Development of Authentic Assessment Instrument for Improving Extensive Capacity of Metacognitive Perspective

Irfai Fathurohman[1], Riyan Dwi Cahyaningsih[2]

[1,2]Program Studi Pendidikan Bahasa dan Sastra Indonesia Universitas Muria Kudus, Indonesia
*Corresponding author: Irfai.fathurohman@umk.ac.id[1]

Abstract

The purpose of this study is to find out the authentic assessment instrument factors in the ability to read English extensively from a metacognitive perspective. The study was designed with an instrument of developmental research design. The developmental research design refers to the research development proposed by Mardapi. The modification model of the Mardapi instrument development was carried out through three stages, i.e., preliminary stages, development stage and evaluation stage. The results of this study through exploratory factor analysis showed that there were three factors derived from 24 statement items. The loading factor value for each item in each factor is > 0.3 which means that the whole item is valid. Based on the Alpha Cronbach reliability test it was obtained a value of 0.954, the instrument was declared reliable. The conclusion shows an authentic assessment instrument of extensive reading ability in English through metacognitive strategies for PBSI FKIP UMK students was proven to be valid, reliable and feasible to use. Authentic assessment instruments can improve the English reading skills of PBSI FKIP UMK students through a learning process that puts forward an approach according to the characteristics of each student.

Keywords: Authentic, Factors, Instruments, Ability, Reading, Metacognitive Assessment

INTRODUCTION

Assessment in a learning process is important and is the basis for students to get the results of the categories or scores learned. The ideal assessment is done according to the ability of each individual which can be deferred to as authentic assessment. Authentic assessment may be described as an assessment conducted to determine the mastery that is owned by each individual as long as he/she gets the learning. The characteristics of each individual are different, both in ability and cognitive mastery and there are some reasons for using authentic assessment one of which is the accuracy of using authentic assessment. Portrait or description of the learning process that is assessed using authentic assessments according to the conditions in the field, so that the data obtained is in accordance with the facts in the field and the results of research are closely related. As Arikunto (2016) said that to judge is to take a decision on something with good and bad measures.

It was revealed that authentic assessment is a comprehensive assessment carried out to assess starting from the input, process, and output of learning (Prasetyo, 2016). Wulan (in Prasetyo, 2016, p.1183-1184), proposes several advantages and disadvantages of authentic assessment. The advantages of this assessment are (a) students can demonstrate a process, (2) the process that is demonstrated can be observed directly, (c) it provides a more complete and natural evaluation of some kinds of reasoning, oral abilities and physical skills, (d) there is an agreement between the teacher and students about the assessment criteria and tasks to be done, (e) it enables the process of assessing learning outcomes and complex skills outside the application of the concept, (f) it gives great motivation to students, (g) it

encourages learning applications in real life situations and simulates real-world task scientists, and (i) it emphasizes many correct answers and creative solutions. While the weak points of the assessment includes among others (a) the very demanding time and effort, (b) the judgment and scoring which are more subjective, (c) more burdensome to the teacher, (d) having a tendency of low reliability so it requires the development of clear criteria (standard) that really shows competence and (e) students who feel that they have more ability in terms of concepts (remembering) usually feel intimidated. Yusuf (2017, p.292) revealed that the emergence of the concept of authentic assessment in the early 1990s, was as a form of experts' dissatisfaction with the weaknesses of objective tests, especially to assess high-level cognitive abilities in doing something in real life or real word settings.

The problems that have occurred during learning such as teaching abilities, practices, and performance which certainly have to be seen from the difficulties experienced by each individual, whether their abilities always increase, equal, or decrease. Learning process varies according to the conditions and situations experienced by learners. There are problems that have still been unsolved. Megawanti (2015, p.228) explained that problems of education in Indonesia emerges every year. Problems also arise starting from the level of input, process, and output. These three levels are actually interrelated and interconnected to each other. Inputs affect sustainability in the learning process. The learning process also influences the output or results. Furthermore, the output will return to the input in the level of higher education or enter the workforce, where the theory begins to be practiced. The ability to think of each student in the complexity of a problem is certainly a step in improving the ability of each individual student (samsyiah, 2019).

