Berkala Ilmiah Pendidikan Fisika
ISSN 2337-604X (print)
ISSN : 2549-2764 (online)

Vol 6 No 3 2018
https://ppjp.ulm.ac.id/journal/index.php/bipf
301-314

Development of Multi Representation Based Cognitive Instrument on Newton Law Material

Ramlah, Mustika Wati, Sarah Miriam, Saiyidah Mahtari
Physics Education Program, Faculty of Teacher Training and Education
Universita Lambung Mangkurat
ilah.nina13@gmail.com

DOI: 10.20527/bipf.v6i3.5294

Received : 10 September 2018 Accepted : 24 Oktober 2018 Published : 31 Oktober 2018

Abstract: Questions used in school have not been analyzed by the item and only in one representation. This study aims to produce quality multi-representation-based cognitive instruments on Newton's law material for junior high school students in Banjarmasin. This study specifically aims to describe (1) the validity of cognitive instruments, (2) the reliability of cognitive instruments, (3) the level of difficulty of cognitive instruments, and (4) the discrimination power of cognitive instruments based on multi-representation. The method used is the Research & Development (R & D) method by using the adaptation of the procedure models on Borg & Gall. Data were analyzed using classic formulas and through Rasch applications. The sample of the study was 204 eighth grade students from Public Junior High Schools 14 Banjarmasin, Public Junior High Schools 25 Banjarmasin, and Public Junior High Schools 28 Banjarmasin. The results showed that (1) the validity of cognitive instruments developed is considered valid, (2) the reliability of cognitive instruments developed is relatively reliable, (3) the level of difficulty of cognitive instruments developed is divided into two categories; very difficult and very easy, and (4) the discrimination power of cognitive instruments developed is divided into three categories; repaired, accepted but needs to be repaired, and accepted. It can be concluded that the multi-representation-based cognitive instrument on Newton's law material on eighth-grade students of junior high school in Banjarmasin is suitable for assessment of learning outcomes. So that the instruments developed can be used by teachers to assess student learning outcomes of the material of Newton's law.

Keywords: cognitive instruments, multi-representation based.

INTRODUCTION

Indonesia Government Regulation No.19 of 2005 concerning National Education Standards explains that each education unit conducts the process, implementation, assessment, and supervision of the learning process for the implementation of effective and efficient learning. The assessment is regulated in Chapter IV article 22 paragraph 1 & 2; the assessment of learning outcomes at the education level using several assessment techniques according to the competencies that must be mastered. The assessment test is in the form of written tests, observations, practice tests, oral test, performance test, portfolio, formative asessement, sumative asessment, individual or group assignments and so on (Arifin, 2009).

Learning outcomes can be grouped into three domains; cognitive, affective, and psychomotor (Arikunto, 2013). Cognitive domain is a domain that consists of the ability to remember, understand, apply, analyze, evaluate and create (Nurbudiyani, 2013). Measurement of learning outcomes can be done using tests. The test can be classified into two which are summative tests and formative tests. In general, the test consists of several questions that must be answered to find out the understanding and the mastery of concepts and material. These tests can be made with various forms of representation, because based on each format the representation that is completed will provide information on how far the student's mastery of each type of representation is presented (Aulia, 2015).

Multi representation is a way to express a concept in various ways and forms (Yusup, 2009). Multi representation encourages the formation of an understanding of information. Multi representation also helps students in describing problems and describing sketches and physical situations of the problem and directing students to understand information and knowledge to solve problems (Astuti, 2013).

Through interviews that were conducted with science subject teachers at State Junior High School 14 Banjarmasin, State Junior High School 25 Banjarmasin, and State Junior High School 28 Banjarmasin about students' cognitive assessment. From the interview, information obtained is that the cognitive assessment instrument used by the teacher was an instrument that came from the Dinas Pendidikan Kota. The instrument has no analysis of the item so the teacher does not know the validity and quality of the instrument used. From the analysis conducted on midterm exam questions, the result showed that only the C1 to C3 levels were used. Cognitive instruments only use words or descriptions and some pictures. When students are given problems in different representations, they will experience difficulties in solving the problem. If the problem is not addressed immediately, it will interfere with learning outcomes because of the low understanding of the concepts in the subjects given.

