JERE 7 (1) (2018) 87 – 93

Journal of Educational Research and Evaluation

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

The Development of Theme Instrument Based Test for The IV Grade of Elementary School in Margorejo District Pati

Ana Sholihatin Septiningsih [1]✉, Totok Sumaryanto Flotentinus[2], Dewi Liesnoor Setyowati[2]

[1.]SDN Sokokulon 02, Pati, Indonesia
[2.] Universitas Negeri Semarang, Indonesia

Article Info

Article History:
Accepted 17 Mei 2018
Approved 12 July 2018
Published 10 March 2018

Keywords: Instrument development, valid, reliable,

Abstract

This study motivated the lack of teacher's capability to apply the 2013 curriculum , so some teachers are less able to develop the independent test instrument. They depens on the LKS. Beside that, teachers are formulating question for examination without considering the indicator or analyzing basic competency that will be achieved students. This research is to develop valid standard instrument, reliable and effective use. This research is a development instrument with R & D modify the Ball & Gall approach. Test validation of the contents of use v'aiken by rating experts, to determine the validation and reliability of the empiric ie with analysis correlation point biserial and kuder richadson (KR-20). The results of validation of the contents of the instrument test-based theme by experts worth use. While the trial court of about developed a total of 30 items about a valid as much as 26 item about. So with reliability in the trial court range 0,89-0,91 with high category. While the test of effecticity, based on the analysis of the questionnaire with 6 of respondents primary school teacher and 28 items questionnaire, get a score of 497 with an average 82,83 by effective category. Based on the results of these studies, the theme-based test instruments developed valid, reliable and effective to use. This research is useful for teachers as a reference and learning materials in measuring the daily assessment of student learning outcomes.

© 2018Universitas Negeri Semarang

✉Correspondence Address:
Jalan raya Pati Kudus Km.04,Pati, Indonesia
E-mail: anasholihatin21@gmil.com

P-ISSN 2252-6544
E-ISSN 2502-356X

INTRODUCTION

Integrated thematic learning is a learning that integrates the various competencies from different subjects into the various themes (Curriculum Document 2013). The theme knows the meanings of basic concepts so that students do not learn basic concepts partially.

In the academic year 2014/2015 the enactment of the 2013 curriculum in all over Indonesia which is the renewal and improvement of the 2006 Curriculum. The basic characteristic of the Curriculum 2013 lies in the approach used in developing the curriculum. The 2013 curriculum emphasizes a scientific approach to the concept of thematic learning at the primary to secondary level. Weeb & Pearson (2012: 19) states down the initial activity in thematic learning that is analyzing the theme to be chosen .

One of the considerations in the development of the 2013 curriculum is in terms of assessment. The curriculum, learning, and assessment are the components of learning. Based on Kemendikbud Year 2016 on the Assessment Guide for Primary School one of which is the assessment of daily assessment (PH), in which activities undertaken by education to measure the achievement of basic competencies. Assessment itself is a series of activities to obtain, analyze and interpret data about the process and learning outcomes of learners are done in a systematic and continuous so that it can be meaningful information in decision making (Daryanto, 2014: 111).

All learning content in the classroom can be integrated in thematic learning. *Implementation od school currilum 2013 that had implemented by developing integrated thematic learning model scientific based using discovery learning technique* (Mudiono, Gipayani, Madyono: 19-27).

Based on interviews with several elementary school teachers in Pati Regency. Teachers are less self-sufficient in developing daily judging apps in the form of theme based test instruments, teachers are encouraged in LKS books without formulating indicators or reviewing the basic competencies to be achieved by students.

. Daily assessment is one form of learning outcomes that must be achieved students, if the instrument is not a good question how can be said as a standard test instrument. The quality of the test and item are largely determined by validity, reliability, objectivity, practiceability, distinguishing power, degree of difficulty, choice effectiveness and efficiency (Solichin, 2017: 192-213). The development of problem banks in Puspendik divided into several stages namely: (1) preparation of the lattice, (2) writing questions; (3) review and revision (review and improvement); (4) assembly questions; (5) test questions; (6) quantitative analysis; (7) selection of questions) (Balitbang Kemendikbud: 2017).