Ability assessment as it is in learning English is assessing the ability to read which is an important thing that must be assessed in stages and in accordance with the abilities controlled by the learner. The level of understanding of reading ability is seen from the ability to read according to spelling, understanding the meaning and meaning of reading, intonation that is appropriate and the ability to express words according to situations and conditions. In this study, the situation and condition of the learners seem to have difficulty. The situation of students who are not within the scope of the English study program and are within the scope of the Indonesian Language and Literature Education study program is one of the difficulties or interests of students in learning English. A lack of self confidence hinders students' progress in learning English. Reading according to Haryadi (in Widianto and Subyantoro, 2015, p.2) is an interaction between the reader and the writer. The interaction is not direct, but it is communicative. Mushon (in Aziz, Kustiono, and Lestari, 2019, p.100) revealed that education is an inseparable part in the process of human maturity and has a major contribution to the development of science and technology. Santi, Setiawan, Khumaedi (2019, p.117) explained that evaluations in learning are important to be carried out by educators but must be carried out with full calculation because they have an impact on the mentality of students and the learning programs implemented.

The ability to read in English is divided into two, namely extensive reading and intensive reading. Both of these skills support one another. Based on the results of observations conducted on PBSI students (Indonesian Language and Literature Education Study Program) of Muria Kudus University, it was obtained that students have difficulty in reading English text, i.e., extensive reading and intensive reading. This difficulty is in general caused by being less accustomed to reading English texts. English courses are taught for four semesters at PBSI UMK. Within a period of four semesters students will apply their abilities to understand and use the ability to read and to analyze ancient texts that require a deep theory to study, one of which is derived from English-language texts.

The need for authentic assessment in solving the difficulty of learning the ability of English texts is the right solution to do in this study. Authentic assessment can be used to determine the level of learners' abilities and learner's shortcomings. However, the use of authentic assessment alone in learners is not enough, an appropriate learning strategy is also needed, and is able to improve the development of learners in mastering the ability to read English texts. Metacognitive strategies are appropriate for use in developing English reading skills. Metacognitive strategy is a strategy used in learning ranging from selection or planning, stages of the learning process, identification, reflection, and provides appropriate solutions to overcome the problems faced. Maulana (in Namira, 2014, p.1272) revealed that learning with metacognitive strategies is learning that instills awareness about how to design, monitor and control what they know, what is needed to do it and how to do it. Tuysuz (in Widiyana, Rusilowati, Priyanto, 2019, p.91) states that educators need sensitivity in misunderstandings that occur to students

through effective learning processes.

Based on the problems that have been described in this article the researcher focuses on conducting research on "Development of Authentic Assessment Factor Instruments in Increasing the Metacognitive Perspective Extensive Reading Ability". The purpose of this study is to find out the authentic instrument assessment factors in the ability to read extensive English metacognitive perspectives.

METHODOLOGY

This research was designed with an instrument of developmental research design. The developmental research design used refers to the research development proposed by Mardapi. Authentic assessment factor instruments for improve the ability to read extensive metacognitive perspective. In the preliminary stage, the source of data in this study was obtained from preliminary observations made in the study program of Indonesian Language and Literature Education at Muria Kudus University. The next activity is a document study which aims to see the results of an authentic assessment of extensive reading skills in English through metacognitive strategies. The next stage is development. At the stage of development, the data sources in this study were obtained from the results of validation by experts and the data from field trials. The subjects of this study were students of the Indonesian Language and Literature Education study program at Muria Kudus University. In this study, to calculate the contents of an instrument based on expert judgment using the Aiken's V Formula. In the two way anava test, researchers used an analysis of variance proposed by Hoyt, item scores in this case were considered to be factorial designs of the two paths known as treatment X subject design (Mardapi, 2017, p.78). The Hoyt formula is as in (1), where s2r is residual variant which in the x subject treatment analysis is the mean squared interaction between items and subject MKs and s2s is subject variants are squares between subjects, namely MKs Reliability.

$$\mathrm{rxx} = 1 - \frac{s2r}{s2s} \tag{1}$$

RESULTS AND DISCUSSION

Results and discussion in this study served in an authentic instruments in the ability to read English incentives metacognitive perspective. Analysis via SPSS 16.0, the results of analysis of the total variance factor analysis.

