the results of field trials; (8) Conduct operational field trials; (9) Perform final revision of field test products; and (10) Dissemination and implementation (Yani, 2016). To clarify the design stages of the ten stages above, they can be combined into four stages, namely: (1) Needs analysis and formulating goals, (2) Designing procedural knowledge instruments, (3) Developing procedural knowledge instruments, and (4) Implementing and disseminating instruments procedural knowledge (Effendi & Hendriyani, 2018). The stage of needs analysis and formulating goals, in this first stage is done by collecting information based on literature studies and field studies to prepare requirements related to procedural knowledge instruments. The results of literature studies and field studies are used to formulate objectives. The stage of designing the procedural knowledge instrument, at the design stage, is carried out in two steps, namely the first step by determining the type of questions to be used in the instrument. The type of questions that will be developed is the type of multiple choice questions in simple form (consisting of four answer choices). The second step is to compile a grid of procedural knowledge instruments. Which is fully described in Table 1.

Table 1. Procedural Knowledge Instruments Grid

Subjects	Basic Competences	Indicators	Number of Items
Pancasila and civic education	Explain the benefits of the diversity of individual characteristics in everyday life	a. Determine how to maintain the diversity of individual characteristics	4
		b. Determine how to maintain the religious diversity in the vicinity	1
Indonesian	Exploring content and mandate of poetry presented orally and in writing with the aim of pleasure Verbalizing personal poetry with proper pronunciation, intonation, and expression as a form of self-expression	a. Shows how to make poetry	1
		b. Analyzing how to make poetry	1
Natural Sciences	Comparing the life cycles of several types of living things and linking them with conservation efforts upaya Make a life cycle scheme for several types of living things in the surrounding environment, and slogans for their conservation efforts	c. Examine the content and message of a poem	2
		a. Examine the relationship of poetry with proper pronunciation, intonation, and expression ekspresi	1
		a. Studying the cycle of living things	2
		b. Describe the life cycle of living things	1
Social Sciences/Social Studies	Identify the characteristics of space and the use of natural resources for the welfare of the community from the city/district level to the provincial level Presenting the results of the identification of spatial characteristics and the use of natural resources for the welfare of the community from the city/district level	a. Combining pictures of the stages of growth of animals and plants	1
		b. Deciding on pictures of the stages of growth of animals and plants	1
		a. Analyzing the relationship between the characteristics of space and the natural resources in its environment	1
		b. Analyze the distribution of animal or plant resources in the area	2
		a. Evaluating the results of observations on the use of natural resource sustainability	2

Subjects	Basic Competences	Indicators	Number of Items
Cultural Arts and Crafts	Demonstrating local dance movements	a. Analyzing how to make a regional creation dance	2
	Create collages, montages, apps and mosaics	a. Designing collages and montages with the right technique	2
		b. Rearrange the work of applications and mosaics with the correct technique	1

The stage of developing procedural knowledge instruments, this stage consists of three steps, namely (1) Arranging the items of the instrument referring to the question grid. (2) Validation through judgment. Judgment is carried out by five judges, namely five lecturers. Furthermore, a discussion was held between the researcher and the judges to obtain a decision on whether the items were revised or not. After that, a content analysis and revision I was carried out based on suggestions and input from judges and decision-making criteria as a result of the validation. (3) The field test is limited to the revised items of judgment results followed by a number of students on a small scale. In its implementation, it is done by distributing the items from the revised phase I. The data received from the implementation of the limited trial is then analyzed statistically. Statistical analysis carried out included the level of difficulty, discriminatory power, item validation, reliability of the question text, and analysis of distractors for each item. After that, a decision is made based on the results of statistical analysis and revision II.