The development of theme-based test kits is essential for immediate development as the theme-based test tool is the most important assessment part of classroom learning. development of test instruments to find out that the standard developed instruments are in accordance with the developed grid (Whardhani & Putra, 2016: 75-82). The test instrument is a tool used to perform tests to find out the achievement of learning or competence achieved by learners for a particular field. As Mardapi (2016: 97) points out, the test results are characteristic information for a group of people.

With the theme-based test is a test tool used in thematic learning, the purpose of the theme-based test instrument as well as the general results test results to measure the progress of students. The test of learning outcomes is one type of test used to measure the progress or learning progress of learners, after they follow the learning process (Sudijono, 2009:99). *Effective in meeting the needs of all students while implementing new thematic* (J.John, 2015: 172- 187).

Based on the description, the formulation of research problem are: (1) how feasibility of theme based test instrument according to some experts ?; (2) how to design a valid and reliable theme-based test development design; (3) how is the effectiveness of theme based test instrument in class IV Primary School?.

In line with the formulation of the problem, the researcher aims to arrange a theme-based test instrument that is valid content, valid grain, reliable and effective to use. Theme-based test instrument is developed for teachers and students so that it complies with the principles of the appraisal standards of Permendiknas No.9 year 2016. The development of test instruments should be through various test instruments, namely qualitative test and quantitative test (Fitrifianofa, Waskito, Budiharti: 2013: 90). Because the assessment instrument is needed teacher for evaluation tool.

Theme-based test instrument developed based on the 2013 curriculum for Grade IV Primary School with a theme "Cita-Citaku" sub tema "Hebatnya Citaku-citaku".

METHODS

This research was conducted in class IV in Margorejo Sub-district, Pati Regency, which uses the 2013 curriculum of 6 schools in Moch Syafi'I cluster. The subjects of this research are (1) students as subject for simulation and product testing, (2) experts as much as 3 people for validation of test instrument made, (3) subject teacher to know the effectiveness of theme based test instrument. Expert Team Validator consisting of 3 experts, namely 2 Lecturers UNNES evaluation experts, and 1 teacher expert test legitimacy test question

This research is a development research. The research model used as a reference is the development stage of the R & D Model by Borg & Gall (2003), by modifying it into nine steps consisting of; potential problems, information gathering and literature studies, product design, design validation, design revisions, limited trials, product revisions, field trials, and final product revisions.

The theme-based test instrument developed in this study is the daily replication of the short stuff in accordance with the theme "Cita-Citaku" Sub Theme "Great Cita-Citaku" by reviewing the core competence (KI) and Basic Competence (KD) contained in the theme, the , determine the grid, the preparation of theme-based test instruments based on the lattice.

Product development trial is done through 3 stages, namely expert test, individual test, field trial. Expert tests are conducted to determine the validity of the contents of the instrument so that the theme-based test instrument can be used on field tests. Field tests included small-scale test subjects of 30 students as a minimum sample on statistical tests (Baley, 2011) in 6 schools, and a broad-scale test conducted with the subject of 126 students in 6 school

Data analysis technique in this research is to test the feasibility of the instrument / content by Expert using questionnaire of content validity validity, usage test that is test the validity of grain, reliability, difficulty level and differentiating power problem. Testing the effectiveness of theme-based test instruments using a questionnaire of effectiveness assessment of 6 teachers.

RESULT AND DISCUSSION

The results show and discuss three important points namely: 1) the feasibility of test-based instruments according to some experts; 2) designing valid and reliable theme-based test instrument design; 3) the effectiveness of theme-based test subjects developed.