Table 1. Results of Analysis of Total Variance Explained Factor Analysis

Component	Initial eigen values			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of Variance	Cumulative %
1	11,882	49,509	49,509	11,882	49,509	49,509	5,911	24,628	24,628
2	1,494	6,226	55,735	1,494	6,226	55,735	4,551	18,962	43,590
3	1,201	5,005	60,740	1,201	5,005	60,740	4,116	17,150	60,740

The deciding factors in authentic assessment are understanding the content of the text, understanding the detailed contents of the text, the smoothness of the expression, the accuracy of the diction, the accuracy of the sentence structure, the meaning of the speech. Triyono, Sunarto, and Lestari (2019, p.66) revealed that test assessments were used to test students' cognitive abilities. In this study

the factors used in the study, are the results of authentic assessment instruments used to measure the English reading ability of Indonesian Language and Literature Education Study Program students, the Teaching and Education Faculty of Muria Kudus University.

These factors are an inseparable part of this research and are closely interrelated. This study contained 24 instruments used for research. Through exploratory factor analysis with the help of the SPSS program there are 3 factors that are formed and can represent the number of indicators. Based on the 24 items analyzed, it was found that 3 factors had initial eigenvalues above 1. It means that the 24 items could be grouped into 3 groups of factors. Factor 1 has a value of 11,882 and is able to explain the variance of 49,509%, factor 2 has a value of 1,494 and is able to explain a variance of 6,226%, factor 3 has a value of 1,201 and is able to explain a variance of 5,005%. Therefore, the three factors are able to explain the variance of 60.740%. Overall results of the Total Variance Explained are illustrated in the following 4.1 Scree plot.
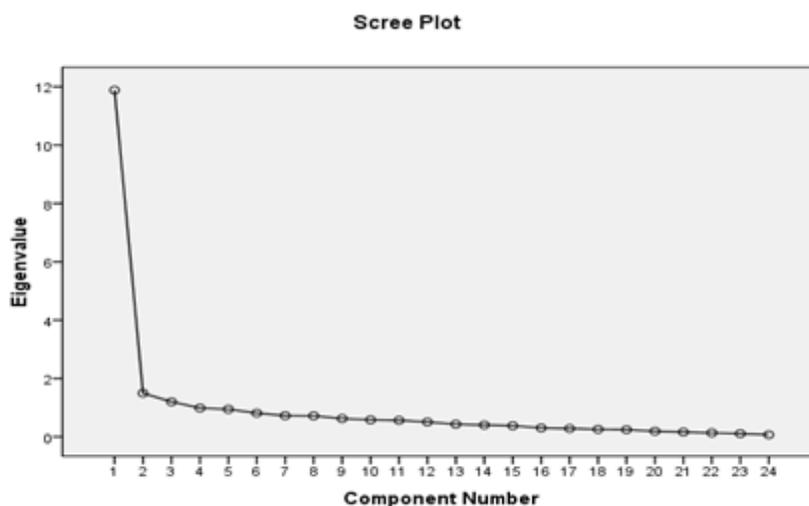
Figure 1. Scatterplot of the Total Variance Explained results.

Based on the Scree Plot, it can be seen that from factor 1 both factors 2 declined quite sharply, then from factor 2 to factor 3 it was still decreased, while factor 4 to factor 24 was below 1. Furthermore, the determination of each item would be included in the factor which of the three factors is present. Grouping of items and the size of the loading factor of one factor can be seen from the value of the loading factor which has a value of > 0.3. Grouping items into factors can be done by looking at table 2 about the Rotation Component Matrix.