The last stage is the stage of implementing and disseminating procedural knowledge instruments. At this stage, it is carried out in two steps, namely, (1) the broad stage test is carried out by giving back the procedural knowledge instrument that was revised in stage II and then submitted to a number of samples. This test is intended to produce a final product in the form of a set of procedural knowledge instruments that have been validated and are accurate. Furthermore, revision III was carried out as a revision of the final product based on the decision-making results of statistical analysis. (2) The final product revision (revision III) is intended to obtain a final product in the form of a valid and reliable procedural knowledge instrument. The subjects used in this study were 5 material expert lecturers, 22 students in small groups, and 14 students in large groups. The procedural knowledge test instrument is in the form of questions with a total of 30. The type of questions developed is multiple choice with four alternative answers. The questions compiled contain aspects of procedural knowledge. The teaching materials are contained in the learning process for class IV theme 6 (Cita-citaku). The correct answer will be given a score of 1 while the wrong answer or no answer will be given a score of 0.

Testing the validity of the instrument was carried out by giving the instrument to 5 experts, then the test results were analyzed using the CVR and CVI formulas. Based on the results of the CVR test, it was found that all the items of the instrument made were declared valid and suitable for use with the results of $\sum CVR$ of 30. After knowing the results of the CVR, the analysis proceeds to the calculation of CVI. The data obtained from the research results were then analyzed using quantitative descriptive analysis and qualitative descriptive analysis. Quantitative descriptive analysis was carried out by determining the validity of the items with biserial point correlation, the criteria for the instrument items were declared valid if the calculated pbi was greater than r_{table} ($pbi > r_{table}$) with a significance level of 5% (0.05), otherwise if the calculated pbi was smaller than r_{table} ($pbi < r_{table}$) then the item concerned is declared invalid (Candiasa, 2010). After obtaining the results of the validity of the items, then the difficulty level is calculated with the criteria used are the smaller the index obtained, the more difficult the question is. On the other hand, the higher the index obtained, the easier the problem will be. Clearer criteria can be seen in Table 2.

Table 2. Difficulty Level Criteria

Difficulty Level	Criteria
0,00 - 0,25	Hard
0,26 - 0,75	Medium
0,76 - 1,00	Easy

(Fatimah & Alfath, 2019)

Distinguishing Power calculations can be classified and interpreted in Table 3.

Table 3. Classification and Interpretation of Distinguishing Power

Criteria	Qualification	Interpretation
$D \geq 0,40$	Very good	The discriminatory power is very good, the items work very well
$0,30 < D \leq 0,39$	Good	Good discrimination, little or no need for revision
$0,20 < D \leq 0,29$	Medium	Medium discriminating power
$D \leq 0,19$	Bad	Weak discriminatory power, should not be used
Negative	Very bad	Bad discrimination, should be thrown away

(Yusuf, 2015)

Reliability using the KR-20 technique. This technique is used for dichotomous test types (objective tests). (Candiasa, 2010). The calculated reliability can be seen the results through criteria such as Table 4.

Table 4. Reliability Criteria

Criteria	Qualification
$0,00 < r \leq 0,20$	Very low
$0,20 < r \leq 0,40$	Low
$0,40 < r \leq 0,60$	Enough
$0,60 < r \leq 0,80$	High
$0,80 < r \leq 1,00$	Very high

(Candiasa, 2010)

Each item of multiple-choice questions is equipped with alternative answers. One alternative answer is the correct answer and the other is a distractor or incorrect answer. The multiple choice questions made consist of four alternative answers so that three of them serve as a distractor. Which is then continued by conducting a qualitative descriptive analysis, namely describing the calculated data and linking it with the results of previous research.

3. RESULT AND DISCUSSION

Results

The first stage is the stage of needs analysis and formulating goals. This instrument needs analysis activity is carried out by collecting information based on literature studies and field studies to prepare requirements related to procedural knowledge instruments. Based on the analysis stage that has been carried out, the results show that in the preparation of the instrument, the teacher still uses the previous questions, so there is a lack of innovation in making questions. In addition, teachers also rarely give questions or implement procedural knowledge in learning. Learning so far tends to only conceptual and factual knowledge abilities. Therefore, to improve students' procedural knowledge skills, an instrument of procedural knowledge is needed in elementary school. The second stage is the stage of designing the instrument. The design stage is done by determining the type of questions to be used, namely the type of multiple choice questions in simple form (consisting of four answer choices). And then compile a grid of procedural knowledge instruments. After the design is approved, the research can proceed to the stage of developing the instrument.