The Feasibility of Theme-Based Test Instrument According to Some Experts

The product developed in this research is the test instrument that is theme based test instrument there is thematic learning for the fourth grade of elementary school. The preparation of the instrument grid refers to the guidelines for the assessment of learning achievement kembbud, as well as from the curriculum content of 2013. After the grid is made the next step is the preparation of the question items as much as 30 questions in the form of a short entry with the theme "Cita-citaku" sub theme "great my ideals ". The initial design is shown in table 1.

Table 1. Theme Instrument Test

No.	Subject	KD	N of Items
1	PPKn	3.3	6
2	Bahasa Indonesia	3.3	6
3	Matematika	3.11	6
4	IPA	3.5	6
5	IPS	3.5	6
N of Items			30

The initial design of the instrument is tested content validation / expert. Previously on the theme-based test instrument needs to be fixed in some questions in accordance with the input and suggestion of the three validators that is about the improvement of the system punctuated sentence, the indicator is not in accordance with the item. Here are the results of the analysis of expert judgments of theme-based test instruments.

Table 2. Experiment Validation Analysis Results

Butir Soal Nomor	Koefisien Aiken's V	Keterangan
1-30	0.56-1.00	Proper to use

Based on the results of expert validation analysis using Aiken's V formula indicated that 30 items of theme-based test instrument for the fourth grade of elementary school were said to be worth mentioning, because this criterion is based on koefiesien (v) ≥0,300.

Designing Valid and Reliable Theme Based Test Instrument Design

Developing the theme-based test instrument for elementary school class IV is done with several stages in accordance with the

guidance of the writing of the Ministry of National Education of 2018, among which are: (1) determine the purpose of assessment; (2) pay attention to SK & KD; (3) determine the type of measuring instrument; (4) arranging the test grille; (5) writing theme-based test items.

After the preparation of the appropriate instrument, carried out to the next trial that is field trials. Trial for 3 times, namely trial 1, trial 2, trial 3. The test results obtained information time needed to complete the theme-based test questions approximately 70 minutes. In addition through the interpretation of the analysis of items can be known the quality of the problem based on the characteristics of the item with known grain quality questions that include; content validity, difficulty level, distinguishing power, and reliability.

To test the validity of short item, using biserial correlation pount because the data obtained dichotomy, the result obtained as follows:

Tabel 3. The Validity of Theme-Based Test Instrument

Uji Coba Lapangan	*Corelation Point Biserial* (t_{tabel})	Jumlah Butir Valid
Trails 1	2.05	23
Trails 2	1.98	27
Trails 3	1.98	26

Based on the analysis of the validity of the above items of 30 items, the experimental 1 instrument was tested to 30 students (small scale) to know the validity of the grain and the reliability of the instrument. Valid items if $t_{hitung} \geq t_{tabel}$ (30 subjects (Torang, 2014)). Tests 1 obtained 23 items declared valid ($t_{hitung} \geq 2{,}05$) and can be continued in the next test that is test 2 and 3. Then some items need to be revised in accordance with the power of different items. So in test 2 and test 3 produce all valid items.

In the test it can be determined that the tested problem is good, it looks at the reliability coefficient value, the average level of difficulty and the average of the different power of all questions. However, based on the analysis of each item there are items that need to be revised content (Table 4)

Table 4. Differentiation Power Problem

catagory	Items	Number of item
accepted	1,3,4,5,7,10,13,14, 16,17,18,21,23,28, 29,30	16
Accepted repaired	2,11,20,24,26,27	6
repaired	8,9,15,19,22	5
thrown away	6,12,26	3

The results of the test Table 5 is used as an ingredient for the revision of the contents of the problem, this improved problem is a matter that will be used in the next test of large-scale trials (trial 2 and trial 3). To know the category of questions can be seen based on the level of difficulty of the problem. In accordance with the conclusion of Thorndike & Hagen (1961) that good tingka difficulty is 0.3 to 0.7 or moderate category. Here is the analytical hasl for the problem level of Table 5:

