



A STUDY ON THE QUALITY OF FINAL EXAM ITEMS MADE BY THE TEACHER AT XII GRADE STUDENTS OF SENIOR HIGH SCHOOL IN SIDRAP REGENCY

AUTHORS INFO

A.S. Alonemarera
Universitas Sembilanbelas November Kolaka
alonemarerausn@gmail.com
+6285299688649

ARTICLE INFO

E-ISSN: 2721-0804
Vol. 2, No. 1, June 2020
URL: <http://usnsj.com/index.php/biology>
© 2020 JBSE All rights reserved

Suggestion for the Citation and Bibliography

Citation in Text:

Alonemarera, A. S. (2020)

Bibliography: Alonemarera, A. S. (2020). A Study on The Quality of Final Exam Items Made By The Teacher at XII Grade Students of Senior High School in Sidrap Regency. *Journal of Biological Science & Education*, 2 (1, June), 32-41.

Abstract

Bad quality of the questions is one of the factors that can lead to errors in the grading of student learning outcomes. It gives a lower students' score than the appropriate score and will give an impact on school grades (NS). In this case, school grades (NS) have an effect of 40% in determining the passing of national exams (UN). Therefore, this study was conducted to assess the quality of first semester final exam items made by biology class XII teacher based on some aspects; they are the material, construction, and language. In this study, the researcher used purposive sampling technique in order to collect the data. Items were assessed using a multiple-choice item evaluation sheet which has been declared valid by an expert validator. By combining the material aspects of the assessment data, construction, and language, research results obtained in general show that of the 120 items assessed, there are 104 (86,7 %) items with excellent criteria and 16 (13,3%) items with right criteria. Thus, the items made by biology class XII teachers are still suitable for final semester exam of the 2013/2014 academic year Sidrap Regency.

Keywords: Quality of Questions, Multiple-Choice

A. Introduction

Evaluation program is one of the activity carried out by educational institution in order to improve their performance and quality. Therefore, evaluation becomes one of the essential elements in the education system. According to Permendikbud no. 23 in 2016 stated that evaluation or assessment has a role in measuring and controlling the quality of education (Kemendikbud, 2017). In addition, Dikdasmen (2014) stated that assessment or evaluation conducted by educator and education unit is an internal evaluation, while the assessment undertaken by the government is an external evaluation that makes students as the object, including student learning outcomes.

The scoring system to maintain the results level of students in the senior high school (SMA) conducted by the government in national test Program (UN) (Republic of Indonesia Government Regulation number 19 of 2005 concerning National Education Standards). The implementation of the national examination (UN) and the determination of graduation of students in the education unit refer to Badan Standar Nasional Pendidikan Number 0022/P/ BSNP/XI/2013 concerning the Standard Operating Procedure (POS) for the Implementation of the 2013/2014 Academic Year National Examination. The graduation of participants students from the

education unit are determined through teacher council meetings based on the following conditions/criteria: (1) complete all learning programs, (2) get a good minimum score, (3) pass the school examination (US), and (4) pass the national examination (UN) (BSNP, 2013)

Therefore, if the students accomplish the graduation requirements set by the education unit based on the final grade (NA) obtained, it can be declared that students passed the national examination (UN). Final grade (NA) is the combination of school grades (NS) of 40% and UN score of 60%, from subjects tested nationally. In this case, the combination of school exam scores (US) with an average semester report scores of 3,4, and 5 with a standard of 30% and 70% for average grades report card (Badan Standar Nasional Pendidikan, 2013). Thus, it can be concluded that the special scores of students while studying at a particular educational institution also influence the determination of graduation. These special scores should be used as a basis for developing the right quality learning achievement test instruments.

Based on observations from one of the high schools in Sidrap Regency, it was found that the teacher never knew how the quality of the questions had been arranged. It means that there was no effort to improve the quality of the questions. On the other hand, the quality of questions that are not good can be one of the factors in the presence of errors in providing learning outcomes to students. Therefore, this study was undertaken to uncover quality of final exam items made by biology teachers who have been tested in the academic year 2013/2014, at XII grade students of senior high school in Sidrap Regency based on aspects of the material, construction, and language.