Thus, using multi representation is expected to assist students in building a deeper understanding of concepts, so that they can solve physic problems in different forms of objects such as verbal, image, graphic, or mathematical forms. When students are able to represent a

concept in the form of representation that is different from before it will help students in solving physical problems. Based on the description above, it is necessary to develop instruments in assessing student learning outcomes. The objective in this study was to produce multi-representation-based quality cognitive instruments on Newton Law material for junior high school students in Banjarmasin. In the matter of Newton's law allows many multi representations in learning.

LITERATURE REVIEW

Assessment is a systematic way that includes the activities of collecting, analyzing, interpreting the information used to make conclusions. Some of the things that become principles in an assessment are, a) the assessment process is an integral part of the learning process, b) assessment is a reflection of real world problems rather than the school world, c) using various sizes, methods, and criteria that are appropriate, d) it is holistic which covers all aspects of learning (Supranato, 2012). Bloom classifies learning outcomes into three, namely cognitive, affective, and psychomotor. A domain that emphasizes the development of intellectual abilities and skills is called cognitive. In 2001, Anderson and Krathwohl made a revision of Bloom's taxonomy to remembering, understanding, applying, analyzing, assessing, and creating (Gunawan & Palupi, 2016).

Multi representation can be interpreted as several ways to express a concept through different ways, forms, or formats (Astuti, 2013). Representation in physics learning can be used to minimize difficulties experienced by students in the process of learning physics. Multi representation has three benefits, namely as a complement to information, limiting interpretation, and understanding builders (Widianingtiyas & Fauzi, 2015).

Some important reasons for using multi representation are; multi-intelligence, visualization of the brain, helping construct other types of representations, are useful for qualitative reasoning, and are used for quantitative reasoning in abstract mathematical representations. In physics, many types of representations can be used. These types include, among others, 1) verbal descriptions are used to define a concept. 2) Images / diagrams are used to visualize something that is abstract. Diagrams that are often used in physics, namely motion diagrams, object free diagrams, field line diagrams, electric circuit diagrams, light diagrams, wavefront diagrams, and state energy diagrams. 3) Graphs are used to explain a concept that has a long

explanation and 4) mathematics is used to solve quantitative problems (Yusup, 2009).

Item Response Theory (IRT) or commonly referred to as modern test theory is a review of items using the answer theory of the item. This theory is a theory that uses the function of mathematics to connect between opportunities to answer correctly a scale with students' abilities (Fatkhudin, Bayu, & Agus, 2014). One of the main advantages of IRT compared to classical test theory is in the concept of IRT where problem statistics such as the level of difficulty, distinguishing power lies in the same scale as the measured ability of students (Alwi, 2012).

One of the advantages of IRT is the probability that the subject to answer the item correctly depends on the subject's skills and the characteristics of the item (Rosidah, 2018). The test score on the IRT has a value when compared to the characteristics of the item and the performance of a participant can be predicted by a set of factors (Ridho, 2007). The IRT has the ability to predict lost data based on individual response patterns (Amelia & Wati, 2018).

METHOD

This research is a Research & Development (R & D) using the adaptation of the procedure model from Borg and Gall (Gall, Gall, & Borg, 2003). The development procedure carried out in this study with the following steps, 1) potential and problems by analyzing the needs to know the instruments used by the teacher in learning, 2) data collection with interviews with teachers in schools, 3) product design by designing a product in the form of development multi-representation-based cognitive instrument of description questions by adjusting competency standards, basic competencies, and indicators based on KTSP curriculum syllabus, 4) design validation provides assessment of product designs made based on rational thought by validators, 5) improvement of design by improving products that have been made and perfecting products developed before the product is tested, 6) product trials by testing to students on a small scale to find out the problem received and questions that need to be omitted, 7) revision of the product by correcting the questions received so that they can be tested for use, 8) trial of usage by testing on a wide scale to assess the products that have been developed are feasible and have advantages, 9) revision of the final product by refining the product that has been developed and in accordance with the conditions that exist in the field based on the results of the

usege trial so that it can be known which questions are suitable for use.

