ISEJ Indonesian Science Education Journal

ISEJ: Indonesian Science Education Journal
ISSN: 2716-3350
Vol. 4, No. 1, January 2023, Hal 10-21

Development of a Two-Tier Multiple Choice Question Assessment Instrument to Measure Students Science Process Skills on Acid-Base Material

Meri Andaria[1], Hadiwinarto[2]

[1,] *Mahasiswa Program Studi Doktor Pendidikan FKIP Universitas Bengkulu*
[2] *Fakultas Keguruan dan Ilmu Pendidikan Universitas Bengkulu*

Correspondent e-mail:
[1] mery_swaraunibfm@yahoo.com,

Abstract
This development research aims to: 1) develop a two-tier multiple choice question assessment instrument that is able to measure the science process skills of high school students, 2) determine the feasibility of a two-tier multiple choice question assessment instrument to measure high school students' science process skills, 3) determine the response teachers regarding the two-tier multiple choice question assessment instrument applied in schools. This research uses the Borg and Gall procedure which has been reduced to 9 stages, namely: 1) preliminary stage, 2) planning stage, 3) initial product design development stage, 4) initial trial stage, 5) initial product revision stage, 6) field trial stage, 7) field trial product revision stage, 8) operational trial stage, 9) final product revision stage. The research was conducted at SMA N 1 Bengkulu and SMA N 2 Bengkulu in the 2014/2015 academic year. Data collection techniques were carried out through questionnaires, observation, and tests. The types of data obtained are qualitative and quantitative data. The effectiveness of the assessment instrument product was carried out by means of a mean difference test using a sample of two classes for each school. This test is to compare the scores of students in the experimental class using the two-tier multiple choice question assessment instrument and the control class using the multiple choice question assessment instrument. The results showed that: 1) a two-tier multiple choice question assessment instrument that was able to measure the science process skills of high school students was successfully developed through 9 stages of Borg and Gall. 2) The feasibility of the TTMCQ assessment instrument product is guaranteed through the content validity which is considered good by material experts, the construct validity is considered good by the assessment instrument expert, the validity of the items is good, the reliability is high, has a difficulty level with the proportion of 53% easy, 42% moderate, and 5% difficult, has sufficient distinguishing power with minimal interpretation, and has a good level of practicality, which means it is suitable for use in learning. 3) Teacher responses include that the teacher has never used questions in the form of two-tier multiple choice questions, the question form is good and neat, teachers are willing to reuse multilevel multiple choice questions on the grounds that this form of questions challenges students to improve science process skills.

Keywords: Assessment Instrument, Two-Tier Multiple Choice Question, Science Process Skills

INTRODUCTION

The teacher process is a key person in the classroom so that the teacher has a very vital and fundamental role in guiding, directing, and educating students in the learning process. Apart from being a teacher and educator, one of the teacher's roles that cannot be ignored is as an evaluator, namely the teacher plays a role in carrying out assessments, compiling assessment instruments, and assessing student work (Suparlan, 2006).

Assessment is a series of activities to obtain, analyze, and interpret data about the process and student learning outcomes which are carried out systematically and continuously, so that it becomes meaningful information in decision making (BSNP 2006). The objectives of the assessment include the teacher being able to determine the success of achieving learning objectives, student mastery of lessons, the effectiveness of teaching methods and knowing the position of students in the class, the teacher can also classify whether a student is a group of students who are smart, moderate, or less in their class compared to their peers. his friend (Rusfidra, 2006).

Science is not just a product, it's a process. Science develops through research and investigation. To be able to do this, students need to master a number of basic science skills called science process skills. Assessment of science learning processes and outcomes demands more comprehensive techniques and methods of assessment (Stiggins, 1994). In addition to the aspects of learning outcomes that are assessed to be comprehensive, assessment techniques and assessment instruments should be more varied. One of the assessments that can be carried out by science teachers in the implementation of the teaching and learning process is the assessment of science process skills in accordance with Permendiknas 22 of 2006 concerning Content Standards and Permendiknas 23 of 2006 concerning Graduate Competency Standards.

