MAPPING READING QUESTIONS OF SUMMATIVE TEST FOR ELEVENTH GRADE OF SMK ST. NAHANSON PARAPAT SIPOHOLON

*Ivan Fernando Simanjuntak

**Marisi Debora, S.Pd., M.Hum.

ABSTRACT

This study aimed to analyze the distributions of the cognitive process dimensions of Revised Bloom's Taxonomy and the distribution of indicators on reading questions of summative test. The design of this research was a descriptive qualitative study with a survey method. This research was conducted in SMK St. Nahason Parapat Sipoholon on 2020/2021 academic year. The data of this research was the reading questions of summative tests used by the teacher in the odd semester examination. The source of the data was a summative test for eleventhgrade students in the 2020/2021 academic year. The data were analyzed using Revised Bloom's Taxonomy's operational verb dimension of the cognitive process and grouped according to basic competencies and indicators in the syllabus. The findings revealed that: (1) the distribution of the cognitive process dimensions of Revised Bloom's Taxonomy on the summative reading test are not evenly distributed, and the question is dominated by the dimensions of the level process cognitive remembering (C1) and understanding (C2). The average distribution of dimensions of the cognitive process of Revised Bloom's Taxonomy on reading test was the cognitive process dimensions level C1 of 12 questions out of 18 in total (56%), C2 consist of 4 questions (22 %), C3 was not included (0 %), C4 covered 4 questions (22 %), C5 by 0 %, and C6 by 0 %. (2) The distribution of indicators on the odd semester exam questions of reading test was not evenly divided. There are 26 indicators arranged by the English teacher for odd semester, only 2 indicators are included in the reading test items for total 9 questions out of 18 in total. Some indicators are not included in the questions, and there are 9 questions (50%) out of 18 in total were not related to the indicator.

Keywords: *Reading Questions of Summative Test, Cognitive Process Dimension of Revised Bloom's Taxonomy, Indicator*

*Graduate Status
**Lecturer Status

I. INTRODUCTION

A. Background of The Study

A test is a type of evaluation instrument that is used to assess the teaching and learning process. Tests are also utilized for a variety of purposes in the fields of vocational education and technology. Test results can be used to place students in the most appropriate classes or programs. Exams are an important aspect of the teaching and learning process if they are integrated into daily classroom teaching and designed to be a part of the cumulative learning process. According to Hughes, the test is designed to assess student accomplishment and the program's overall success (Pradanti et al, 2018).

According to Hughes, the test is also used to identify a student's strengths and shortcomings, as well as to determine what learning still needs to be done (Bernasela, 2014). Djiwandono states that there are two kinds of tests. The first is the Standardized test, and the second is the Teacher-Made test (Asyhar, 2017). An example of the Teacher-Made test is a summative test. Furthermore, Arikunto argues that summative test can be likened to semester exams which are usually carried out at the end of each semester (Asyhar, 2017). This is test that is made based on material and specific objectives created by the teacher for his own class. The summative test will be a source of information for the teacher to find out students' understanding of the material for one semester and also determines whether the student can continue his studies to the next class or not.

Based on the results of interviews that have been conducted with two English teachers, summative test was also being conducted in SMK St Nahanson Parapat. SMK St Nahanson Parapat is a private vocational school that implements the 2013 curriculum, where this school is known to have carried out the odd semester exams in December 2020. The 2020/2021 odd semester exam questions used are in the form of multiple choices and there are 25 test items compiled by the English teachers. In conducting the test, the

final exam questions that have been made are rarely analysed by the teacher. As for, the educators makes the questions by looking at textbooks, or student worksheets that have been made from several book publishers, or revised questions that have existed from years ago, so it is not yet certain whether the questions made are in accordance with the basic competencies and indicators contained in the syllabus.

The result of the English summative test at SMK St Nahanson Parapat Sipoholon was very good. According to the English teachers, there are 115 students in the eleventh grade. The results of the odd semester exam show that 95% of the students have passed the minimum score of 75 (KKM) required by the school. 20% of the students managed to score 80-85, 30% of them scored 86-90, and 45% are successfully completed the exam with scores above 90. In more detail, there are 20% of the students who managed to get perfect scores on the odd semester exams of 2020/2021 academic year. It means that all of the eleventh grader could solve the test.