Based on Table 2, there are 3 factors that formed the formation of a factor based on the existence of a strong relationship between the items with factors that can be seen through the value of the loading factor. The following is the grouping of each item into the factors formed. In factor 1, there are several items that have the highest loading factor in factor 1 and low in other factors, namely items. Ability to use synonymous vocabulary with meaning and derivatives related to text content (0.793), Ability to use sentences with subjects subject to work by linking text content in them (0.703), Ability to show passive sentence patterns in text (0.680), and Ability to show a problem related to phenomena / current situation with in-text content (0.278). The validity value of the factor 1 construct is 0.614.

In factor 2 the items that have the highest loading factor are factor 2 and the value of the loading factor is low in other factors ability to communicate using body language (0.753), ability to use vocabulary appropriately according to the contents of the text (0.703), ability to express problems in the text using appropriate intonation, pronunciation, and pauses (0.644), and ability to communicate in language that is easy to understand (0.014). The construct validity value of factor 2 is 0.528.

In factor 3 items that have the highest loading factor with a factor of 3 and a low loading factor value in other factors, i.e. Ability to use language that is easy to understand and accordance with the contents of the text (0.751), Ability to use title, introduction, and content accurately in the text (0.643), Ability to communicate in submitting reports / documentation (0.604), and Ability to use synonymous

vocabulary with intents and derivatives related to text content (-0,030). The construct validity value of factor 3 is 0.492.

Table 2. Results of the Rotation Component Matrix

Rotated Compomenent Matrix[1]			
Aspects Assessed	**1**	**2**	**3**
Ability to communicate good and correct language	0,303	0,474	0,434
Ability to communicate a language that easy to implement	0,664	**0,014**	0,532
Ability to communicate in submitting report or documentation	0,614	0,047	**0,604**
Polite communication skills and courteous with other people and groups	0,347	0,521	0,404
Body language communication skills	0,109	**0,753**	0,075
Ability to use easy language according to text	0,641	0,274	0,312
Ability to use headings, intro, and content accurately in text	0,332	0,405	**0,643**
Ability to use easy language according to text	0,50	0,133	**0,751**
Ability to show differences, similarities, and strong relationship, when commenting on problems with text content	0,352	0,188	0,590
Ability to express problem in the text using correct intonation, pronunciation, and pauses	0,044	**0,644**	0,488
Ability to use correct pronunciation pronunciation and combining the appropriate disclosures in the contents of the text	0,380	0,466	0,367
Ability to use pauses according to the contents of disclosure in the text	0,307	0,593	0,404
Ability indicate a problem related phenomena / current situation with the content in text	**0,278**	0,452	0,564
Ability to relate problems in the text with environment in which to study	0,418	0,475	0,460
Ability to utilize vocabulary appropriately according to text content	0,314	**0,703**	0,040
Ability to use the right choice of words according to situation and content in text	0,656	0,295	0,365
ability to use synonymous vocabulary with intents and utterances related to the content of text	**0,793**	0,111	-**0,030**
Ability to show different word variations with proper explanation according to text content.	0,642	0,386	0,259
Ability to use sentences in which the subject takes action according the relationship with the text	0,407	0,476	0,318
Ability shows the pattern of active sentence in the text	0,566	0,517	0,100
Ability to use sentences with subject to work by linking content of the text	**0,703**	0,407	0,283
Ability to show passive sentence pattern in text	**0,680**	0,256	0,202
Ability to demonstrate mandate or messages that are in the text	0,558	0,269	0,268
Ability to show positive meaning word, vocabulary or sentences	0,637	0,491	0,340

The Results of Instrument Item Reliability Test

After conducting the construct validity results, the next step is to conduct a reliability test. The reliability test of the authentic assessment instrument for reading English through metacognitive strategies in this study was carried out using the Alpha Cronbach reliability test through the SPSS 16.0 program. The following are the results of the reliability test using SPSS 16.0 contained in Table 3.

Table 3. Results of Field Trial Reliability Tests - Reliability Statistics

Reliability Statistics	
Cronbach's Alpha	**N of Items**
.954	24

(Source: results of SPSS 16.0 data processing)

The results obtained from the reliability test was 0.954, which means that an authentic instrument of English reading ability through metacognitive strategies is consistently used in making assessments.