According to Ango (2002) science process skills are an important component in the implementation of the learning process because they can affect the development of student knowledge. Therefore, the assessment of science process skills is an important component in learning chemistry. Process skills are all directed scientific skills used to find a concept, principle, or theory to develop a pre-existing concept, or to refute an invention. Process skills involve cognitive or intellectual, manual, and social skills. Cognitive or intellectual skills are involved because by doing process skills students use their minds. Manual skills are clearly involved in process skills because they involve the use of tools and materials, measurement, arrangement of tools. Social skills are meant that they interact with each other in carrying out teaching and learning activities with process skills. Science process skills are thinking skills used to create knowledge, reflect on problems, and formulate results (Aydin, 2013).

In general, science process skills are classified into basic process skills and integrated process skills. Basic process skills include observing, measuring, inferring, predicting, and classifying. Integrated process skills include controlling variables, interpreting data, formulating hypotheses, defining variables operationally, and designing experiments. The indicators are observing, grouping or classifying, interpreting, predicting, asking questions, formulating hypotheses, planning experiments, using tools or materials, applying concepts, and communicating (Depdikbud, 2013).

Chemistry as a science has different characteristics compared to other sciences. This difference lies in the complexity of the components which include processes, products, and attitudes. Thus, chemistry learning should be oriented towards process skills carried out through scientific work, not just memorizing concepts. It is through this process that scientific skills (scientific process skills) can be developed, so that the correct experience of science can be obtained.

Based on observations made at SMA N 1 Bengkulu and SMAN 2 Bengkulut about the analysis of the fulfillment of 8 SNPs carried out in November as a basis for communicating about educational development issues, it was obtained that standard data had not been met in both SMAs. Standards that are lacking in implementation at SMAN 1 Bengkulu are standard 2 (Process Standards), standard 4 (Educators and Education Personnel Standards) and standard 8 (Assessment Standards). Of the three standards, the standard that the author might improve is the assessment standard. It can be seen that there is a quite high difference of 2.32 from the ideal score of 39 (BAN, 2012) obtained a score of 34. Whereas in SMAN 2 Bengkulu the standard which is very poorly implemented is standard 8 (Standard Assessment). It can be seen that there is a fairly high gap of 3.71 from the ideal score of 39, only a score of 31 was obtained. This indicates that the fulfillment of the assessment standards needs to be developed.

Referring to the results of the mapping, interviews were conducted with subject teachers. From the results of the discussion, it turned out that the cause of incompleteness was due to the lack of practice questions as part of the learning evaluation. This results in students' abilities being less honed and developing. Another cause is that subject teachers only use two to three assessment techniques which cause the teacher to be less able to know the progress and difficulties of student learning. It appears that the assessment system for learning outcomes measures more cognitive aspects in the form of memorization. Assessment of science process skills has not been carried out. This is because there is no instrument yet. The results of the needs analysis in schools concluded that the teacher needed an assessment instrument that was able to measure students' science process skills.

METHOD

This study used a research and development research design. The research and development model carried out refers to Borg and Gall (1983). The following development steps:
1). Conducting preliminary research and information collection (research and information collecting), namely recognizing problems in the field, analyzing the teacher learning process, analyzing UN results, analyzing curriculum, analyzing questions, and conducting literature studies.

2). Planning, namely determining basic competencies, formulating learning indicators, and formulating indicators of science process skills. 3). Developing a preliminary form of product, namely preparing learning materials and making an assessment instrument grid. 4). Performing preliminary field testing, namely validating the initial form of the product to experts and users on a limited scale. 5). Make revisions to the main product (main product revision), namely making improvements in accordance with the suggestions from the results of preliminary field testing. 6). Performing main field testing, which is validating development products on a wider scale and compared to control products if possible. 7). Revising operational products (operational product revision), namely revising products based on suggestions from the results of the main field testing. 8). Performing operational field testing, namely conducting validation tests on the operational products produced. 9). Make revisions to the final product (final product revision), namely revising products such as suggestions from the results of operational field testing. 10). Disseminating and implementing products (dissemination and implementation), namely making reports on products at professional meetings and in journals, collaborating with publishers to conduct commercial distribution, and helping to provide quality control. The development step carried out in the research is only up to the ninth step because for the tenth step it requires high costs and very broad coverage in a long time.

RESULTS AND DISCUSSION

Characteristics of the Two-Tier Multiple Choice Question Assessment Instrument that Measures Students' Science Process Skills

Science process skills are thinking skills that are used to create knowledge, reflect on the symbols, and formulate the results (Aydin, 2013). Akinbobola and Afolabi (2010) see science process skills as mental and physical abilities and competencies as tools needed to learn science and technology effectively as well as cognitive and psychomotor skills used in problem solving, problem identification, data collection, transformation, interpretation. and communication. So that learning based on science process skills is very suitable to be implemented in this 2013 curriculum. Because the demands in the Content Standards (inquiry by providing direct learning experiences through the use and development of scientific process skills and attitudes) and based on the nature of science (products, processes, technology applications, attitudes) aim to increase the meaningfulness of science learning.