Table 5. Trouble Level Limited Trial Problem

Classification	Items	amaount
easy (0,7<p 1,00)	1,4,5,14,16,21, 22,23,24	9
medium (0,3 p 0,7)	2,3,6,8,9,10,11, 13,15,17,18, 19,28,29,30	15
difficult (P < 0,3)	7,12,20,25,26,2 7	6

Test reliability of the instrument using reliability test Composite reliability by method (KR-20) as retnawati opinion (2016: 90-91) Kuder Richardson formula used to estimate the reliability of dichotomous instrument or 1 and 0. Test reliability is done against the revised item from the differentiator power test on a limited trial of 30 point-based test instruments. based on Kuder Richardson's reliability measurements using SPPS software with the results in the following table.

Tabel 6. Instrument reliability on a small scale

Instrument	KR-20	N of Items
Theme-Based Test	0.910	27

The reliability coefficient of the theme-based test instrument obtained 0.91 can be interpreted as very high reliability, as Guilford (1956: 145) categorizes the reliability coefficient of 0.8000 - 1,000 very high categories.

Instruments can be used or tested further on a broad class sample with a subject of 126 students. This trial to know the consequences of developed instruments. As Gronlund (1981: 169), other terms of the instrument are good "*reliability refers to the consistency of measurement, that is to how consistent test scores of other evaluation result are from one measurement to other*".

Large / Large Scale Test

Large trials were conducted twice, trial 2 and trial 3 ie after trial 1 (limited scale test) was done on a large / large scale with a sample of 126 fourth grade students at SDN in Pati . As in a limited-scale trial, extensive trials through the grain validity test stage, discrimination test, difficulty test and instrument reliability test.

Based on the results of small-scale trials the number of items consists of 30 questions with feasible question used in the next trial that is as much as 27 items. From 27 items tested to 126 students on a large scale 2, then on trial 3 as many as 26 items with the results of all validi. So it can be continued in the next stage that is on the differentiator power test. here are the table 7 test results on the differentiating power of test 2 and trial 3

Table 7. test results of 2 and 3 differentiating power questions

Katagori	Uji Coba 2		Uji Coba 3	
	Items	N of Items	Items	N Of Items
accepted	1,2,3,5,6,8,9,10,11,12,13,14,15,16,17,18,19,21,22,23,25,26,27	23	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26	26
Accepted repaired	20,24	2	-	-
repaired	7	1	-	-
thrown away	4	1	-	-

Different power analysis of the problem aims to determine whether or not a problem differentiates between the upper and lower groups. As according to Arikunto (2012: 232) the differentiator of the ability to distinguish high-ability students with low-ability students with a good discrimination index of 0.41-0.70 and a very good 0.71-1.00. As Wardhani & Putra (2016: 75-82) studies to develop standard instruments need to be tested for differentiating power.

Next test the level of difficulty question. The degree of difficulty is the opportunity to correctly answer a problem at the level of ability of the provisions usually expressed in the form of index. can be seen in Table

Table 8. Trouble Level 2 Testing and Trial 3

catagory	*Trial 2*		*Trial 3*	
	Items	*N*	*Items*	*N*
Easy (0,7<p 1,00)	5,7,9,10,13,17,25	7	4,6,8,9,12,16,24	7
Medium (0,3 p 0,7)	1,2,3,4,6,8,11,12,14,15,18,19,20,21,22,23,24,26,27	20	1,2,3,5,7,10,11,13,14,15,17,18,19,20,21,22,23,25,26	19
difficult (P < 0,3)	-	-	-	-

Based on the analysis of table 8, most of the theme-based test instruments developed in the medium classification, the theme-based test instruments developed include feasible use and either meet the criteria of good instruments, as Sudijono (2009: 370) points of test item learning can be expressed as good item items, when item items are not too difficult and not too easy in other words the difficulty of the item is moderate or sufficient. Lewis R Aiken (1994: 66) states classically this index of difficulty levels expressed in the form of a proportion of magnitude ranging from 0.00-1.00.