After knowing the maximum factor components formed are 3 factors, the next step to determine each indicator/item will be included in these 3 factors. Component Matrix results show that there are 3 factors that are formed is the most optimal amount. Component Matrix shows the distribution of the 24 items on 3 factors that are formed. A factor will be formed if the value is >0.30. To see the correlation between items with the factor to be formed is determined by looking at the value of loading factors on each item 1 to factor 3 which is seen in the table above. Item of ability to communicate using body language, the correlation between these variables with a factor of 1 is 0.524 (strong enough, while the correlation with a factor of 2 is 0.418 (strong enough (the sign-only shows the direction of the correlation). Communicating using body language and the rotation process needs to be done. The followings are the results of the Rotation Component Matrix shown in Table 4.

Table 4. Results of the Rotation Component Matrix

Result Rotated Component Matrix			
	1	**2**	**3**
Ability to communicate good and correct language	.303	**.474**	.434
Ability to communicate a language that easy to implement	**.664**	.014	.532
Ability to communicate in submitting report or documentation	**.614**	.047	.604
Polite communication skills and courteous with other people and groups	.347	**.521**	.404
Body language communication skills	.109	**.753**	.075
Ability to use easy language according to text	**.641**	.274	.312
Ability to use headings, intro, and content accurately in text	.332	.405	**.643**
Ability to use easy language according to text	.050	.133	**.751**
Ability to show differences, similarities, and strong relationship, when commenting on problems with text content	.352	.188	**.590**
Ability to express problem in the text using correct intonation, pronunciation, and pauses	0.44	**.644**	.488
Ability to use correct pronunciation pronunciation and combining the appropriate disclosures in the contents of the text	.380	**.466**	.367
Ability to use pauses according to the contents of disclosure in the text	.307	**.593**	.404
Ability indicate a problem related phenomena / current situation with the content in text	.278	.452	**.564**
Ability to relate problems in the text with environment in which to study	.418	**.475**	.460
Ability to utilize vocabulary appropriately according to text content	.314	**.703**	.040
Ability to use the right choice of words according to situation and content in text	**.656**	.295	.365
ability to use synonymous vocabulary with intents and utterances related to the content of text	**.793**	.111	-.030
Ability to show different word variations with proper explanation according to text content.	**.642**	.386	.259
Ability to use sentences in which the subject takes action according the relationship with the text	.407	**.476**	.318
Ability shows the pattern of active sentence in the text	**.566**	.517	.100
Ability to use sentences with subject to work by linking content of the text	**.703**	.407	.283
Ability to show passive sentence pattern in text	**.680**	.256	.202
Ability to demonstrate mandate or messages that are in the text	**.558**	.269	.268
Ability to show positive meaning word, vocabulary or sentences	**.637**	.491	.340

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation coverged in 8 iterations.

Based on Table 4 that shows the results of the rotation factor, it can be seen that the grouping of indicators into factors and the magnitude of loading factors obtained is seen in table 4. It appears that the determination of indicator inputs to certain factors follows the magnitude of the correlation between variables and factors, namely to those with large correlations. Thus the factors formed along with the items are presented in Table 5.

Table 5. Results of the Grouping of Indicators into Factors

Details	Factors formed	Correlation factor value	Factor name
Ability to communicate good and correct language		.474	
Ability to communicate a language that easy to implement		.664	
Ability to communicate in submitting report or documentation		.614	
Polite communication skills and courteous with other people and groups		.521	
Body language communication skills		.753	
Ability to use easy language according to text	1	.641	K1
Ability to use headings, intro, and content accurately in text		.643	
Ability to use easy language according to text		.751	
Ability to show differences, similarities, and strong relationship, when commenting on problems with text content		.590	
Ability to express problem in the text using correct intonation, pronunciation, and pauses		.644	
Ability to use correct pronunciation pronunciation and combining the appropriate disclosures in the contents of the text		.466	
Ability to use pauses according to the contents of disclosure in the text		.593	
Ability indicate a problem related phenomena / current situation with the content in text		.564	
Ability to relate problems in the text with environment in which to study		.475	
Ability to utilize vocabulary appropriately according to text content	2	.703	K2
Ability to use the right choice of words according to situation and content in text		.656	
ability to use synonymous vocabulary with intents and utterances related to the content of text		.793	
Ability to show different word variations with proper explanation according to text content.		.642	
Ability to use sentences in which the subject takes action according the relationship with the text		.476	
Ability shows the pattern of active sentence in the text		.566	
Ability to use sentences with subject to work by linking content of the text		.703	
Ability to show passive sentence pattern in text	3	.680	K3
Ability to demonstrate mandate or messages that are in the text		.558	
Ability to show positive meaning word, vocabulary or sentences		.637	