B. Literature Review

1. Characteristics of Evaluation of Learning Outcomes

Institution or students must ultimately obtain information from their instructors or educators (Purwanto, 2012). Evaluation of learning outcomes is as a field activity, has characteristics that distinguish it from other areas of activity. According to Sudijono (2012), one of the traits possessed by the evaluation of learning outcomes is the difficulty of avoiding errors in measurement. Measuring instruments consist of written tests that are used orally by educators in measuring students as an effort to assess their learning outcomes. Arikunto (2013) states that educators give scores called "scores" based on the number of correct answers or quality answers provided by students. The question that then arises from the marking event is whether the score that has been given by the educator to the students is the same or can be considered the same as the actual score or not. The real score is interpreted as a score that truly reflects the learning achievement of the students concerned.

They are two things that may occur related to the grading according to Sudijono (2012), namely: (a) the possibility of the value being given is lower than the amount that should be, (b) the potential of the value being given is higher than the value that should be. If there is one of the two possibilities that have been mentioned, then comes what is referred to as an error/error in the assessment of learning outcomes. Several factors can cause mistakes that occur in the assessment of learning outcomes. One factor is the measuring instrument, where the measuring device used in the evaluation cannot measure correctly/precisely what should be measured.

2. Item Test Qualitative Analysis

According to Anastasia & Urbina (1997), to conduct a qualitative analysis of items, it can be examined in terms of content and form or procedures/rules for improving quality in judgment. Meanwhile, according to Mardapi (2004) before testing, theoretical or qualitative analysis of questions is first carried out, taking into account the suitability of the items (which have been made) with the necessary abilities and indicators to be measured and the fulfilment of requirements for material, construction and language aspects.

Before a test tool is used, then the item analysis must be done qualitatively regarding the rules of writing written questions. Moderator or discussion techniques and panel techniques can be used as qualitative analysis techniques on items. In the moderator or discussion technique, there is one person as a mediator. Each piece is discussed and solved together by examining and discussing the writing conventions. Also, participants/reviewers were allowed to comment on their fields of expertise. While panel technique is a technique carried out by participants by working independently, it may be in different places, in assessing/examining the items. The reviewers will consider the issues, both in terms of content (material), construction and discussion (Ali & Khaeruddin, 2012). Related to the previous explanation, it

can be concluded that item analysis is qualitatively done through a review of the material aspects, construction aspects, and language aspects.

C. Methodology

1. Research Design

This research is a descriptive study that illustrates the quality of first semester final exam items made by biology class XII teacher in the 2013/2014 academic year in the high schools in Sidrap Regency based on the assessment of the material, construction, and language aspects. The sampling technique used was *purposive sampling*. In this study, the school chosen as a place to take items was a school that used multiple-choice questions. The three selected schools are SMA 1 Pangsid, SMA 1 Watang Pulu, and SMA Negeri 2 Panca Rijang.

2. Instrument

The instruments used in this study were 120 items of first semester final exam documents and assessment sheets of validated multiple-choice items developed based on the format Direktorat Pembinaan SMA (2010). Example of Indicators for each aspect of assessment can be seen in table 1. Data collection was done using non-test techniques, namely interviews and documentation to obtain a set of odd semester final exam questions along with learning tools (syllabus and lesson plans) and examine or assess documents (*documentary analysis*) items using the quality assessment sheet items.

Table 1. Example of Aspect's Indicators

Aspect	Material	Construction	Language
Example of Indicators	The cognitive level of the problem corresponds to the cognitive level of the item's indicator	The item's subject is clearly and decisively formulated	Use Indonesian language properly/official Indonesian Spelling System (EYD)
	Material items are based on the content in item indicators	The formulation of subject items and the choice of answers are statements that are only needed.	Use communicative language
	Write the equivalent answer choices (homogeneous) The answer choices come from the same material in the item's subject.	The item's subject does not give directions to the correct answer Answer choices in the form of numbers or times are arranged in the order of the size of the number, or chronologically the time	Do not use local/taboo language The answer choices do not repeat the same words/groups of words unless they are a unity of understanding

3. Technique of Data Analysis

Research data obtained were analyzed through descriptive analysis techniques that are displayed in tables and diagrams. The results of the quality assessment of each for aspects of the material, construction, language are obtained in the form of categories by calculating the percentage of the number of indicators fulfilled for each aspect of the assessment. Furthermore, a general category is obtained for each item by averaging the value of the percentage of assessment results from the material, construction and language aspects. The categorization of item quality can be seen in Table 2