As according to Khuerudin (2017: 97-129) Instruments developed must meet the analysis of quantitative and quantitative items. Quantative analysis, which includes the level of difficulty / difficulty (TK) items, distinguishing power, power pengoh, and reliability in accordance with the principles of assessment.
The next stage is to test the reliability of theme-based test instruments on large-scale trials 2 and 3. Here's a table serving 9.

Tabel 9. Analysis of Trial 2 and Trial 3 Reliability

Instrument	KR-20	N of Items
Test-Based Theme Test 2	0,89	27
Test-Based Theme Test 2	0,91	26

The results of the analysis were done on the item of questionnaire used in test 2 and test 3. The reliability coefficient indicated that the test items were arranged based on the theme-based test on high clasfication. This is in accordance with the opinion reinforced by the opinion of Benninga, et al (2003), the instrument

exceeds 0.8 so that the reliability in the category good and can be analyzed further.

The Effectiveness of Theme-Based Testing Instrument

Instrument effectiveness analysis is done by spreading the questionnaire of the effectiveness of theme based test instrument to 6 teachers. The effectiveness of the instrument is performed to see the quality and effectiveness of the developed instrumental so that the product is beneficial to teachers and students. This is because the testing of the test instrument is based on needs analysis into new products with its effectiveness test first, thereby producing products that function and benefit the community (Wijiyanti, et al 2015).

The effectiveness test obtained from the teacher's response to the questionnaire-based test instrument effectiveness instrument for Elementary School has a score of 497 with an average score of 82.83 such results are consulted with the following table.

Table 10. Criteria for Effectiveness of Theme-Based Testing Instrument

Rentang Skor	Katagori
1-28	Very ineffective
29-56	Ineffective
57-84	Effective
85-112	Very effective

From the table, it is known that the instrumers are included in the "effective" category, this result is in accordance with H.Emerson (1994: 16) explanation which states that effectiveness is a measurement in the sense of achieving predetermined goals. Even to the effectiveness of the Ratopo test, Waskito & Wibowo (2014) compared the effectiveness of summative physics test instruments using self and feeback revision methods to produce standard instruments.

From the test results of the effectiveness of theme-based test instruments it is known that the theme-based test instrument for fourth grade students of SD is effectively used by teachers, students, and schools, both in whole and in every aspect.

CONCLUSION

The conclusions of research and development of theme based test instrument for fourth grade students in Elementary School are: (1) Theme-based test instrument consists of 26 items of standard that are feasible to be used; (2) Theme base test instrument is proven by quantitative analysis that has grain validity, problem-breaking power, problem level and good reliability. The proof is done by relibilyas coefficient analysis through KR-20 method has also exceeded the high category limit that is 0,80; (3) the effectiveness of theme-based test instruments in fourth grade thematic learning elementary school with the theme of "Cita-Citaku" sub theme "Terrific Cita-Citaku" entered into effective category. This is evident from the results of questionnaire effectiveness (teacher) a number of 6 people, and analyzed shows a score of 82.83 mengkatagorikan effective instrument.

ACKNOWLEDGMENT

Acknowledgment of the researcher to the parties who have helped during the process of this research, including: Department of Education and Culture of Pati Korwil District of Margorejo and also to the supervisor for guidance, criticism and suggestion. Thanks also to the expert validator of research and education evaluation

REFERENCES

Arikunto. (2012). *Dasar-Dasar Evaluasi Pendidika*n. Jakarta: Bumi Aksara.

Aiken, Lewis R. (1994). *Psychologial Testing and Asessment*, Eight Edition. Boston: Allyn and Bacon

Balitbang Kemdikbud. (2017). *Panduan Penulisan Soal 2017 SD/MI*. Pusat Penilaian Pendidikan Badan Penelitian dan Pengembanagan Kementrian Pendididikan dan Kebudayaan.