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation coverged in 8 iterations.

Table 5 shows the results of the grouping of 24 items into 3 factors formed and the naming of each factor. The results of the extraction from the factor analysis there are 3 factors formed, namely factor 1, there are 9 indices with loading values of 0.474, 0.664, 0.614, 0.521, 0.753, 0.641, 0.643, 0.751, and 0.590, then given the name KI. Factor 2, there are 9 indicators formed each having loading value of 0.644, 0.466, 0.593, 0.564, 0.475, 0.703, 0.656, 0.793, 0.642 named K2. Factor 3 there are 6 items, namely the loading load value respectively 0.476, 0.566, 0.703, 0.680, 0.558, and 0.637.

Factor Formation

Judging from the Total Variance Explained table, an eigenvalue value greater than 1 is obtained, namely 3 factors. This is in accordance with existing criteria, namely factors that have an eigenvalue of more than 1 will be maintained and factors that have an eigenvalue of less than 1 are not included in the model (Supranto in Wiratmanto, 2014, p.33). Furthermore, from the table also can be seen that the

cumulative total variance is quite large, namely 60.740% of the variability to the 24 original items. In addition to seeing the total variance table explained many factors can also be done by looking at the scree plot. Based on the existing scree plot, it starts to flatten on the initial grain extraction into 3 factors. Thus the maximum factor that can be formed is 3 factors. The names of the factors formed are the ability to use synonymous vocabulary with intent and description related to the contents of the text, the ability to communicate using body language, and the ability to use language that is easily understood and in accordance with the contents of the text.

The comparison of the results obtained by researchers with previous studies, i.e., in research conducted by Jamilah and Purnawan (2009) with the title of the study "Pengembangan Instrumen Pengukuran Hasil Pembelajaran Mata Kuliah Pronunciation". Research conducted by Jamilah and Purnawan (2009) different power tests for recognition tests resulted in 32 items that have the ability to distinguish test participants who are in the smart group and not smart, while 8 items require revision. The highest power difference is 0.612 and the lowest power difference (for non-deciduous items) is 0.308. Other research conducted by Patria (2016) with the title of the research "Pengembangan Instrumen Penilaian Unjuk Kerja Praktik Lompat Jauh Berbasis Rolling Assessment dalam Pembelajaran Penjasorkes Kelas VIII SMP" In the analysis of the score of the product trial results there are 16 items each valid on performance and cognitive instruments. Performance instrument reliability in product trials is 0.86 and usage test is 0.921, while cognitive instrument reliability in product trials is 0.93 and in the usage trial is 0.726. The performance instrument developed was constructing valid because it was able to explain the overall variation in the long jump material by 74,062% the practicality analysis of the squat-style long jump instrument obtained a practical score of 169 and was categorized as "practical".

Based on research that has been done that there is an increase in the extensive reading ability of English students of PBSI FKIP UMK. It happens in every learning by using authentic assessment. The increase in extensive reading skills has a significant effect because each student is measured by an authentic assessment according to the performance of the practice of reading English extensive.

CONCLUSION

The influence of the use of authentic assessment in learning English language skills is a critical success factor in learning the learner's characteristics. Difficulties, strengths and weaknesses that occur in the learning process can be measured properly and can be used as material or reference in improving reading skills. Authentic assessment provides a description of the assessment according to the reality that occurs in each individual, so that the ability possessed by the learner is recorded properly and is able to be used as reflection material to achieve the best results.