The subject of this study is a cognitive instrument based on multi representation. Research was conducted from January to November 2017.

The data collection techniques in this study were in the form of tests and non-tests. Non-test technique is in the form of interviews, and test techniques are in the form of a validation questionnaire and giving a test in the form of a description problem.

Data analysis used in this study is by using classical theory in small-scale product trials and Rasch modeling with item response theory in wide-scale usage trials. Analysis of the results of validation is the average score obtained from the assessment of experts and practitioners and adjusted to the criteria based on (Widoyoko, 2009).

Table 1 Instrument Reliability Criteria

Reliability Coeficient	Criteria
$0{,}80 \leq r$	High Reliability
$0{,}40 \leq r \leq 0{,}80$	Intermediate Reliability
$r < 0{,}40$	Low Reliability

(Ratumanan & Laurens, 2011)

The results of the reliability analysis of the validator are adjusted to the instrument reliability criteria. The reliability criteria can be seen in Table 1. Validity analysis of the results of small-scale design trials is calculated using the product moment formula with rough numbers.

The results of the reliability analysis on product trials, adjusted to the criteria based on (Ratumanan & Laurens, 2011). After a large-scale usage trial, an analysis of the validity of the question was done using the Rasch program. The validity of an item depends on the Outfit Mean Square (MNSQ) value, the Z-Standard Outfit value (ZTSD), and the correlation of the resulting measurements. The value can be said as suitable as long as there is no MNSQ value, ZSTD value, and correlation measurement value that is outside the criteria simultaneously. In addition to MNSQ values and ZTSD values, correlation measurement values are located between 0.4 - 0.85 (Sumintono & Widhiarso, 2015). The criteria for accepted MNSQ Outfit is 0.5 <MNSQ <1.5, the received ZSTD Outfit value is -2.0 <ZSTD <+2.0, and the accepted Point Measure Correlation (Pt Mean Corr) is 0.4 <Pt Measure Corr 0.85.

In general, the reliability value of the test can be measured using Rasch modeling which can be indicated by the individual separation value and item separation and the Alpha Cronbach value. The criteria for interpreting individual separation values and item separation values for an instrument can

use criteria from (Sumintono & Widhiarso, 2015).

While the Cronbach Alpha value that is used to measure the interaction between individuals with the overall items is interpreted using Table 2.

Table 2 Intrepretation of Alpha Cronbach Values

Criteria	*Alpha Cronbach*
Bad	< 0,5
Poor	0,5 – 0,6
Intermediate	0.6 – 0,7
Good	0,7 – 0,8
Very Good	> 0,8

(Sumintono & Widhiarso, 2015)

Analysis of the level of difficulty of the problem based on the results of small-scale product trials can be calculated using the following formula based on (Rusilowati, 2014). Then the results obtained are adjusted to the criteria for the assessment aspects determined based on (Rusilowati, 2014).

The level of difficulty of the problem in a wide-scale usage trial can be analyzed using the logit number contained in the problem measurement column with Rasch modeling. The higher the logit value, the higher the level of difficulty of the problem. The criteria that can be used to interpret the level of difficulty of the problem based on (Sumintono & Widhiarso, 2015).

Discrimination power (DP) index can be obtained by calculating the formula based on (Rusilowati, 2014). Furthermore, the refractive index obtained in product trials and usage trials is interpreted using (Rusilowati, 2014).