The fact that the English Summative test in SMK St Nahanson Parapat are rarely analyse and was made by the teachers become the good consideration in conducted the research. Another consideration of the researcher to do the test analysis and mapping the reading questions was because reading questions are considered more difficult than other types of questions in English test, such as, grammar test, vocabulary test, and etc. Shape (2000) identified eight types of questions commonly used in reading tests. The 8 types are: previewing, reading for main ideas, using context for vocabulary, scanning for details, making inferences, identifying exceptions, finding references, and referring to the passage (Solikhah, 2015). Hence, when working on English questions such as in the Semester exams, UAS, UN, SBMPTN, TOEFL, IELTS, and so on, the reading test often makes the reader stuck or confusing. Apart from having to read long texts, the time was limited (Solihati and Suharto, 2020).

Based on the explanation above, this research focused on the English summative test in SMK St Nahanson Parapat. The analysis based on the cognitive domain of Revised Bloom's Taxonomy. The differences between this research with the previous study were the research design and the data and data sources. This research was descriptive qualitative research, and the data and data sources comes from vocational school, on the other hand, the previous study were commonly quantitative research and the data comes from Junior or High School students.

II.   REVIEW OF LITERATURE

A.   Theoretical of Framework

In supporting the idea of this study, the research presented some theories that helped the researcher. This section describes several points about the study which consists of Syllabus, Tests, Types of Tests, Reading Test, Summative Test, Item Analysis, and Revised Bloom's Taxonomy.

1.   Syllabus

The syllabus is defined as a summary of each lesson or the results of curricular translation into subject matter, learning activities, and assessments in general. A syllabus is a learning plan that comprises competency standards, basic competencies, topic/learning resources, learning activities, and assessments for a given subject or group of subjects/themes.

A syllabus is described as a learning plan for a topic that comprises core competencies, basic competencies, learning materials, learning activities, assessments, time allocation, and learning resources in Permendikbud Nomor 59 Tahun 2014 tentang kurikulum 2013. Based on the above definition, a syllabus is a learning plan that includes   Competency   Standards,   Basic   Competencies,

subject/learning materials, learning activities, indicators, and competency achievement for assessment, time allocation, and learning resources for a specific subject and/or group of subjects or themes. Simply put, a syllabus is a reference for creating a learning framework for each subject.

2. **Test**

Test is an assessment technique used to measure the ability of students to achieve certain competencies or learning objectives, through quantitative data analysis, which results in numbers. Based on that number, the level of mastery of student competence is interpreted (Sanjaya and Andi, 2017). For the needs of evaluating the teaching and learning process, standardized tests and teacher-made tests can be used.

The standardized test is a test that in its preparation takes a long time, in which to obtain a standard test through procedures such as compilation, testing, analysis, revision and editing. In addition, standardized tests are based on materials and general objectives of schools across the country so that they cover a wide range of aspects and knowledge or skills with only a few test items for each skill or topic. The standard test also has high reliability (Arikunto, 2017).

Meanwhile, the teacher-made test is a learning outcome test prepared by the teacher itself to measure and assess student learning outcomes, both on each presentation of one learning unit, as well as on formative and summative exams (Toijo, 2013). The quality of the test is also important because one of the functions of the test is to reveal students" ability in the material given (Asyar, 2014).

3. **Reading Test**

According to Syafiie (1999), reading is a process of developing knowledge of written discourse. This process occurs by matching or linking previously possessed schemata of information and experiences in order to build a comprehension of the discourse being read (Solikhah, 2015). Furthermore, Burns (1996) asserted that reading comprehension is divided into two categories: low reading comprehension and strong reading comprehension. The low reading group is also called literal reading. The high reading group is also called critical reading, and consists of three: interpretive understanding, critical understanding, and creative understanding.

Shape (2000) identified eight types of questions commonly used in reading tests. The 8 types are: previewing, reading for main ideas, using context for vocabulary, scanning for details, making inferences, identifying exceptions, finding references, and referring to the passage (Solikhah, 2015). In the question of the meaning of vocabulary that refers to context, the answer can be found by looking at the word, phrase, or sentence that is around the word being asked for its meaning. That is what is called context, which is a word or sentence in the reading. Context serves to guess meaning. Generally, if the reader knows the general words of a sentence, the two will also know the meaning of the general meaning in the sentence. So, guessing the meaning through context can be used to guess the meaning of certain vocabulary in question.