The third stage is the instrument development stage. At this stage, the preparation of instrument items refers to the question grid, which is followed by a validation test through judgment. The expert test results were analyzed using the CVR and CVI formulas. Based on the results of the CVR test, it was found that all the items of the instrument made were declared valid and suitable for use with a CVR result of 30. After knowing the results of the CVR, the analysis continued to the CVI calculation. From the CVI analysis that has been carried out, the CVI result is 1, so it can be stated that the procedural knowledge instrument of elementary school students has met the requirements very well. Subsequently, a limited trial was conducted. The limited trial of the procedural knowledge test instrument was given 25 items in the form of multiple choice conducted on 22 class V students. The data from the test results are then analyzed, validated, reliability, level of difficulty and discriminatory power.

A test can be said to be valid if the test can measure the object that should be measured. In this study, item validity was calculated using the biserial point correlation formula. The biserial point index (ypbi) obtained from the calculation results was consulted to the table at a significance level of 5%. By using a sample of 22 students, $r_{table} = 0.43$. If $pbi > r_{table}$, then all questions are said to be valid. Based on the

results of the analysis of the limited test items, all procedural knowledge test items in the form of multiple choice tests were declared to be used as research instruments. The reliability of the questions was measured using the KR-20 formula. The interpretation of the reliability coefficient (r_{11}) is if $r_{11} > 0.80$ then the items tested have very high reliability or reliable, the reliability of the questions obtained in the limited trial shows a very high scale, namely 0.934. The distinguishing power of an item if the criteria is $D \geq 0.40$ is declared very good, in this study it shows that all items have very good distinguishing power. Next, a distractor analysis was performed. The level of difficulty of a question is stated to be easy with the criteria of 0.76-1.00, while the medium level is between 0.26-0.75. In this study, it showed that there were 8 moderate questions and 17 questions with easy criteria. The quality analysis of the distractors of the questions is seen from the results of the distribution of the answers to the questions in each item. This research has a good distractor.

The fourth stage is the stage of implementing and disseminating procedural knowledge instruments. At this stage, a broad trial of the procedural knowledge test instrument was given 25 items in the form of multiple choice conducted on 141 class V students. The tests used have been validated by competent experts in their fields so that the research data obtained can achieve the expected goals. The data from the test results are then analyzed, validated, reliability, level of difficulty and discriminatory power. In this study, item validity was calculated using the biserial point correlation formula. The biserial point index (γ_{pbi}) obtained from the calculation results was consulted to the table at a significance level of 5%. By using a sample of 141 students, then $r_{table} = 0.17$. If $pbi > r_{table}$, then the question is said to be valid. Based on the results of the analysis of the broad test items, all procedural knowledge test items in the form of multiple-choice tests were declared to be used as research instruments. The reliability of the questions was measured using the KR-20 formula. The interpretation of the reliability coefficient (r_{11}) is if $r_{11} > 0.80$ then the items tested have very high reliability or reliable, the reliability of the questions obtained in the limited trial shows a very high scale, namely 0.934. The distinguishing power of an item if the criteria is $D \geq 0.40$ is declared very good, in this study it shows that all items have very good distinguishing power. Next, a distractor analysis was performed. The level of difficulty of a question is stated to be easy with the criteria of 0.76-1.00, while the medium level is between 0.26-0.75. In this study, it showed that there were 8 moderate questions and 17 questions with easy criteria. The quality analysis of the distractors of the questions is seen from the results of the distribution of the answers to the questions in each item. This research has a good distractor.