According to the Ministry of Education and Culture (2013), in general science process skills are classified into basic process skills and integrated process skills with indicators, namely observing, grouping, interpreting, predicting, asking questions, formulating hypotheses, planning experiments, using tools or materials, applying concepts, and communicating. . Science process skills are important to be developed in schools because they are in accordance with Permendiknas 22 of 2006 concerning Content Standards (SI) and Permendiknas 23 of 2006 concerning Graduate Competency Standards (SKL), in which chemistry as part of natural science is a science that is born and develops based on observation and experiment. Thus, studying chemistry is not enough just to memorize facts and concepts that have been made, but also required to discover these facts and concepts through observation and experiment. It is through this process that scientific skills (scientific process skills) can be developed, so that the correct experience of science can be obtained (Rustaman, 2003).

Measuring science process skills requires a valid and clear assessment, but the development of an assessment instrument to measure science process skills has not been widely used by educational practitioners. The formative assessments found in schools provide little opportunity for students to develop deeper knowledge (Cullinane, 2011). An alternative assessment that is able to measure students' science process skills is needed by teachers in schools. One of them is developing a two-tier multiple choice question form (multilevel multiple choice questions). The form of this question is in the form of rtnoyuasesrama with ordinary multiple choice. The form of a two-tier multiple choice question (TTMCQ) was adapted from Treagust (2006). The form of the TTMCQ question consists of two levels of questions, the first level is the content of the question which has two alternative answer choices, namely true and false and the second level is the reason for the answer given on the basis of the first choice. Inclusion of reasons at the second level aims to avoid students who sometimes guess answers, see students 'ability to understand reasons, and make it easier for teachers to measure students' science process skills compared to ordinary multiple choice questions (Cullinane, 2011).

The development of the TTMCQ assessment instrument began with the creation of a development draft. This draft is the initial form of the product made and contains an introduction, steps for the development of an assessment instrument, and an assessment instrument which includes an assessment instrument grid, a sheet containing the assessment instrument items, an answer sheet provided to work on questions, answer keys to questions, as well as a question assessment rubric. The assessment instrument set was then validated to material experts, assessment instrument experts, and senior teachers.

Material experts involved in the validation are experts in the field of acid-base materials. The purpose of validating material experts is to ensure the validity of the contents of the assessment instrument items developed so that there are no misconceptions. Expert assessment

instruments used in validation are experts in the field of learning evaluation. The purpose of validation for assessment instrument experts is to ensure the validity of the instrument constructs developed so that it can measure science process skills. The senior teachers involved are chemistry teachers who have experience teaching in high school. The purpose of validating senior teachers is to ensure the appropriateness of the assessment instruments before they are applied in schools.

The results of the validation of the assessment instruments concluded that they had good content validation and construct validation and were feasible to be applied in schools after several revisions were made. Revisions made include correcting the concept of material that is still wrong and inaccurate, improving the appropriateness of the use of learning indicators and scientific process skills indicators, improving question construction, improving the suitability of the subject matter with the answer reasons, improving the distractor balance of the answer reasons, correcting writing errors, and simplifies question writing. The responses and suggestions from experts and senior teachers were used to improve the assessment instrument at the limited trial stage.

The initial trial of the assessment instrument was carried out in two stages, namely one-on-one trials and small group trials (Borg and Gall, 2007). The one-on-one trial was carried out on six students. The results of the one-on-one trials concluded that the assessment instrument was quite easy for students to understand. Students are happy with the form of questions that have never been given by the teacher at school. Small group trials were conducted on fourteen students and two teachers who were randomly selected and did not include students in the previous trial. The results of the limited trial concluded that the two-tier multiple choice question assessment instrument could be applied in schools after going through several revisions. Revisions made were correcting ineffective sentences and correcting unclear images. The students' responses on average indicate that this question is a difficult question and requires thinking time to analyze the two levels of the question. The responses and suggestions given in the initial trial are analyzed to improve the assessment instrument that will be applied to field trials.