4. **Summative Test**

The summative test has been covered in the preceding sub-chapters. According to Gronlund, summative tests are often administered at the end of a course. It is intended to assess the extent to which instructional goals have been met. Furthermore, the exam

may be used to establish course scores in order to confirm that students have mastered the necessary learning goals (Kristiana, 2014). According to Sudjiono, summative tests are designed to assess students' abilities after they have completed all subject content. It is given at the conclusion of a course of study (Suryani, 2017).

Summative assessment, which normally occurs at the end of the course, tries to measure what students have learned. A course's final test is an example of summative evaluation. It summarizes what the pupil has learned, implying a glance back to see how successfully the student met the aim (Yohana, 2009). To create a successful summative exam, the test creator must consider the test items that are most suited for evaluating student progress, are aligned with the learning objectives, and are as reliable as feasible. Teachers can use the results of learning outcomes assessments to determine how well students have grasped the subject matter's material or the amount to which they have met learning objectives, and teachers can motivate students to improve their learning abilities (Kristiana, 2014).

5. **Item Analysis**

Item analysis is an important activity in the preparation of questions so that quality items are obtained (Kusaeri, 2012). Aieken states, Item analysis aims to improve the quality of the test items and find out student diagnostic information. Quality questions are questions that can provide information as precisely as possible, so that it can be seen whether students have mastered the material or not (Kurniawan, 2015). Furthermore, in carrying out the item analysis, the question writers can analyze qualitatively, in relation to its content and form and quantitatively in relation to its statistical characteristics or judgment improvement procedures and empirical improvement procedures (Kurniawan, et al, 2017).

The qualitative analysis that is carried out before the questions are used is in the form of a study which is intended to analyze the questions in terms of material, construction and language. Material analysis is in the form of studies related to the scientific substance being asked in the questions (the suitability of the questions with the indicators or the suitability of the content of the material with the level, type of school and grade level). The construction analysis is in the form of a review which is generally related to the question writing technique. Language analysis is in the form of examining questions related to the proper and correct use of Indonesian (Giani et al, 2015).

6. **Revised Bloom's Taxonomy**

Anderson et al. altered Bloom's Taxonomy in 2001 to a two-dimensional Bloom's Taxonomy with a cognitive process dimension and a knowledge dimension. The cognitive process dimension is made up of six verbal levels: remembering, comprehending, applying, analyzing, evaluating, and creating. These six stages, designated as C1 through C6, are frequently utilized in the formulation of learning objectives (Effendi, 2017). In the knowledge dimension, there are four levels represented by nouns: factual knowledge, conceptual knowledge, procedural knowledge, and metacognitive knowledge.

In detail, the levels in the Two-Dimensional Bloom's Taxonomy can be explained as follows: first, the dimensions of cognitive processes. The levels in the cognitive process dimension are intended to provide a comprehensive package of classifying students' cognitive processes that enter into learning objectives and then become a reference in making evaluation tools.

III. **RESEARCH METHODOLOGY**

This research was designed as a descriptive qualitative study using a survey approach. According to Sugiyono (2012: 29), a descriptive approach is one that is utilized to describe or analyse a study outcome but not to draw broad generalizations (Amelia, 2015). Furthermore, Tika (1997: 9) states specifically that "a survey is a research approach that tries to gather vast volumes of data in the form of variables, units, or persons at the same time; data is acquired through individuals or specified physical samples to generalize to what is investigated" (Gunawan, 2012).

The data of this research were the reading questions of summative tests used by the teacher in the odd semester examination and the interviewed with the English teacher. The source of the data were a summative test for eleventh-grade students in the 2020/2021 academic year of SMK St Nahanson Parapat Sipoholon and the interviewed via chat WhatsApp with the English teacher.

In this study, WhatsApp was used to interview the English instructor and collect preliminary data, as well as the Taxonomy Bloom Cognitive Aspects Analysis Checklist, which is as follows:

Table 3.1. Checklist of Cognitive Aspects Analysis of Revised Bloom's Taxonomy and suitability for Basic Competencies and Indicators.