Discussions

The procedural knowledge instrument from the analysis is declared valid, because it is in accordance with the characteristics of students, in accordance with the cognitive domain of students. The instrument is said to be valid when it can reveal data from the variables correctly and does not deviate from the actual situation (Eivind & Ytterhaug, 2020; Mulholland, 2016). Validity is concerned with the extent to which the measurement is accurate in measuring what is intended to be measured (Anita et al., 2018; Hellstrand et al., 2020). In addition, the validity of a question can be seen by examining the curriculum so that it is able to measure students' mastery of learning materials, in line with the statement stated that the validity of the questions was seen with the curriculum and the data obtained were in accordance with the actual situation in the field (Ndiung & Jediut, 2020). The validity of the instrument can be proven by some evidence. These evidences include content, otherwise known as content validity or content validity, and constructively known as construct validity (Haviz, 2018). Content validity assessed by experts (Pradipta et al., 2020; Yusup, 2018). After conducting a content validity test to the expert, then the instrument was revised according to the advice/input from the expert. The instrument is declared content valid depending on the expert. Content validity can be determined by comparing the content contained in the learning outcomes test with Basic Competence (KD) in each of the existing subjects. Construct validity focuses on the extent to which a measuring instrument shows measurement results that match its definition. The definition of variables must be clear so that the assessment of construct validity is easy. The definition is derived from theory. If the definition is based on the right theory, and the question or item statement is appropriate, then the instrument is declared valid with construct validity.

Instruments of procedural knowledge from the results of the analysis are declared reliable. Instruments are said to be reliable when they can reveal reliable data (Schildkamp et al., 2020). This can be shown by the level of constancy (consistency) of the scores obtained by the subjects measured by the same measuring instrument at different conditions and at different times. The ability of students can affect the constancy of a question, in line with this study that states the reliability of the questions is influenced by the ability of students, the more heterogeneous or the more different students' abilities, the higher the reliability of the test (Dewi & Prasetyo, 2016). The number of students can affect the reliability of the questions, because the more test takers, the more diverse their abilities (Iskandar & Rizal, 2018). In addition, the length of the questions also affects the stability of a question, a large number of test items by

examining several objectives will be more reliable than a small number of items, because it will be more representative. However, the number of test items that are too many will be tiring and interfere with concentration so that the results obtained are no longer correct (Loka Son, 2019).

The level of difficulty in this study shows questions at an easy to moderate level. If a question has a proportionally balanced level of difficulty, then it can be said that the question is good (Nurjanah & Marlianingsih, 2015). Item items can be stated as good items, if the items are not too difficult and not too easy, in other words the degree of difficulty of the item is moderate or sufficient (Arisana & Ismani, 2016; Yulistianti & Megawati, 2019). Questions that are too easy do not stimulate students to solve problems (Rahmi Nur Fauziah et al., 2020). On the other hand, questions that are too difficult will cause students to become desperate and not have the enthusiasm to try again because they are out of reach (Iskandar & Rizal, 2018). The power of difference about the procedural knowledge studied stated that it was very good. A question can be said to have different power if it can be answered by high-ability students and cannot be answered by low-ability students (Fitriani & Artikel, 2017; Suzana, 2017). If a question can be answered by smart or poor students, it means that the question has no distinguishing power, so also if the question cannot be answered by smart students and poor students, it means that the question is not good because it does not have distinguishing power. The higher the discriminating power coefficient of an item, the more capable the item is to distinguish between students who master competence and students who do not master competence. (Hanifah, 2014).

The distractors contained in the procedural knowledge instrument are good. Good items, distractors will be chosen evenly by students who answer wrongly. On the other hand, items that are not good, the distractors will be chosen unevenly (Hery Susanto, Achi Rinaldi, 2015). There are usually three or five options, and of the possible answers attached to each item, one of them is the correct answer (answer key) while the rest are incorrect answers. That wrong answer is commonly known as "distractor" (distractor: pegecoh). The purpose of installing a distractor on each item is so that of the many students taking the test there are those who are interested in choosing it, because they think that the distractor they chose is the correct answer (Mujianto, 2017; Rusmawan, 2018). The more students are fooled, the more the distractors can carry out their functions as well as possible. On the other hand, if there is no selection of the distractor installed, then the distractor cannot perform its function properly. The distractor is declared to have functioned properly if the distractor has been selected at least 5% of all test participants (Arifin, 2017; Loka Son, 2019).