Field trials were carried out on fifty-three students and two teachers. The results of the field trials were analyzed to obtain the item validity, reliability, distinction power, and the difficulty level of the questions. The results of the field trials concluded that the two-tier multiple choice question assessment instrument from the average results of the two SMAs had good validity, high reliability, the average difficulty level of the two schools with the proportion of 53% easy, 42% moderate, and 5% is difficult, has sufficient distinguishing power with minimal interpretation, and has a good level of practicality so that it is in accordance with the quality standards of the assessment instrument according to Arikunto (2007), namely valid, reliable, and practical.

The operational trial was conducted on sixty-six students and two teachers who were randomly selected and did not include students in the previous trial. The results of operational trials concluded that educational products are ready for use in schools without the presence of developers based on Borg and Gall, (1983). In this operational trial stage, it can also be seen that the product's feasibility and excellence in operational practice according to Sukmadinata (2006). So that the two-tier multiple chcoic assessment instrument can already be applied in schools after going through several revisions. The revision made was correcting less effective sentences. Responses and suggestions given to operational trials are analyzed to improve the assessment instruments that have become the final product.

The two-tier multiple choice question assessment instrument that is able to measure students' science process skills has characteristics that are developed based on indicators of science process skills from the Ministry of Education and Culture (2013) including the skills of

observing, grouping, interpreting, predicting, asking questions, formulating hypotheses, planning experiments, using tools or materials, applying concepts, and communicating and having good validity, high reliability, difficulty level of questions with the proportion of 53% easy, 42% moderate, and 5% difficult, has sufficient distinguishing power with minimal interpretation, and has a practical level of questions rated good. So that it becomes the final product of an assessment instrument that can be used to measure science process skills. The use of this product is able to stimulate students to carry out scientific thinking processes.

Product Feasibility Assessment Instrument Two - Tier Multiple Choice Question to Measure Science Process Skills

The development outcome assessment instrument aims to help teachers prepare questions that are able to train students' science process skills on acid-base material. The purpose of the development was because according to UN data and daily tests of acid-base material for the 2013/2014 academic year there were still incomplete students. This acid-base material requires students to carry out a scientific process, therefore learning in this material needs to be further developed in order to achieve completeness.

A good assessment instrument must have validity, reliability, objectivity, and practicality. The validity of the assessment instrument is guaranteed through the validity of the content and the validity of the constructs assessed by material experts and assessment instrument experts prior to the application of the assessment instrument widely. Validation is also carried out for senior teachers. The validity of the content is important to ensure that the items of the assessment instrument developed do not have a misconception. The validity of constructs is important to ensure that the construction of the questions developed is able to measure students' science process skills. Validation of senior teachers is important to assess the theory of assessment instruments that will be applied in schools. Testing of the assessment instrument is continued to the users gradually.

The first stage of testing is a one-on-one trial, which is testing limited to three students from SMAN 1 Bengkulu and SMAN 2 Bengkulu. The trial results showed that the readability of the initial product development results was quite good. The student assessment states that the development results have a fairly good sentence structure, the meaning of the questions is quite clear, the terms used are quite easy to understand, and the instructions for working on the questions are quite clear. Written errors on several questions were still found in this one-on-one trial. However, the overall legibility of the development results assessment instrument was quite good.

The second stage of testing is the small group trial, which is testing the questions in small groups, namely one teacher and seven students each from SMAN 1 Bengkulu and SMAN 2 Bengkulu who are not included in the previous trial. The purpose of small group trials is to assess the feasibility of the product by user students and teachers. The results of the product trials concluded that the development outcome assessment instrument was suitable for use in schools.

The third stage of testing is field testing, which is testing the questions in the field, namely fifty-three students, each with twenty-one students at SMAN 1 Bengkulu and thirty-two students at SMAN 2 Bengkulu. The purpose of field trials is to test the validity, reliability, differentiation power, and difficulty level of the developed items. The results of the analysis of the items in the field test showed that the development results were good. The questions have good validity, high reliability, difficulty level with the proportion of 53% easy, 42% moderate, and 5% difficult, have sufficient distinguishing power with minimal interpretation, and have a good level of practicality.