No	Test Items	Basic Competencies	Indicator	Cognitive Level					
				C1	C2	C3	C4	C5	C6
1									
2									
Etc.									
25									
Total									
Percentages (%)									

IV.   RESEARCH FINDINGS AND DISCUSSION

A. Research Findings

The following are the findings from an analysis of the distribution of the dimensions of the cognitive process of Revised Bloom's Taxonomy, as well as the conformity with the Basic competencies and indicators of reading test on the odd semester final exam questions for English of eleventh-grade students at SMK St. Nahanson Parapat:

(1) The Distribution of Cognitive Process Dimension of Revised Bloom's Taxonomy on Summative Test Questions.

In the 2020/2021 academic year, the distribution of the level of the cognitive process dimension of Revised Bloom's Taxonomy on the summative test questions for eleventh-grade students of SMK St Nahanson Parapat is not evenly distributed

Table 4.1. The Distribution of Cognitive Process Dimension on Summative Test Questions.

The Distributions of The Cognitive Process Dimension							
Cognitive Level	LOT			HOT			TOTAL
	C1	C2	C3	C4	C5	C6	
Total number of the reading test items	10	4	-	4	-	-	18
The percentage of the reading test items	56%	22%	-	22%	-	-	100%

The cognitive process dimensions are shown in Table 4.4. The level of Revised Bloom's Taxonomy, which is widely used in final exam questions for English in eleventh-grade, namely the understanding level (C2) consists of 4 questions (22%), the remembering level (C1) consists of 10 questions (56 %), the level of applying (C3) did not exist in the reading questions, the level of

analyzing (C4) consists of 4 questions (22%), and the level of creating (C6) and Evaluating (C5) consists of no questions. The proportion of the cognitive process dimensions of Revised Bloom's Taxonomy for eleventh grade questions is 78% easy questions, 22% moderate questions and no difficult questions are found.

(2) The Distribution of the Indicators on Reading Questions of Summative Test.

There are total 24 Indicators arranged by the teacher in the RPP for eleventh-grade students in the odd semester (Appendix A). The results obtained from the analysis of the distribution of Basic competencies and Indicators on the summative test questions for eleventh-grade, there were only 2 indicators applied in the test items.

Table 4.2. The Distribution of the Indicators on Reading Questions of Summative Test for Eleventh-grade.

The Distributions of the Test Items Relating to the Indicator in Reading Test				
NO	Indicators	The question's Number	The Total of the Test Items Relating to the Indicator	The Percentage of the Test Items Relating to the Indicator
1	3.5.2 Mengidentifikasi isi, struktur teks, dan unsur kebahasaan dalam surat pribadi.	6, 7, 8, 9, 12, 13	6	33%
2	3.5.3 Menjelaskan kegunaan teks khusus dalam bentuk surat pribadi	5, 10, 11	3	17%
TOTAL	2	-	9	50%

The Test Item Distributions Unrelated to the Indicator in the Reading Test				
NO	TOPIC	The question's Number	The Total of the Unrelated Test Items	The Percentage of the Unrelated Test Items
1	Procedure Text	14, 15, 16, 17, 18, 19, 20, 21, 22.	9	50%

Table 4.5. indicates that the indicator 3.5.2 "Mengidentifikasi isi, struktur teks, dan unsur kebahasaan dalam surat pribadi" is the most commonly used with 6 questions. Followed by indicator 3.5.3 "Menjelaskan kegunaan teks khusus dalam bentuk surat pribadi" with total of 3 questions. Only one basic competency (3.5) is used in the conducted 18 questions of reading test of summative test for the odd semester, out of a total of five basic competencies. On the other hand, there 26 indicators in the syllabus for English subject of eleventh grade in the first semester. Only 2 (8%) of the available indicators were used (out of a total of 26). There were 9 (50%) unrelated test items to the indicators which covered about procedure text. This topic was not related to the indicators arranged by the English teacher.

B. Discussion

The cognitive domain is one in which there is more mental activity or cognitive function. In the Revised Bloom's Taxonomy cognitive domain, there are six levels of thinking processes, going from low to high: remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). After analyzing the data in the research findings, it can be stated that the questions utilized as evaluation instruments were inadequate. This is due to the unequal distribution of Revised Bloom's Taxonomy's cognitive process

dimensions at each level. Questions are dominated at the level of remembering (C1) and level of understanding (C2), which indicates the questions are focused on rote topics/materials. The reading questions of summative test for eleventh-grade students of SMK St Nahanson Parapat only reached the analysis level (C4), and no applying (C3) or evaluation level (C5) nor created level (C6) was identified, according to the distribution of Revised Bloom's Taxonomy's cognitive process dimensions.