RESULT AND DISCUSSION

The results of the procedure research on the development of multi-representation-based cognitive instruments based on the steps of making multi-representation based cognitive instruments that include nine steps. The results of these steps have been carried out as follows, potential and problems; the potential and problems in this study relate to the ability of teachers to make a test instrument for student learning outcomes in schools well, in addition to these problems, the test instrument consists of levels of C1 to C3 which mostly use verbal representation. Data collection; data collection in this case by means of interviews with science teachers in the schools of Public Junior High School 14 Banjarmasin, Public Junior High School 25 Banjarmasin Public Junior High School 28 Banjarmasin.

Product design; at this stage researchers developed a multi-representation-based cognitive instrument on Newton's law with 15 questions. The first thing to do is to determine the indicator of the question

according to Basic Competency in the material. The instrument developed are questions in the form of descriptions, instruments developed using several representations, namely verbal, mathematical representations, images, graphs, and tables.

In the Validation phase the instrument design then was validated by two validators who came from academics and a practitioner who was expert in his field. The results of the analysis of each validator can be seen in Table 3.

Table 3. Validity Result of Multi representasion based cognitive instrument

Assessment Aspect	Mean of Assessment	Criteria
General Construction	3,5	Very good
Item Validity	3,7	Very good
Mean	3,6	Very good
Reliability	0,68	Interme-diate

Based on the results of the instrument validation from the validators, it was obtained valid results with a slight revision for general construction and item validation. The results of the validation of the experts and practitioners showed that an overall total score of an average of 3.6. The average results are included in the very good category. While the results of reliability obtained were 0.68 and included in the medium category. In this design improvement phase researchers make improvements based on comments and suggestions given.

Small-scale product trials were conducted on 22 students. The number of items that were given to students are 15 items in the form of multi-representation descriptions. The results of the analysis of the validity and reliability of the product trials are in Table 4.

Table 4 Validity and reliability of the product trial

No.	Criteria	Item Number
1	Invalid	-
2	Poor Validity	-
3	Low Validity	-
4	Intermediate Validity	-
5	High Validity	1, 2, 3, 4, 8, 10, 11, 12, 13, 14, 15
6	Very High Validity	5, 6, 7, 9

Based on Table 4, it is known that valid questions are 15 questions and there are no invalid questions. This shows that the questions made can be used in extensive trials.

Table 5 Analysis Result of Item Difficulty Level

Note	Item Number
Easy	-
Intermediate	1, 2, 3
Difficult	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15

Analysis of the level of difficulty of the questions in the design tests can be seen in Table 5, from the results obtained there are two categories of difficulty levels of the problem, namely the problem with the Intermediate level of difficulty is 3 questions and questions with a high level of difficulty or difficult are 12 questions. Questions included in the medium category are questions with verbal, mathematical, and image representations which are questions about Newton law I. As for the questions about Newton law II and III included in the difficult category with verbal, mathematical, table, and graph representation. Judging from the results turns out that the problem with table and graph representation is a difficult category problem because many students cannot work on the problem with the representation. Comparison of the difficulty level of an instrument is 30% of the questions are in the easy category, 40% of the questions are in the medium category, and 30% of the questions are in the difficult category. However, from the result it was obtained that the questions are only divided into two categories and the proportion of questions are in too many difficult categories. According to (Arikunto, 2013) a good question is a matter that is not too easy or not too difficult. Problems that are too easy make students not stimulated to solve a problem. Whereas questions that are too difficult can make students become desperate so they do not have the motivation to solve them (Taufiq, 2015). The analysis results of the discrimination power are in the following product trials.