Table 2. Categories of Item Quality

Percentage of Fulfillment Rating Indicator Indicator (%)	Criteria	Information
0-20	Not good	Rejected
21-40	Not good	Repaired
41-60	Pretty good	Repaired
61-80	Well	Accepted, but should be fixed
81-100	Very good	Accepted

(adapted from Riduwan, 2012)

D. Findings and Discussion

1. Findings

a. Material Aspect

Table 3. Results of Item Points Assessment in Material Aspect

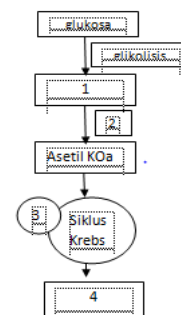
Assessment aspects	Category	Number of items			Total items	%
		SMAN 1 Pangsidi	SMAN 1 Watang Pulu	SMAN 2 Panca Rijang		
Material	Very good	31	31	4	66	55
	Well	9	9	22	40	33,3
	Pretty good	-	-	4	4	3,3
	Not good	-	-	10	10	8,3
	Not good	-	-	-	-	-
Total		40	40	40	40	120

Table 4. The Example Question Corresponding to the Cognitive Level of Item Indicators

Test Indicator	Explain the process of protein synthesis
Question Number	23
Question	<p>If the sequence of nitrogen bases in the DNA template chain is as follows:</p> <p style="text-align: center;"> $\overleftarrow{\text{A T S}} \quad \overrightarrow{\text{G G A}} \quad \overrightarrow{\text{S S T}} \quad \overrightarrow{\text{A A G}} \quad \overrightarrow{\text{A A S}}$ </p> <p>Then the sequence of nitrogen bases in the mRNA of the transcription results is ...</p> <p>a. TAG-SST-GGA-TTS-TTG b. UAG-SSU-GGA-UUS-UUG c. AUS-GGA-SSU-AAG-AAS d. ATS-GGA-SST-AAG-AAS e. UAG-SST-GGA-TTS-UUG</p>
The Answer	B

Table 5. The Example Question that is not Corresponding to the Cognitive Level of Indicators

Test Indicator	Explain the stages of the respiratory reaction
Question Number	12
Question	<p>Respiration is illustrated in the scheme below!</p> <p>From the following scheme, pyruvic acid and pyruvic acid decarboxylation are numbered</p> <p>a. 1 and 2 b. 2 and 3 c. 2 and 4 d. 3 and 2 e. 4 and 3</p>
The Answer	A



c. Language Aspect

Table 10. Results of Item Points Assessment in Language Aspects

Assessment aspects	Category	Number of items			Total items	%
		SMAN 1 Pangsid	SMAN 1 Watang Pulu	SMAN 2 Panca Rijang		
Language	Very good	-	1	-	1	0,8
	Well	40	39	40	119	99,2
	Pretty good	-	-	-	-	-
	Not good	-	-	-	-	-
	Not good	-	-	-	-	-
Total		40	40	40	40	120

Table 11. The Example of Questions that are not based on the EYD (Official Indonesian Spelling System)


Question Number	3
Question	Etiolation event in sprouts occurs because of a. stunted growth due to lots of light b. rapid growth due to no light c. auxin buildup on the stem d. inhibition of auxin due to lack of light e. grain dormancy due to moisture
The Answer	B
Question Number	22
Question	In human, ovum there are.... a. 22 autosomes + X b. 22 autosomes + Y c. 22 autosomes + XX d. 22 autosomes + XY e. 44 autosomes + XX
The Answer	A
Question Number	23
Question	If the sequence of nitrogen bases in the DNA template chain is as follows: <div style="text-align: center;">  </div> Then the sequence of nitrogen bases in the mRNA of the transcription results is ... a. TAG-SST-GGA-TTS-TTG b. UAG-SSU-GGA-UUS-UUG c. AUS-GGA-SSU-AAG-AAS d. ATS-GGA-SST-AAG-AAS e. UAG-SST-GGA-TTS-UUG
The Answer	B