The fourth stage of testing is operational testing, which is testing questions in the field that are close to school operations, namely one teacher and thirty-four students from SMAN 1

Bengkulu and thirty-two students from SMAN 2 Bengkulu. The objective of operational trials is to assess product viability and excellence in operational practice. The results of the operational trials concluded that the development outcome assessment instrument was feasible to be applied in schools operationally into the final product

The practicality of the problem is that the probability of reusing it is high. The practicality of using two-tier multiple choice questions has been proven by Kartimi (2012) who developed a critical thinking measurement tool using two-tier multiple choice in the concept of thermochemistry. Salirawati (2011) developed a Chemical Balance Misconception Detection Instrument (IPMKK) which is the result of a combination of two-tier multiple choice questions with the inclusion of open reasons. The results of Salirawati's research (2011) prove that the feasibility test for high school chemistry teachers does not experience difficulties in implementing IPMKK and analyzing it.

Teacher Response to the Two-Tier Multiple Choice Question Assessment Instrument

The teacher's response to the development outcome assessment instrument was obtained from the results of questionnaires and interviews. The results include that the teacher has never used questions in the form of two-tier multiple choice questions, questions in the form of multilevel multiple choice are more complicated than ordinary multiple choice, the shape of the questions is good and neat, the form of this question is more challenging to ability. thinking of students, this question is suitable for seeing the ability of students to think to what extent, this question is suitable to be used to see the effectiveness of learning carried out by the teacher as well as understanding the concepts of the students 'counter-students' being taught.

The results of the questionnaire showed that not all teachers were willing to reuse multilevel multiple choice questions to measure students' understanding of the acid-base material. The form of this question is considered more suitable as a matter of diagnosing thinking skills and is not relevant for scoring. However, other teachers were willing to reuse multilevel multiple choice questions on the grounds that this question form challenged students to improve science process skills that were not just memorizing, could be used as daily test questions because they could focus on this material only and the questions were simple but students had to really understand the concept of the material, and are also able to avoid students guessing answers like when working on ordinary multiple choice questions because of the level of questions, namely the subject matter and the reasons for the answers are interrelated.

The teacher's response to the assessment instrument for development results is in accordance with Sampson's (2006) research which states that the advantage of using two-tier multiple choice questions for teachers is as an insight into making a form of assessment that challenges student knowledge and provides a way to assess students' conceptual understanding. Research from Treagust (1995) states that the form of two-tier multiple choice questions is more effective for determining alternative material concepts for students and for seeing whether a learning is meaningful or not.

Tuysuz's (2009) research is also in line with the results of the study proving that the use of two-tier multiple choice tests can help teachers teach better and help students learn better. The form of two-tier multiple choice test questions can be used by teachers to increase student knowledge and maintain students' alternative understanding and help improve the teaching and learning process when compared to the regular multiple choice test.

The results of Nofiana's (2013) research also revealed that the application of an evaluation instrument of two-tier multiple choice questions to measure higher-order thinking skills designed in such a way can have a positive or significant effect on student learning outcomes.

Susilo (2013) developed a science process skills test instrument using a multiple choice form with four alternative answers to the digestive system material. Susilo's research (2013) reveals that this test in the form of 35 multiple choice questions has been successfully developed by experts. In this study also revealed several weaknesses in this form of test questions, including the possibility of students guessing answers or answering questions luckily.

Another study that is in line with the results of the study is research from Halaydina and Downing (1989) which states that the two-tier multiple choice question form can be used to test student understanding and measure cognitive skills at a higher level (higher order thinking). Treagust's research (2006) also proves that the form of two-tier multiple choice questions can be used to improve students' thinking skills. Another supportive research is research by Cullinane (2011) which states that the use of a two-tier multiple choice question form can improve learning assessment and deeper thinking skills.

The results of this development research are also in accordance with the research of Kilic and Saglam (2009) which concluded that the multilevel multiple choice test form can be used to determine students' additional understanding of the concept of genetics. Kwen and Cheng's research (2005) also concluded that the use of multilevel multiple choice questions has several advantages compared to ordinary multiple choice questions, including providing creative thinking or graded thinking to students, providing in-depth knowledge of the reasons given by students and knowing the misconceptions that occur. .

The results of the mean difference test of the assessment instrument between the two-tier multiple choice question and the multiple choice question used the two-party t-test to find out whether there was a difference in the mean scores of the two classes from SMAN 1 Bengkulu and SMAN 2 Bengkulu where the results of both data were in outside the critical area, it means that the mean acid-base values of the TTMCQ and MCQ classes are not significantly different. This two-party t-test score explains that students give the same response in working on the questions, using either the TTMCQ or MCQ assessment instruments. This means that if students are able to work on questions in the MCQ form, then students are also able to work on questions in the TTMCQ form. The conclusion obtained from testing the mean difference is that students give the same response to both forms of assessment instruments.