Metacognitive strategies in learning the ability to read English appropriately are used to facilitate the learning process to be more directed. Each stage in metacognitive, such as selection or planning, stages of the learning process, identification, reflection, and providing appropriate solutions to overcome the problems encountered is a process to obtain ideal learning and in accordance with the learning process.

REFERENCES

Arikunto, S. (2016). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.

Aziz, K., Kustiono., Lestari, W. (2019). Need Analysis for Development of a Cube and Beam Volume Assessment Instrument in Macromedia Flash Based Elementary Mathematics Learning. *Journal of Educational Research and Evaluation*, 8(2), 99-107.

Jamilah., Purnawan, A. (2009). Pengembangan Instrumen Pengukuran Hasil Pembelajaran Mata Kuliah Pronunciation. *Jurnal Diksi*, 16, 122-133.

Mardapi, D. (2017). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Yogyakarta: Parama Publishing.

Masrukhin. (2014). Pengembangan Instrumen Penilaian Otentik Mata Pelajaran Pendidikan Agama Islam Untuk Meningkatkan Kemampuan Evaluasi Dalam Pembelajaran. *Prosiding Seminar Nasional*. Semarang: Universitas Negeri Semarang.

Megawanti. (2015). Meretas Permasalahan Pendidikan Di Indonesia. *Jurnal Formatif*, 2(3), 227-234.

Namira, Z..B., Kusumo, E., Prasetya, A.T. (2014). Keefektifan Strategi Metakognitif Berbantu Advance Organizer Untuk Meningkatkan Hasil Belajar Kimia Siswa. *Jurnal Inovasi Pendidikan Kimia*, 8(1), 1271-1280.

Patria, L. D. (2016). Pengembangan Instrumen Penilaian Unjuk Kerja Praktik Lompat Jauh Berbasis Rolling Assessment dalam Pembelajaran Penjasorkes Kelas VIII SMP. *Tesis*. (Tidak dipublikasikan). Semarang: Universitas Negeri Semarang.

Prasetyo, A.P.B., Suaesih, S., Hadiyanti, L.N. (2016). Pendampingan Guru Biologi Kota Semarang dalam Mengembangkan Instrumen Penilaian Otentik Pada Kurikulum Berbasis Kompetensi. *Unnes Science Educational Journal*, 5(1), 1182-1190.

Santi, E., Setiawan, M.H., Khumaedi, M. (2019). Developing Assessment Instrument on Writing on Descriptive Paragraph Based on Picture of Junior High School Students. *Journal of Educational Research and Evaluation*, 8(2), 116-122.

Triyono., Sunarto,. Lestari, W. (2019). Need Analysis in the Development of Tolerance Attitude Assessment Instrument in Android-Based PPKn Learning. *Journal of Educational Research and Evalution*, 8(1), 65-72.

Widianto, E., Subyantoro. (2015). Peningkatan Keterampilan Membaca Teks Klasifikasi Menggunakan Metode SQ3R dengan Media Gambar. *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 4(1), 1-9.

Widiansah, K.N., Kartono., Rusilowati, A. (2019). Development of Assessment Instruments Mathematic Creative Thinking Ability on Junior High School Students. *Journal of Educational Research and Evaluation*, 8(1), 84-90.

Widiyana, S., Rusilowati, A., Priyanto, AT Sugeng. (2019). Three Tier Multiple Choise Test Instrument Design For Analysis of Student Misconceptions In The Subject of PPKn. *Journal of Educational Research and Evaluation*, 8(2), 91-97.

Wiratmanto. (2014). Analisis Faktor dan Penerapannya dalam Mengidentifikasi Faktor-Faktor yang Mempengaruhi Kepuasan Konsumen Terhadap Penjualan Media Pembelajaran. *Skripsi*. Yogyakarta: Universitas Negeri Yogyakarta.

Yusuf, A.M. (2017). *Asesmen dan Evaluasi Pendidikan: Pilar Penyedia Informasi dan Kegiatan Pengendalian Mutu Pendidikan*. Jakarta: Kencana.