After analyzing the results, it was discovered that the lack of Evaluating (C5), and Creating (C6) level questions in the summative test corresponded with the data gathered from the researcher's interview with the English teachers, which summarized as, when administering the test, the final exam questions that were created are rarely analyzed by the teachers. They develop the questions by looking through textbooks, student worksheets created by various book publishers, and revised questions from previous years. It is also known that, the reason why there was no Applying (C3) level for reading summative test because Applying (C3) was actually related to the productive skills in English, which is Writing and Speaking skill. It understandable that the reading question did not covered the level of Applying (C3). Applying includes using a procedure to solve a problem or do a task. Therefore, applying is closely related to procedural knowledge. Furthermore, This category includes two kinds of cognitive processes: executing and implementing (Widodo, 2006).

Indicators are measurable changes in behavior, such as attitudes, knowledge, and skills, indicate that Basic Competencies has been achieved (Delafini et al, 2014). The distribution of indicators on the odd semester English final exam questions of eleventh-grade is categorized as inadequate based on the findings of the analysis. In the eleventh grade, there are 26 indicators. There are 9 (50%) indicators-related questions on

the final exam. Furthermore, 9 questions (50%) do not refer to indicators out of a total of 18 reading questions. The odd semester exam is a form of assessment used by educators to assess student competency at the end of the semester. The indicators that indicate all basic capabilities are included in the scope of the questions tested.

The findings reveal that the distribution of Revised Bloom's Taxonomy dimensions of cognitive process, as well as the distribution of Basic Competencies and Indicators, does not match what they should be. There are questions in the exam questions for which the basic competencies and indicators are unknown; this can occur because the exam questions were not prepared in a sequential procedure of preparing questions. If the questions on the odd semester final exam are arranged according to established methods, the questions will accurately represent the curriculum content. Furthermore, the proportion of cognitive process dimensions at each level will be evenly distributed, because the questions are in the form of indicators containing operational verbs, allowing the number of questions to be adjusted according to the proportions of Revised Bloom's Taxonomy's cognitive process dimensions at each level throughout question preparation.

V.   CONCLUSION AND SUGGESTION

A. Conclusion

The findings of the study lead to the following conclusions:

(1) The distribution of each level of Revised Bloom's Taxonomy cognitive process dimension on the odd semester final exam questions of reading test for English in SMK St Nahanson Parapat Sipoholon in 2020/2021 academic year, namely at the cognitive process dimensions level C1 of 12 questions out of 18 in total (56%), C2 consist of 4 questions (22 %), C3 was not included (0 %), C4 covered 4 questions (22 %), C5 by 0 %, and C6 by 0 %.

(2) The distribution of indicators on the odd semester exam questions of reading test for English in SMK St Nahanson Parapat Sipoholon for eleventh-grade of the 2020/2021 academic year is similarly not evenly divided. There are 26 indicators arranged by the English teacher for odd semester, only 2 indicators are included in the reading test items for total 9 questions out of 18 in total. Some indicators are not included in the questions, and there are 9 questions (50%) out of 18 in total were not related to the indicator.

B. Suggestions

Derived from the previous conclusions, the following suggestions as a follow-up in this study:

(1) To teachers, teachers are advised to pay more attention to the rules of writing questions when compiling the questions, as well as the distribution of the proportion of exam questions based on the Revised Bloom's Taxonomy so that the questions are better and refer to the HOTS (Higher Order Thinking Skills) questions.

(2) To further researchers, when conducting additional research, include other indicators to broaden and deepen research on final exam questions.

REFERENCES

Amelia, D., Susanto, Arif, F., (2015), Analisis Hasil Belajar Matematika Siswa pada Pokok Bahasan Himpunan Berdasarkan Ranah Kognitif Taksonomi Blom Kelas VII-A di SMPN 14 Jember, *Jurnal Edukasi* UNEJ, *2(1)*. */bitstream/123456789/24330/1/Dira%20Mustarah.pdf,* Diakses pada 4 Januari 2018. *8(6),* 141-146.

Arifin, Z. (2011), *Penelitian Pendidikan*, PT Remaja Rosdakarya, Bandung.

Arikunto, (2017), *Dasar-Dasar Evaluasi Pendidikan*, Bumi Aksara, Jakarta.