4. CONCLUSION

Based on some of the discussions above, it can be seen that this procedural knowledge instrument is valid and reliable. This indicates that the test text will always give the same results to the same group and are given the test at different times and situations. The limitation of this instrument is that it can only measure the dimensions of procedural knowledge.

5. REFERENCES

- Agung, A. A. G. (2014). *Metodologi Penelitian Pendidikan*. Aditya Media Publishing.
- Anita, A., Tyowati, S., & Zuldafrial, Z. (2018). Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Edukasi: Jurnal Pendidikan*, 16(1), 35–47. <https://doi.org/10.31571/edukasi.v16i1.780>.
- Anshori, M., & Iswati, S. (2009). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://dx.doi.org/10.18592/tarbiyah.v7i1.2100>.
- Arifin, Z. (2017). Kriteria Instrumen dalam Suatu Penelitian. *Jurnal Theorems (the Original Research of Mathematics)*, 2(1), 28–36. <http://dx.doi.org/10.31949/th.v2i1.571>.
- Arisana, A. L., & Ismani. (2016). Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 14(1), 78–97.
- Candiasa, I. M. (2010). *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Undiksha Press.
- Carrington, S., Saggars, B., Webster, A., Harper-Hill, K., & Nickerson, J. (2020). What Universal Design for Learning Principles, Guidelines, and Checkpoints are Evident in Educators' Descriptions of Their Practice When Supporting Students on the Autism Spectrum? *International Journal of Educational Research*, 102, 1–11. <https://doi.org/10.1016/j.ijer.2020.101583>.
- Dewi, N. D. L., & Prasetyo, Z. K. (2016). Pengembangan Instrumen Penilaian IPA untuk Memetakan Critical Thinking dan Practical Skill Peserta Didik SMP. *Jurnal Inovasi Pendidikan IPA*, 2(2), 213. <https://doi.org/10.21831/jipi.v2i2.11963>.
- Effendi, H., & Hendriyani, Y. (2018). *Pengembangan Model Blended Learning Interaktif dengan Prosedur Borg*