Table 6 Analysis Result of Discrimination Power of Item

Note	Item Number
Omitted	-
Improved	1, 2, 11, 15
Accepted but needs to be repaired	3, 8, 10, 12, 13, 14
Accepted	4, 5, 6, 7, 9

Based on Table 6, it can be concluded that there are three categories, namely the problem that is repired, the problem accepted but needs to be corrected and the question accepted. This is because problems with improved categories can be solved by students with high abilities but cannot be completed by students with low abilities. Problems with categories are accepted but need to be fixed and can be completed by students with high abilities and students with low abilities but unable to complete correctly. According to (Arikunto, 2013) a good question is a question that can distinguish the ability between smart student and student who are not smart. Based on the representation used in the instrument, it is known that the problem with image representation is a fixed and

accepted category question but needs to be repaired rather than a question with other representations.

Based on the results of the analysis obtained from product trials, the researcher revised the instrument, namely, the vocabulary and numbers contained in the instrument and discarded 4 questions. Usage trials are conducted with a large number of students. Trials were conducted on three public schools, namely Public Junior High School 14 Banjarmasin, Public Junior High School 25 Banjarmasin Public Junior High School 28 Banjarmasin with a total of 204 students. The number of questions tested was 11 items in the form of multi representations. The analysis in this section uses rasch method analysis which is obtained by the validity of the questions, reliability of the questions, and the level of difficulty of the questions and the ability of each student in solving the given questions.

Table 7. Item Suitability

Entry	***Outfit***		***Pt-Measure***
Number	**MNSQ**	**ZSTD**	***Correlation***
2	1,30	2,5	0,68
6	1,27	1,8	0,58
5	1,25	1,7	0,62
1	1,16	0,6	0,46
4	0,84	-1,1	0,53
3	1,02	0,2	0,66
7	0,91	-0,6	0,63
10	0,84	-0,8	0,41
8	0,70	-1,7	0,46
9	0,58	-2,4	0,46
11	0,38	-4,1	0,51
Mean	0,85	-0,7	
S.D	0,28	1,9	

Based on Table 7 above, it can be seen from MNSQ, ZSTD, and Pt-measure correlation values. Based on these criteria it appears that all questions fall into the valid category. This is consistent with the statement (Sumintono & Widhiarso, 2015) that there is no misfit item if the items that are simultaneously outside the criteria for Outfit Means Square values, Z-standard Outfit, and Point Measure Correlation values. According to (Sumintono & Widhiarso, 2015), the suitability value of items consisting of MNSQ, ZSTD, and Pt Measure Corr Outfit was strongly influenced by the sample size. Then the large sample size in this study has a positive effect on the results of the study.

Tabel 8 Detection of Deferential Item Function

Item number	**Probability**
1	0,3256
2	0,0156
3	0,0505
4	0,0181
5	0,0010
6	0,1091
7	0,7200
8	0,4758
9	0,9690
10	0,4494
11	0,1162

The validity of the item questions can also be known through the bias question by looking at the probability value. The question is bias if the probability value of items is less than 5%

(Sumintono & Widhiarso, 2015). Based on Table 8 above there are three biased questions, namely questions number 2, 4 and 5. The item is called a bias if it is found that one individual with certain characteristics is more beneficial than an individual with other characteristics (Sumintono & Widhiarso, 2015).

Tabel 9. Statistic Summary

	Respondent		**Item**	
	Outfit		*Outfit*	
	MNSQ	ZSTD	MNSQ	ZSTD
Mean	1,03	-0,2	0,93	-0,3
Separation	1,79		29,04	
Person reliability	0,76		1,00	
Cronbach alpha	0,99			

The reliability of the usage test can be seen in Table 9. The value of student reliability obtained in this study amounted to 0.76 with Intermediate category. The reliability value of the items obtained is 1.00 with very good category. While Cronbach's alpha value is 0.99 with a very good category. So, based on the value of the reliability of students and the value of reliability of the items can be concluded that the consistency of the students' answers is sufficient and the quality of the items is excellent. In addition to the value of student reliability and reliability of the item, it is also known that the separation person value is 1.79 and the separation item value is 29.04. The greater the value of separation, it is the better the quality in terms of overall students and items. According to (Azizah, Wati, Salam, & Mahtari, 2017) good test questions that have high validity also have high reliability. In addition to the values of validity and reliability, also known the level of difficulty of the items as follows.