Table 12. Examples of Very Good Criteria for Language Aspects

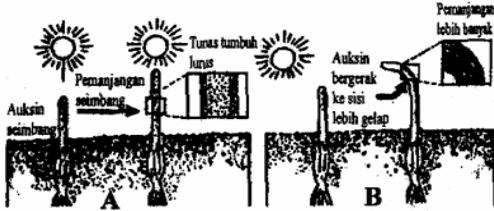
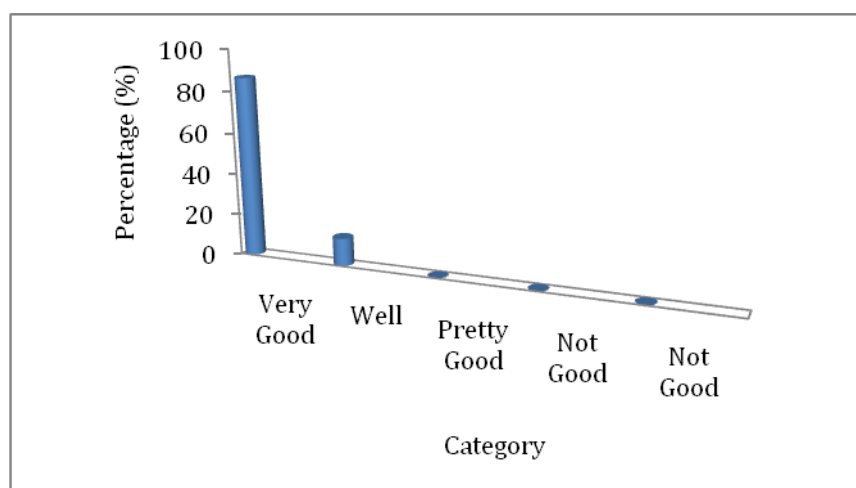
Question Number	2
Question	<p>Look at the picture on the effect of light on hormones and plant growth below!</p>  <p>Which statement is true about the chart above?</p> <p>A. Light accelerates the production of auxin in plant B, while light inhibits the production of auxin in plant A so that it grows straight.</p> <p>B. Plants A and B have different directions of growth because the intensity of the received light is not the same.</p> <p>C. Plant A gets light from above, so auxin goes to the bottom, while Plant B receives light from the side so that auxin goes to the side.</p> <p>D. Plant A grows slowly, while Plant B proliferates, so Plant B turns toward the light.</p> <p>E. Plant A grows straight because the distribution of auxin is evenly distributed, while Plant B grows toward the light. After all, auxin decomposes if exposed to light.</p>
The Answer	E

Table 13. Results of the Overall Item Assessment

Category	Number of items			Total items	%
	SMAN 1 Pangsidi	SMAN 1 Watang Pulu	SMAN 2 Panca Rijang		
Very good	40	38	26	104	86,7
Well	-	2	14	16	13,3
Pretty good	-	-	-	-	-
Not good	-	-	-	-	-
Not good	-	-	-	-	-
Total	40	40	40	120	100

**Figure 1.** Results of the Overall Item Assessment

2. Discussion

a. Material Aspect

The results of the assessment of material aspects showed that in general the questionable items in the evaluation are indicator number 1, which amounted to 41,7 % items had a cognitive level that was not base on the level of cognitive indicators. It can be stated that the suitability of the cognitive level between items and indicators of multiple-choice questions made by biology teachers was generally categorized as inappropriate. This result was also supported by Lodang & Bara's research (2012) which found that in general the level of appropriateness between formative evaluation instruments and cognitive goals of biology subject in SMP Negeri 1 Watansoppeng was categorized as less or inappropriate.

The items in table 4 have very good criteria, namely the items about the process of reading DNA into RNAm, by the indicator questions that explain the process of protein synthesis. It also has a homogeneous and logical choice of answers and has one correct answer. Specifically for question indicators and item questions, it is at the C2 cognitive level (understanding), this is reinforced by the presence of the verb "explain" on the question indicator. In Anderson & Krathwohl (2010), the verb "to explain" takes place when students can create or use a causal model in a system. This is in line with the content of the questions that require students to be able to arrange a series of nitrogen bases in the form of mRNA if a series of DNA nitrogen bases is provided as in the item.