The mean difference test shows that the mean value of the classes tested by the application of the TTMCQ assessment instrument with the class being tested, the application of the MCQ assessment instrument for the two SMAs is not significantly different. The result of the analysis showed that the average score of students' science process skills at SMAN 1 Bengkulu which was tested with the TTMCQ form was 3.05 and those tested with the MCQ form was 3.15. Whereas in SMAN 2 Bengkulu, the one tested with the TTMCQ form was 2.73 and the one tested with the MCQ form was 2.82. The two classes that were treated showed that there was a difference in the average score of students' science process skills, but these results did not show a significant difference based on the results of the mean difference test. 1 Bengkulu is superior to students at SMAN 2 Bengkulu.

The average score of science process skills on the two forms of instruments in both SMA shows a score below the KKM applied in their respective SMA. This means that students' science process skills have not shown maximum results. Science thinking process skills cannot occur spontaneously because to have this ability requires a lot of practice. Learning to develop science process skills will be successful if you do lots of practice or tests (Sagala, 2011). The value of science process skills can be obtained well if the teacher accustoms students to work on questions that empower science process skills.

As for the mastery of students' science process skills in each indicator in both SMA, the results of the analysis showed that the overall mastery of science process skills at SMA N 1 Bengkulu was higher than in SMA N 2 Bengkulu with higher basic science process skills than

integrated process skills . Furthermore, for the achievement of the highest indicator on indicator 1, namely observing with a value of more than 90% of students can master this indicator in both SMA N 1 Bengkulu and SMA N 2 Bengkulu. While the lowest achievement is in indicator 7, which is planning an experiment with a value of less than 40% of students who can master it.

The results of the analysis are in line with Ongowo's (2013) research, which showed that over a period of ten years, the students' science process skills with the highest achievement were observing or observing skills. These results are also in line with the research of Afolabi and Akinbobola (2010) which states that basic science process skills are higher than integrated science process skills.

Research on the development of assessment instruments aims to obtain alternative formative evaluations that are able to measure students' science process skills supported by Cullinane (2011) which states that the form of TTMCQ questions can be used as an alternative to formative assessment. Formative assessment is an assessment activity that aims to seek feedback on the teaching and learning process that is or has been implemented (Purwanto, 2010). Formative evaluation is given at the end of each learning material. Arikunto (2007) revealed the benefits of formative evaluation for students and teachers in schools. Benefits for students, among others, can be used to determine the mastery of learning material by students, as reinforcement for students, and as a diagnostic tool to find out which parts of the learning material are still difficult to master. Benefits for teachers include knowing the extent to which the subject matter can be accepted by students and in accordance with the objectives to be achieved. Science process skills are a demand from the Content Standards and Competency Standards for high school graduates. If students are often trained to use a two-tier multiple choice question assessment instrument that empowers science process skills, it is expected that student learning outcomes and science process skills will improve.

CONCLUSION

The conclusions from the results of this study are as follows 1). The characteristics of the two-tier multiple choice question assessment instrument that measure science process skills, among others, are developed according to the Ministry of Education and Culture (2013) which are generally classified into basic process skills and integrated process skills with indicators namely observing, grouping, interpreting, predicting, asking questions, formulating hypotheses , planning experiments, using tools or materials, applying concepts, and communicating and having good validity and high reliability. 2) The feasibility of the two-tier multiple choice question assessment instrument product is guaranteed through the validity of the content which is considered good by the material expert, the validity of the construct that is considered good by the assessment instrument expert, the validity of the items is good, the reliability is high, has a difficulty level of questions with a proportion of 53 % easy, 42% moderate, and 5% difficult, has sufficient distinguishing power with minimal interpretation, and has a good level of practicality. 3) The teacher's response to the application of the two-tier multiple choice question assessment instrument was obtained through the results of a questionnaire and interview. The results include that the teacher has never used questions in the form of two-tier multiple choice questions, questions in the form of multilevel multiple choice are more complicated than regular multiple choice, the form of questions is good and neat, this form of questions is more challenging for students' thinking skills. suitable to be used to see the effectiveness of learning carried out by the teacher and understanding of concepts in students of the material being taught. The results of the questionnaire show that teachers who are willing to reuse multilevel multiple choice questions on the grounds that this question form challenges students to improve science process skills that are not just memorizing, can be used as daily test questions because they can focus on this material only and are also able to avoid students from guessing the answer is like doing the usual

Socaolmpmiliithaton ugsaenrda because there are levels of questions, namely the subject matter and the reasons for the answers are related. While the results of the mean difference test show that the average class score tested is the application of the TTMCQ assessment instrument with the class being tested, the application of the MCQ assessment instrument for the two high schools does not differ significantly, meaning that students give the same response to the two forms of assessment instruments so the development instrument is suitable for use in schools.