Tabel 10 Item Difficulty Level

Difficulty Level	**Item Number**
Very Difficult	10, 8, 9, 11
Difficult	-
Easy	2, 3, 1
Very Easy	5, 7, 6, 4

Based on Table 10, it is known that there are three levels of questions, which are very difficult at 36.36%, easy at 27.27%, and very easy at 37.37%. The proportion of questions in the medium category is less than the category of easy and difficult categories. Problems with image representation, verbal, mathematical, and table are in very difficult categories. Problems with easy categories are questions with verbal, mathematical, and image representations. As well as questions with very easy categories are questions with mathematical representation, graphics, images, and verbal. Judging from these results, the problem with image representation is more difficult than the

problem with other representations. According to (Amalia & Widayati, 2012) the items were good if the items were not too easy and not too difficult.

Based on Table 11 it can be seen that the discrimination power of the items is divided into four categories, namely the category of questions that are not used or discarded, questions that are repaired, problems accepted but need to be repaired, and questions accepted. From these results it is known that the problem that is omitted is a problem that cannot distinguish abilities between students who have high abilities and students who have low abilities. This is in accordance with (Akbar, 2013) questions that can be done by smart children or less intelligent children are not good questions. A good question is a question that can only be done by clever students.

Tabel 11 Analysis Result of Item Discrimination Power

Note	Item Number
Omitted	1, 2, 3
Repaired	4, 5, 6
Accepted but needs to be repaired	7
Accepted	8, 9, 10, 11

In addition to validity, reliability, level of difficulty, and discrimination power is also known the mean value of the representation of students. The results of the students' multi representation average analysis can be seen as follows.

Tabel 12 Mean of Multi-Representation

Question	Verbal	Mathe-matical	Image	Table	Chart
Newton Law I	9,9	9,3	9,2	-	-
Newton Law II	7,7	7,7	7,2	-	8,2
Newton Law III	4,2	4	3,8	4	-

Based on Table 12 it can be seen that in Newton's first lawproblem it is known that the average representation of students who can do the most problems is in verbal representation, which is equal to 9.9. In Newton's second law problem it is known that the average representation of students who can do the most problems is in the graphical representation of 8.2. In Newton's third law problem, it is known that the average representation of students who can do the most problems is in verbal representation, which is 4.2, on average, on verbal representations. Judging from the overall questions provided it appears that more students are working on problems with verbal representations than other representations and the most difficult problem for students to do is the problem

with image representation. Judging from these results it is known that the problem with image representation is a problem with categories that are difficult compared to other representations.

Based on the results of the analysis on the usage test, the researcher revised the vocabulary and numbers contained in the instrument. The final results obtained are 5 questions that are omitted, and produce a final product of 6 questions which are equipped with a grid and answer key.

CONCLUSION

Based on the results of the development and testing in the field, it can be concluded that the multi-representation-based cognitive instrument on Newton's law material on the eighth-grade junior high school students in Banjarmasin can be said as feasible to be used for assessment of learning outcomes. This is in accordance with the findings as follows: 1)the validity of the instrument developed in this study is classified as valid, 2) the reliability of instruments developed in this study is relatively reliable, 3) the level of difficulty of items developed on instruments is divided into two categories, namely 4 very difficult category questions and 2 very easy category questions, and 4) the discrimination power of the items developed is divided into three categories: 1 category problem is repaired, 1 category problem is accepted but needs to be repaired, and 4 category questions are accepted. So that the instruments developed can be used by teachers to assess student learning outcomes in the material of Newton's law.

REFERENCES

Akbar, S. (2013). *Instrumen Perangkat Pembelajaran*. Bandung: PT Remaja Rosdakarya.

Alwi, I. (2012). Kriteria Empirik Dalam Menentukan Ukuran Sampel Pada Pengujian Hipotesis statistik dan Analisis Butir. *Jurnal Formatif*, *2*(2), 140–148.