The question indicator in table 5, which explains the stages of respiration reaction, is at the cognitive level C2 (understanding), this is reinforced by the word "explain" on the question indicator. In Anderson & Krathwohl (2010), the verb "to explain" takes place when students can create or use a causal model in a system. So students are really required to be able to construct the meaning of the learning material. The point problem is at the cognitive level C1 (remembering), because students are only required to remember parts of the aerobic respiration scheme. The components of the aerobic respiration scheme have been explained by the teacher when giving the subject matter. So that if it is raised again in the form of questions as above, then only the ability of students to remember what has been learned previously is measured, not the ability of students to understand the subject matter in indicators. Thus, the item is said to be incompatible with the question indicator in terms of cognitive levels.

b. Construction Aspect

The results of the construction aspect assessment show that generally, the problem item in assessment is indicator number 12, namely from 45 things that have answer choices in the form of numbers, as many as 31,1% items that do not arrange answer choices based on the order of the size of the number. In order to overcome the problem of how to write or place the choice answers in the form of number, only precision and accuracy are needed when writing items. The results of Arif & Gusryani's research (2010), for analysis of items concerning construction aspects, have an average rating of 82,3 and are included in right categories.

At a glance, the choices of answer questions in table 7, with indicator items about photosynthesis events, are arranged based on the smallest number value to the largest number value. However, if you pay attention, the answer choices (b) have a numeric value higher than the number values in the answer choices (c). So the choice of answers from the items above is said to not meet the assessment criteria, that is, answers in the form of numbers are not arranged based on the order of the size of the numbers. The next discussion on the primary formulation of questions in table 8, with indicators about the matter of protein synthesis. If you pay attention, some letters should not need to be included, the letter "Y" which indicates a type of protein. Without these letters, the problem can be solved because the stages presented by the issues are the stages of protein synthesis that are generally accepted. Arikunto (2013) states that writing items is a difficult task so that if the formulation/ construction of the item is considered to be poor, then the item can be fixed or even replaced. For example, issues that have poor distractors in sentence formulations can be rewritten with changes that are considered appropriate. According to Sumardi (2012) the improvement of the wording of questions is one of the teacher's efforts in developing test instruments. Mardapi (2003) in Sumardi (2012) revealed that one of the goals of developing tests is to improve student progress/learning outcomes.

c. Language Aspect

The results of the assessment of the language aspect showed that generally, the questionable items in the evaluation is indicator number 15, namely from the 120 things that had been

compiled, as many as 99,2% items that did not use language by the rules of the Indonesian language (EYD), especially in use typeface for writing answer choices and giving the number of dots at the end of the subject matter.

According to Bahri (2019) who states that teachers can arrange evaluation questions in good quality, but they have to follow the rules or technical instructions in providing the correct items. One of the rules for writing questions from the aspect of language is the formulation of questions written about the Indonesian language rules, EYD. In Osnal et al (2016) stated that the ability to use the correct Indonesian language is important in producing a right evaluation question formulation, namely a matter of avoiding things such as the bias of meaning or ambiguity. Thus, the items that are arranged to pay attention to EYD become one of the supporting factors that produce a quality test.

Item number 3 (regarding the etiolation event in sprouts) in table 11, there was an error in the selection of letters for writing answer choices, which should be the choice of response letters written using capital letters. In addition, there is also an error in the use of the number of dots at the end of the subject matter; namely there are seven dots, which should only use four dots. In question number 22 (about the number and type of chromosomes in female sex cells), there was also an error in using a colon at the end of the subject, which should end with a dot of 4 pieces. As for question number 23 (about reading the DNA code into RNAm), there was an error in the selection of lowercase letters for writing the answer choices and using the number of dots at the end of the subject matter, i.e. using only three points, which should have ended with four dots.

The items in table 12, about the effect of light on hormones and plant growth, can fulfil all assessment items for the language aspect, especially in the use of Indonesian (EYD) rules, including the procedure for writing multiple-choice questions. The item above is written with a question sentence, so that the beginning of the statement of choice of answers begins with capital and ends with a period. Broadly speaking, the procedure for writing multiple-choice tests, according to Kostania (2013) is as follows; if the body (stem) or subject matter is written with an unfinished sentence, then the beginning of the question sentence must be written in capital letters and the answer choices have the beginning of the statement written in lowercase (except for the name of the self and place name). In addition, because the body of the question is written with an unfinished sentence, then at the end of the sentence the item must be given four points. The first three points are points for the body of the item written with unfinished sentences, while the last point is the endpoint for alternative/answer choices. Thus, at the end of each answer choices are no longer given a dot. If the body of the question is written using a question sentence, then the beginning of the question sentence is written using capital letters and at the end is given a question mark. Each opening of the alternative answers is written in capital letters and end with a period.