- and Gall. 62–70. <https://doi.org/10.31227/osf.io/zfajx>.
- Eivind, K., & Ytterhaug, B. O. (2020). Measuring Mathematical Identity in Lower Secondary School. *International Journal of Educational Research*, 103(7491), 1–12. <https://doi.org/10.1016/j.ijer.2020.101620>.
- Fitriani, V., & Artikel, I. (2017). Analisis Soal Ulangan Harian pada Materi Perubahan Lingkungan di Kelas X SMA di Kota Padang. *Jurnal Biologi Dan Pendidikan Biologi*, 3(1), 36–43. <https://doi.org/10.22202/bc.2017.v3i1.2724>.
- Hanafy, M. S. (2014). Konsep Belajar dan Pembelajaran. *Lentera Pendidikan : Jurnal Ilmu Tarbiyah Dan Keguruan*, 17(1), 66–79. <https://doi.org/10.24252/lp.2014v17n1a5>.
- Hanifah, N. (2014). Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. *SOSIO E-KONS*, 6(1), 41–55. <https://dx.doi.org/10.30998/sosioekons.v6i1.1715>.
- Haviz, M. (2018). Designing and Developing Teaching Material of Human Anatomy With Thinking Map: What Is Internal Relevance and Consistency? *Edusains*, 10(1), 38–45. <https://doi.org/10.15408/es.v10i1.7206>.
- Hellstrand, H., Korhonen, J., Räsänen, P., Linnanmäki, K., & Aunio, P. (2020). Reliability and Validity Evidence of the Early Numeracy Test for Identifying Children at Risk for Mathematical Learning Difficulties. *International Journal of Educational Research*, 102, 1–10. <https://doi.org/10.1016/j.ijer.2020.101580>.
- Hery Susanto, Achi Rinaldi, N. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *Jurnal Pendidikan Matematika*, 6(2), 203–217. https://doi.org/10.18907/jjsre.37.3_343_4.
- Iskandar, A., & Rizal, M. (2018). Analisis Kualitas Soal di Perguruan Tinggi Berbasis Aplikasi TAP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(1), 12–23. <https://doi.org/10.21831/pep.v22i1.15609>.
- Junike, Yusrizal, H. (2016). Pengembangan dan Implementasi Instrumen Tes untuk Mengukur Pengetahuan Prosedural dengan Menggunakan Pendekatan Inkuiri di SMA Negeri 10 Banda Aceh. *Jurnal Pendidikan Sains Indonesia*, 8(2), 45–55. <http://e-repository.unsyiah.ac.id/JPSI/article/view/6584>.
- Khamidah, L. (2017). Pemahaman Konseptual dan Pengetahuan Prosedural Siswa Kelas VIII dalam Penyelesaian Soal Matematika pada Materi Sistem Persamaan Linier Dua Variabel. *Prosiding SI MaNis (Seminar Nasional Integrasi Matematika Dan Nilai Islami*, 1(1), 611–616. <http://conferences.uin-malang.ac.id/index.php/SIMANIS/article/view/297>.
- Laela Umi Fatimah, K. A. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64. <https://doi.org/10.36668/jal.v8i2.115>.
- Loka Son, A. (2019). Instrumentasi Kemampuan Pemecahan Masalah Matematis: Analisis Reliabilitas, Validitas, Tingkat Kesukaran dan Daya Beda Butir Soal. *Gema Wiralodra*, 10(1), 41–52. <https://doi.org/10.31943/gemawiralodra.v10i1.8>.
- Mujiyanto, S. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Butir Tes, Validitas Butir Tes, Interpretasi Hasil Tes Validitas Ramalan dalam Evaluasi Pendidikan. *Jurnal Manajemen Dan Pendidikan Islam* 2, 2(2), 192–213. <https://dx.doi.org/10.26594/dirasat.v2i2.879>.
- Mulholland, S. (2016). ADHD-Specific Knowledge and Attitudes of Teachers (ASKAT): Development and Validation of A New Research Instrument. *International Journal of Educational Research*, 77, 109–116. <https://doi.org/10.1016/j.ijer.2016.03.010>.
- Ndiung, S., & Jediut, M. (2020). Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar Berorientasi pada Berpikir Tingkat Tinggi. *Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), 94. <https://doi.org/10.25273/pe.v10i1.6274>.
- Novferma, N. (2016). Analisis Kesulitan dan Self-Efficacy Siswa SMP dalam Pemecahan Masalah Matematika Berbentuk Soal Cerita. *Jurnal Riset Pendidikan Matematika*, 3(1), 76. <https://doi.org/10.21831/jrpm.v3i1.10403>.
- Nurfarida, L., Yusup, P. M., & Komariah, N. (2017). Tingkat Pengetahuan Masyarakat Sekitar Observatorium Bosscha Lembang Mengenai Polusi Cahaya. *Jurnal Kajian Informasi Dan Perpustakaan*, 5(1), 13. <https://doi.org/10.24198/jkip.v5i1.11327>.
- Nurjanah, & Marliansih, N. (2015). Analisis Butir Soal Pilihan Ganda dari Aspek Kebahasaan. *Faktor Jurnal Ilmu Kependidikan*, 2(1), 69–78. <https://dx.doi.org/10.30998/fjik.v2i1.377.g359>.
- Pane, A., & Darwis Dasopang, M. (2017). Belajar dan Pembelajaran. *FITRAH: Jurnal Kajian Ilmu-Ilmu Keislaman*, 3(2), 333–352. <https://doi.org/10.24952/fitrah.v3i2.945>.
- Pradipta, Sariyasa, & Lasmawan. (2020). Pengembangan Instrumen Kemampuan Berpikir Kreatif dan Literasi Matematika pada Materi Geometri Peserta Didik Kelas IV Sekolah Dasar. 10(1), 21–30. <https://doi.org/10.23887/jpepi.v10i1.3504>.