Amalia, A. N., & Widayati, A. (2012). Analisis Butir Soal Tes Kendali Mutu Kelas XII SMA Mata Pelajaran Ekonomi Akuntansi di Kota Yogyakarta Tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, *10*(1), 1–26.

Amelia, H., & Wati, M. (2018). Pengembangan Instrumen Kognitif untuk Mengukur Penalaran Siswa SMP di Kota Banjarmasin pada Materi Cahaya. *Berkala Ilmiah Pendidikan Fisika*, *6*(1), 83–97.

Arifin, Z. (2009). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.

Arikunto, S. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Astuti, Y. W. (2013). Bahan Ajar Fisika SMA dengan Pendekatan Multi Representasi Pada Mata Kuliah Pendahuluan Fisika Zat Padat. *Jurnal Pendidikan Sains*, *1*(4), 382–389.

Aulia, L. R. (2015). Pengembangan Instrumen Tes Berbasis Multi Representasi Pada Mata Kuliah Pendahuluan Fisika Zat Padat. *Jurnal Inovasi Dan Pembelajaran Fisika*, (1–7).

Azizah, N., Wati, M., Salam, A., & Mahtari, S. (2017). Pengembangan Instrumen Kognitif Keterampilan Proses Sains Siswa SMP Pada Materi Pesawat Sederhana. *Berkala Ilmiah Pendidikan Fisika*, *5*(3), 340–350.

Fatkhudin, A., Bayu, S., & Agus, S. (2014). Item Response Theory Model Empat Parameter Logistik Pada Computerized Adaptive Test. *Jurnal Sistem Informasi Bisnis*, *2*, 121–129.

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational Research*. Boston: Library of Congress Cataloging-in-Publication Data.

Gunawan, I., & Palupi, A. R. (2016). Taksonomi Bloom–revisi ranah kognitif: kerangka landasan untuk pembelajaran, pengajaran, dan penilaian. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, *2*(2).

Nurbudiyani, I. (2013). Pelaksanaan Pengukuran Ranah Kognitif, Afektif, dan Psikomotor Pada Mata pelajaran IPA Kelas III SD Muhammadiyah Palangkaraya. *Pedagogik Jurnal Pendidikan*, *8*(2), 14–20.

Ratumanan, T. ., & Laurens, T. (2011). *Penilaian Hasil Belajar pada Tingkat Satuan Pendidikan*. Surabaya: Unesa University Press.

Ridho, A. (2007). Karakteristik psikometrik tes berdasarkan pendekatan teori tes klasik dan teori respon aitem. *Jurnal Psikologi INSAN*, *2*(2), 1–27.

Rosidah, N. A. (2018). Karakteristik Tes Keterampilan Berpikir Kritis Berdasarkan Teori Respon Butir. *Jurnal Inovasi Dan Pembelajaran Fisika*, *5*(1), 55–63.

Rusilowati, A. (2014). *Pengembangan Instrumen Penilaian*. Semarang: Unnes.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim Komunikata.

Supranato, S. (2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.

Taufiq, A. U. (2015). Pengembangan Tes Kognitif Berbasis Revisi Taksonomi Bloom Pada Materi Sistemreproduksi Untuk Siswa SMA. *Jurnal Biotek*, *3*(2), 1–17.

Widianingtiyas, L. S., & Fauzi, B. (2015). Pengaruh Pendekatan Multi Representasi dalam Pembelajaran Fisika Terhadap Kemampuan Kognitif Siswa SMA. *Jurnal Penelitian Pendidikan Fisika*, *1*(1), 31–38.

Widoyoko, E. P. (2009). *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Pelajar.

Yusup, M. (2009). Multirepresentasi dalam Pembelajaran Fisika. *Jurnal Pendidikan Teknologi Informasi Dan Komunikasi*, *2*(1).