E. Conclusion

Based on the results of the study, by combining the data on the results of the assessment aspects of material, construction, and language obtained as many as 86.7 % items with perfect criteria and 13.3% items with right tests. Thus, the items made by biology class XII teachers are still suitable for use in the final semester exam of the 2013/2014 academic year Sidrap Regency.

F. References

- Ali, S. & Khaeruddin. (2012). *Evaluasi Pembelajaran*. Makassar: Badan Penerbit UNM.
- Anastasia, A. & Urbina, S. (1997). *Psychological Testing (Seventh Edition)*. New Jersey: Prentice-Hall.
- Anderson, L. & Krathwohl, D. (2010). *Kerangka Landasan Untuk Pembelajaran, Pengajaran, dan Asesmen (Revisi Taksonomi Pendidikan Bloom) (Terjemahan oleh Agung Prihantoro)*. Yogyakarta: Pustaka Belajar.
- Arif, H. & Gusryani. (2010). Analisis Butir Soal Ulangan Harian IPA Semester Genap Kelas VII SMPN Se - Kec. Mandau Tahun Pelajaran 2009 / 2010. *Jurnal Pendidikan Sains dan Biologi*, 7 (2).
- Arikunto, S. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Badan Standar Nasional Pendidikan. (2013). *Prosedur Operasi Standar Penyelenggaraan Ujian Nasional Tahun Pelajaran 2013/2014*. Diunduh pada 05 Desember 2013 dari <https://bsnp-indonesia.org/id/wp-content/uploads/2013/12/POS-UN-Tahun-Pelajaran-2013-2014.pdf>.

- Bahri, M. (2019). Teknik Penulisan Soal Pilihan Ganda, Benar Salah dan Menjodohkan. Retrieved 20 May 2020 from <https://www.slideshare.net/MulyadiBahri/teknik-penyusunan-soal-pilihan-ganda>.
- Dikdasmen. (2014). Ujian Nasional: Evaluator Hasil Belajar Peserta Didik. Diakses pada 02 Februari 2014 dari <http://dikdasmen.kemdikbud.go.id/index.php/ujian-nasional-evaluator-hasil-belajar-peserta-didik/>.
- Direktorat Pembinaan SMA. (2010). Petunjuk Teknis Analisis Butir Soal di SMA. Diunduh pada 08 Desember 2013 dari <http://id.scribd.com/doc/106526488/34-juknis-analisis-butir-soal-isi-revisi-0104>.
- Kemendikbud. (2017). Permendikbud Nomor 23 Tahun 2016 tentang Standar Penilaian. Diunduh pada 03 Maret 2020 dari https://bsnp-indonesia.org/wp-content/uploads/2009/09/Permendikbud_Tahun2016_Nomor023.pdf.
- Kostania, G. (2013). Pedoman Penyusunan Soal Pilihan Ganda. Diunduh 10 Desember 2013 dari https://www.academia.edu/32464999/PEDOMAN_PENYUSUNAN_SOAL_PILIHAN_GANDA.
- Lodang, H. & Bara, N.A.S. (2012). Analisis Kesesuaian antara Instrumen Evaluasi Formatif dengan Tujuan Kognitif Pembelajaran Biologi di SMP Negeri 1 Watansoppeng. *Jurnal Bionature*, 13(2).
- Mardapi, D. (2004). *Penyusunan Tes Hasil Belajar*. Yogyakarta: Program Pascasarjana UNY.
- Osnal, Suhartoni & Wahyudi, I. (2016). Meningkatkan Kemampuan Guru dalam Menyusun Tes Hasil Belajar Akhir Semester melalui Workshop di KKG Gugus 02 Kecamatan Sumbermalang Tahun 2014/2015. *Jurnal Pancaran*, 5 (1).
- Peraturan Pemerintah Republik Indonesia. (2005). Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan. Diunduh pada 27 Februari 2020 dari <https://pelayanan.jakarta.go.id/download/regulasi/peraturan-pemerintah-nomor-19-tahun-2005-tentang-standar-pendidikan-nasional.pdf>.
- Purwanto, N. (2012). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya.
- Riduwan. (2012). *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.
- Sudijono, A. (2012). *Pengantar Evaluasi pendidikan (Edisi Revisi)*. Jakarta: Raja Grafindo Persada.