Arikunto, S., (2002), *Dasar-Dasar Evaluasi Pendidikan Edisi Revisis Cetakan Ke 3*, Jakarta : Penerbit Bumi Aksara.

Asyhar, S. (2017). *Item Difficulty on English Summative Test of the 11th Grade Students in MAN 3 Malang* (Thesis). Universitas Brawijaya. Retrived from http://repository.ub.ac.id/695/.

Bernasela. (2014). An Analysis on English Summative Test Items. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, *3 (3)*.

Delafini, R., Holilulloh., dan Yunisca, N., (2014), Pengaruh Kemampuan Guru Dalam Mengembangkan Indikator Pencapaian Kompetensi Terhadap Kesiapan Guru Dalam Mengajar, *Jurnal Pelita*, 1(1) :1.

Desrijenatin, S., (2020), Analisis Soal Ujian Akhir Semester Ganjil Mata Pelajaran Sejarah Indonesia Kelas XI Tahun Pelajaran 2019/2020 di Kota Padang, *Jurnal Kronologi*, 2(2) : 67.

Effendi, R., (2017), Konsep Revisi Taksonomi Bloom dan Implementasinya pada Pelajaran Matematika SMP, *Jurnal Ilmah Pendidikan Matematika*, *2(1)*: 74-76. *Eğitim Fakültesi Dergisi, 20 (2)*: 399-426.

Fajriani, N. (2018). *Pemetaan Kualitas Soal Uas Bahasa Indonesia Yang Dikembangkan Oleh Guru Bahasa Indonesia Alumni Pendidikan Bahasa Dan Sastra Indonesia Fkip Universitas Mataram Di Sma Se-Kabupaten Lombok Barat* (Thesis). Universitas Mataram. Retrieved from https://ejournal.unib.ac.id/index.php/jurnaldiksa/article/view/3460.

Giani., Zulkardi., dan Hiltrimartin., (2015). *Analisis Tingkat Kognitif Soal-Soal Buku Teks Matematika Kelas VII Berdasarkan Taksonomi Bloom*, Palembang, Fakulltas Keguruan dan Ilmu Pendidikan Universitas NegeriSriwijaya.

Gronlund, N.E., 1977. Constructing Achievement Test, 2nd Edition. New York: Prentice Hall, Inc. University of Illinois.

Gunawan, I., dan Anggarini, R. P., (2012), Taksonomi Bloom-Revisi Ranah Kognitif: Kerangka Landasan untuk Pembelajaran, Pengajaran dan Penilaian, *Jurnal Pendidikan Dasar dan Pembelajaran*, *2(2)*: 98-117.

Hamzah, A., (2014), *Evaluasi Pembelajaran Matematika*, Jakarta : Rajawali Pers,. Heaton, J.B. 1975. Writing English Language Test. England: Longman Group Limited.

Hughes, A. 2003. Testing for language teachers, second edition. Cambridge: Cambrige University Press.

Khaerudin, (2017), Administrasi, Analisis Butir dan Kaidah Penulisan Tes, *Jurnal Madaniyah*, 1 Edisi XII : 107.

Khodijah, N., (2014), *Psikologi Pendidikan*, Jakarta : Rajawali Pers.

Kristiana. (2014). *An Analysis on the Content Validity of Summative Test For the Second Grade Students of Junior High School (A Case Study Of Smpt Pp Daarul Amanah Kec Cipondoh- Kota Tangerang)* (Thesis). Syarif Hidayatullah State Islamic University. Retrieved from http://repository.uinjkt.ac.id/dspace/handle/123456789/25380.

Kurniawan, R., Albrian, F., Luqman, H., Retno, M., dan Irin, W., (2017), Pemberian Pelatihan Analisis Butir Soal Bagi Guru di Kabupaten Juombang : Efektif?, *Jurnal of Elementary Education*, 4(1) : 2.

Kurniawan, T., (2015), Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar, *Journal of Elementary Education,* 4(1): 2.

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

Marsiyah, F., (2016), *Analisis Soal Ulangan Harian Buatan Guru Biologi SMA Muhammadiyah 1 Surakarta Tahun Ajaran 2015/2016 Ditinjau dari Tingkat Taksonomi Bloom*, Publikasi Ilmiah.