REFERENSI

Adesoji, F. A. dan Omilani, N.A. 2012. A Comparison of Secondary Schools Students Levels of Conception of Qualitative And Quantitatif Inorganic Analysis. American Journal of Scientific And Industrial Research. ISSN: 2153-649X.

Akinbobola, A. O. dan Afolabi, F. 2010. Analysis of Science Process Skills in West African Senior Secondary School Certificate Physics Practical Examinations in Nigeria. American-Eurasian Journal of Scientific Research 5 (4): 234-240. ISSN 1818-6785

Ango, M.L. 2002. Mastery of Science Process Skills and Their Effective Use in the Teaching of Science: An Educology of Science Education in the Nigerian Context. International Journal of Educology 1 (16) : 11-30.

Agusyana, Y. 2011. Olah Data Skripsi dan Penelitian SPSS 18. Jakarta: Gramedia Aydin, A. 2013. Representation of Science Process Skills In The Chemistry Curricula for Grades 10, 11, and 12/ Turkey. International Journal of Education and Practice, 2013, 1(5) : 51-63.

Arikunto, S. 2007. Dasar-Dasar Evaluasi Pendidikan. Jakarta : PT Bumi Aksara. Ariyati, E. 2009. Keterampilan Proses Sains. Makalah Tidak Dipublikasikan. Program Studi Ilmu Pengetahuan Alam Konsentrasi Pendidikan Biologi-SL. Sekolah Pasca Sarjana Universitas Pendidikan Indonesia.

Borg, W.R dan Gall, M.D. 1983. Educational Research An Introduction (4th Ed). White Plains : Logman Inc.

Borg, W.R., Gall, M.D., dan Gall, J.P. 2007. Educational Research An Introduction (8th Ed). White Plains : Logman Inc.

Budiyono. 2009. Statistika Dasar untuk Penelitian. Surakarta: UNS Press. Chabalengula, V.M., Mumba, F. dan Mbewe, S. 2012. How Pre-Service Teachers Understand and Perform Science Process Skills. Eurasia Journal of Mathematics,

Science & Technology Education, 2012, 8(3), 167-176. Southern Illinois University Carbondale, USA.

Cullinane, A. dan Liston, M. 2011. Two Tier Multiple Choice Question : An Alternative Method of Formative Assessment for First Year Undergraduate biology Students. Limerick : National Center for Excellence In Mathematics and Education Sciences Teaching and Learning.

Depdiknas. 2008. Panduan Penulisan BcuotmirmSoitatlo. Juaskearrta : Depdiknas.

Dewantari, A. 2015. Laporan Pengolahan Hasil Angket SNP SMAN 1 Surakarta dan SMAN 2 Surakarta Tahun Pelajaran 2014/2015. Laporan Tidak Dipublikasikan. Pascasarjana FKIP Universitas Sebelas Maret Surakarta.

Halaydina, T.M dan Downing, S.M. 1989. A Taxonomy of Multiple Choice Item Writing Rules. Applied Measurements In Education, 2(1), 37-50.

Harlen, W. 1992. The Teaching of Science: Studies in Primary Education. London: David Fulton Publishers.

Kartimi. 2012. Pengembangan Alat Ukur Berpikir Kritis pada Konsep Termokimia untuk Siswa

SMA Peringkat Atas dan Menengah. Jurnal JPII 1 (1) (2012) 21-26.

Kilic, D. dan Saglam, N. 2009. Development of Two Tier Diagnostic Test to Determine Students Understanding of Concepts In Genetics. Egitim Arasmalari-Eurasian Journal of Education Research, 36, 227-244.

Kutluay, Y. 2005. Diagnosis of Eleventh Grade Student Misconceptions About Geometric Optic By A Three-Tier Test. Thesis. The Graduate School of Natural and Applied Sciences of Middle East Technical University.