Benninga, S., Jacques., Berkowitz, W., Marvin, K Phyllis, S., & Karen. (2013). The Relationship Of Character Education Implementation And Aademic Achievement In Elementary Schools. *International Journal of Research in Character Education, 1(1), pp. 19-31. ISSN 1543-1223.*

Borg, Walter R & Meredith Damien Gall. (1979). *Educational Research.* Third Edition. United States Of Ameriks: Longman Inc.

Borg, Walter R., and Gall, Meredith Darmein. (1983). *Education Research: An Introduction, 4th ed*. New York: Longman Inc.

Daryanto. (2014). *Pendekatan Pembelajaran Saintifik Kurikulum 2013*. Yogyakarta: Gava Media.

Emerson, H dalam Soewarno Handayaningrat. (1990). *Efektivitas Organisasi*. Jakarta: Rineka Cipta.

Fitrifitanofa, Winda, Waskito Sutadi, & Budiharti Rini. (2013). Pengembangan Instrumen Tes Formatif Fisika Kelas XI Semester Gasal Program Akselerasi. *Jurnal Pendidikan Fisika, 1(1): ISSN 2338-0691*.

Guilford, J.P. (1956). *Fundamental Statistic In Psychology and Education*. 3rd Ed. New York: McGraw-Hill Book Company, Inc.

Gronlund, Norman E. dan Robert L.Linn. (1981). *Measurement and Evaluation In Teaching*. New York: Macmillan Publishing Company.

J. John, Yvonne. (2015). A "New" Thematic, Integrated Curiculum for Primary Schools of Trinidad and Tobago: A Paradigm Shift. *Interantional Journal if Higher Education, 4(3): ISSN 1927-6044*.

Khaerudin. (2017). Administrasi, Analisis Butir, dan Kaidah Penulisan Tes. *Jurnal Madaniyah 1 (XII), ISSN: 2548-6993*.

Mudiono, Alif, Gipayana M, & Madyono Suhel. (2016). Developing of Integrated Thematic Learning Model through Scientific Approching With Discovery Learning Tecnique in Elementary School. *International Academic Institute for Science and Technology, 3(10): ISSN 24543918*.

Retnawati, Heri. (2011). Validitas Reliabilitas dan Karakteritik Butir.

Ratopo, Kenny Anindia, Waskito & Harjunowibowo. 2014. Perbandingkan efektifitas pengembangan instrument tes sumatif fisika menggunakan metode Self dan Feedback Revision. *Prosiding Seminar Nasional Fisika dan Pendidikan Fisika (SNFPF), 5(1) ISSN: 2302-7827*.

Solichin, Mujianto. (2017). Abalsisi Daya Pembeda Soal, Taraf Kesukaran, Validitas Butir Tesm Interpretasi Hasil Tes dan Validitas Ramalan Dalam Pendidikan. *Jurnal Manajemen & Pendidikan Islam, 2(2); P-ISSN: 2503-3506*.

Sudijono, Anas. (2009). Pengantar Evaluasi Pendidikan. Jakarta: Raja Grafindo.

Thorndike, R.L. & Hagen, E. (1969). *Measurement and Evaluation in Psychology and Education*. New York: John Wiley and Sons, Inc.

Wardahani, Dya Febria, Aminuddin Prahatama Putra. (2016). The Development of Standardized Tes Instrument Natural Science Subjects Seventh Grade Junior High School in Kabupaten Banjar. *Procedding Biologi Education Conference*, 13(1), ISSN: 2528-5742)

Weeb, P.I & Pearson, P.J. (2012). Creative Unit and Lesson Planning Approach to Teaching Games for Understanding (TGfU). *New Zealand Physical Educator, 45(3), 17-22*.