Mustarah, D., (2013), Analisis Ulangan Akhir Semester (UAS) Biologi Kelas X Ditinjau dari Taksonomi Bloom, *http://repository.uinjkt.ac.id/dspace*

Oktavianti, F., (2017), Analisis Soal Ujian Akhir Semester Genap Tahun Ajaran 2016/2017 Berdasarkan Taksonomi Bloom Revisi Kelas X dan XI Mata Pelajaran Biologi di SMA Negeri 3 Tanjungpinang, *E-Journal FKIP Universitas Maritim Raja Ali Haji*: 13.

Pradanti, I. S., Martono., Sarosa, T., (2018), An Item Analysis of English Summative Test for The First Semester of The Third Grade Junior High School Students in Surakarta. *English Education Journal*, *6(3)*: 312-318.

Prawira, P. A., (2016), *Psikologi Pendidikan dalam Perspektif Baru*, Yogyakarta : Ar-Ruzz Media.

Purwanto, E., (2014), *Evaluasi Proses dan Hasil dalam Pembelajaran*, Penerbit Ombak, Yogyakarta.

Puspitasari, Tita. (2016).The Content Validity of the Summative test in the Second Semester of the First Year Students (a Case Study of SMAN 1 Cikarang Barat, Bekasi). *PUJANGGA, 2 (1)*: 199-224.

Putri, I. S., (2018), Analisis Kesesuaian Butir Soal Buatan Guru Dengan RPP Pada Materi Protista Dan Keanekaragaman Hayati di SMAN 1 Trumon Tengah Aceh Selatan, *Jurnal Pendidikan*, 1(1) : 56.

Raafi, W. O. N., dan La, N., (2-15), Kualitas Tes Buatan Guru pada Soal Pilihan Ganda Mata Pelajaran Matematika Kelas VIII SMP Negeri 4 Kendari Semester Genap Tahun Pembelajaran 2013/2014, *Jurnal Penelitian Pendidikan Matematika*, 3(1): 170.

Rusyna, A., (2014), *Keterampilan Berpikir: Pedoman Praktis Para Peneliti Keterampilan Berpikir,* Yogyakarta : Penerbit Ombak.

Sanjaya, W., dan Andi, B., (2017), Paradigma Baru Mengajar Edisi Pertama, Jakarta : Kencana.

Semiun, Trivict, T., Luruk, Densiana, F. (2020). The quality of an English Summative Test of a Public Junior High School, Kupang-NTT. *English Language Teaching Educational Journal*, *3(2)*: 133-141.

Septiyana, R. (2016). Analysis of Summative Tests for English. *English Education Journal (EEJ)*, *7(4)*: 433-447.

Solihati, A. T., Suharto, P. P., Mengenali Soal-Soal Reading Comprehension Berekuivalensi Toefl Dan Strategi Mengerjakannya, *IKRAITH-ABDINAS,3(1),* 75-81

Solikhah, I., Pengembangan Tes Reading for Academic Purposes Untuk Program Eap Di Iain Surakarta, *CENDEKIA*, *9 (2)*: 177-194

Sudiyono, A., (2008), *Pengantar Evaluasi Pendidikan*, Jakarta : PT Raja Grafinfo Persada.

Suryani, E.Y. (2017). Pemetaan Kualitas Empirik Soal Ujian Akhir Semester Pada Mata Pelajaran Bahasa Indonesia Sma Di Kabupaten Klaten. *Jurnal Penelitian dan Evaluasi Pendidikan*, *21 (2)*: 142-152.

Thoha, M. C., (1996), *Teknik Evaluasi Pendidikan*, Jakarta : Raja Grafindo Persada.

Toijo, A. K., (2013), Kualitas Tes Buatan Guru, *Jurnal Normalita*, 1(1): 134

Wahyuni, S., (2017), Development Test System Based On Linear Equations Two Variable Resived Taxonomy Bloom to Measure High Order Thinking Skills at Student Class VIII SMPN Sungguminasa Gowa,*Jurnal Daya Matematus, 5(1)*, 134-135.

Widodo, A., (2006), Revisis Taksonomi Bloom dan Pengembangan Butir Soal Buletin Puspendik, *3(2)*: 5-10.

Yulianti, (2016), Pengembangan Alat Evaluasi Hasil Belajar Mata Pelajaran Pendidikan Agama Islam Berbasis Taksonomi Bloom Dua Dimensi, *Jurnal of Islamic Education Studies*, 1(2): 408-416.