Kwen, B.H. dan Cheng, A.K. 2005. Using Two Tier Reflective Multiple Choice Questions to Cater to Creative Thinking. AARE 2005 International Education Research Conference.

Nofiana, M. 2013. Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Question untuk Mengukur Keterampilan Berpikir Tingkat Tinggi pada Materi Kingdom Plantae. Tesis Tidak Dipublikasikan. Universitas Sebelas Maret.

Permendiknas No. 20 Tahun 2007. Standar Penilaian Pendidikan. Jakarta : Depdiknas.

Permendiknas No. 22 Tahun 2006. Standar Isi Pendidikan. Jakarta : Depdiknas.

Permendiknas No. 23 Tahun 2006. Standar Kompetensi Lulusan. Jakarta : Depdiknas.

Purwanto, N. 2010. Prinsip-Prinsip dan Teknik Evaluasi Pengajaran. Bandung : PT Remaja Rosdakarya.

Ramirez, R.P.B. dan Ganaden, M.S. 2008. Creative Activities and Students' Higher Order Thinking Skills. Journal of Education Quarterly, Vol 66 (1), 22-23.

Rusfidra. 2006. Penilaian Proses Belajar Mengajar IPA Di Kelas Melalui Pedagogi. Tersedia di: http://rayapkabel.wordpress.com/ diakses pada 25 April 2014.

Rustaman, N.Y., dkk. 2012. Strategi Belajar Mengajar Biologi. Common Textbook JICA IMSTEP. Bandung: FPMTP A UPI.

Sagala, S. 2011. Konsep dan Makna Pembelajaran untuk Membantu Memecahkan Masalah Belajar dan Mengajar. Bandung : Alfabeta.

Sampson, V. 2006. Two-Tier Assessment. Teacher Toolkit. Arizona : College of Education At Arizona State University In Tempe.

Stiggins, R.J. 1994. Student Centered Classrom Assesment. USA: Macmillanr College Publishing Company.

Subali, B. 2009. Pengembangan Tes Pengukur Keterampilan Proses Sains Pola Divergen Biologi SMA. Makalah disajikan pada Prosiding Seminar Nasional Biologi, Lingkungan dan Pembelajarannya. Jurdik Biologi, FMIPA, Universitas Negeri Yogyakarta, 4 Juli 2009.

Sudjana, N. 2009. Penilaian Hasil Proses Belajar Mengajar. Bandung : PT Remaja Rosdakarya.

Sukiman. 2012. Pengembangan Sistem Evaluasi. Yogyakarta : Insan Madani. Suparlan. 2006. Guru Sebagai Profesi. Yogyakarta: Hikayat Publishing.

Supranata, S. 2004. Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004. Bandung: PT Remaja Rosdakarya Offset.

Susilo, H. 2013. Pengembangan Tes Keterampilan Proses Sains Materi Sistem Pencernaan Kelas XI SMA N 1 Pemalang. Skripsi. Tersedia di lib.unnes.ac.id/18954/1/4401408098.pdf diakses pada 2 Februari 2015.

Treagust, D.F. 2006. Development and Use of Diagnostic Test to Evaluate Students Misconception In Science. International Journal of Science Education, 10, 2 pp 159-169.

Tuysuz, C. 2009. Development of Two-Tier Diagnostic Instrument and Assess Students Understanding In Chemistry. Scientific Research and Essay Academic Journals Vol 4 (6) pp. 626-631, June 2009. ISSN 1992-2248.

Vichitvejpaisal, P., et al. 2011. Developing a Two-Tier Diagnostic Test to Assess Arterial Blood

Gates Learning by Students with Different Background Knowledge In Anesthesiology. South-East Asian Journal of Medical Education Vol 5 No 2, 2011.

Walid, A., Sajidan, S., Ramli, M., & Kusumah, R. G. T. (2019). Construction of the assessment concept to measure students' high order thinking skills. Journal for the Education of Gifted Young Scientists, 7(2), 237-251.

Wartono. 2003. Strategi Belajar MengacjaormFmiistiktao.uMsearlang: JICA.

Widodo, W. 2013. Implementasi Kurikulum 2013 dalam Pembelajaran Sains dan Keterkaitannya dengan Ide-Ide Penelitian Pendidikan Sains Menuju Indonesia Maju. Makalah disajikan dalam Seminar Nasional Pendidikan Sains UNS.

Yuliati, L. 2008. Model-Model Pembelajaran Fisika. Malang : Universitas Negeri Malang.