- Preston, C. C., & Colman, A. M. (2000). Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, and Respondent Preferences. *Acta Psychologica*, 104(1), 1–15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5).
- Rahmi Nur Fauziah, I., Syihabudin, S., & Sopian, A. (2020). Analisis Kualitas Tes Bahasa Arab Berbasis Higher Order Thinking Skill (Hots). (*LISANUNA*): *Jurnal Ilmu Bahasa Arab Dan Pembelajarannya*, 10(1), 45–54. <https://doi.org/10.22373/ls.v10i1.7805>.
- Rusmawan, P. N. (2018). Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Bahasa Inggris Kelas VII SMPN 2 Tegalsiwalan. *Linguista: Jurnal Ilmiah Bahasa, Sastra, Dan Pembelajarannya*, 2(1), 39. <https://doi.org/10.25273/linguista.v2i1.2686>.
- Salamah, U. (2018). *Penjaminan Mutu Penilaian Pendidikan*. 2(1), 274–293. <https://dx.doi.org/10.32478/evaluasi.v2i1.79>.
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative Assessment: A Systematic Review of Critical Teacher Prerequisites for Classroom Practice. *International Journal of Educational Research*, 103, 1–16. <https://doi.org/10.1016/j.ijer.2020.101602>.
- Siregar, N., Armanto, D., & Saragih, S. (2011). Penerapan Pembelajaran Berbasis Masalah untuk Meningkatkan Pemahaman Konsep dan Pengetahuan Prosedural Matematika Siswa SMP. *Jurnal Pendidikan Matematika*, 5(2), 137–150. <http://digilib.unimed.ac.id/id/eprint/754>.
- Suharman. (2018). Tes sebagai Alat Ukur Prestasi Akademik. *Jurnal Ilmiah Pendidikan Agama Islam*, 10(1), 93–115. <http://ejournal.staindirundeng.ac.id/index.php/tadib/article/view/138>.
- Suryanto, A. (2012). Konsep Dasar Penilaian dalam Pembelajaran. *Evaluasi Pembelajaran Di SD*, 5(1), 63–111. <https://pustaka.ut.ac.id>.
- Suzana, A. (2017). Analisis Tingkat Kesukaran dan Daya Beda Butir-Butir Soal Penilaian Akhir Tahun Matematika Kelas X di SMA Negeri 1 Purbalingga. *MathGram Matematika*, 2(2), 1–8. <https://ejournal.unugha.ac.id/index.php/mthg/article/view/172>.
- Tohirin. (2012). *Metode Penelitian Kualitatif dalam Pendidikan dan Bimbingan Konseling*. PT Raja Grafindo Persada.
- Yani, R. (2016). Pengembangan Instrumen dan Bahan Ajar untuk Meningkatkan Kemampuan Komunikasi, Penalaran, dan Koneksi Matematis dalam Konsep Integral. *Jurnal Penelitian Pendidikan*, 13(1), 44–52. <http://jurnal.upi.edu/penelitian-pendidikan/view/1390>.
- Yulistianti, H. D., & Megawati, E. (2019). *Analisis Instrumen Tes Higher Order Thinking Matematika SD*. 13(1), 41–54. <https://pdfs.semanticscholar.org/96cf/d86879287947093721888b83bee0d39950d2.pdf>.
- Yusuf, M. (2015). *Asesmen dan Evaluasi Pendidikan Pilar Penyedia Informasi dan Kegiatan Pengendalian Mutu Pendidikan*. Prenadamedia Group.
- Zahro, I. F. (2015). Penilaian dalam Pembelajaran Anak Usia Dini. *Tunas Siliwangi*, 1(1), 92–111. <https://doi.org/10.22460/ts.v1i1p92-111.95>.