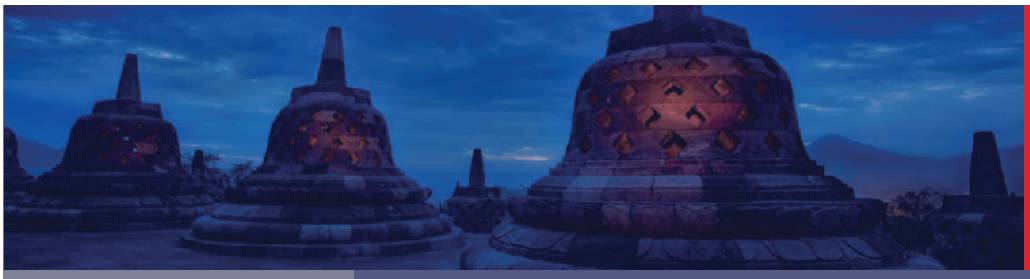

ISSN: 2407-1501

PROCEEDING

INTERNATIONAL CONFERENCE ON EDUCATIONAL RESEARCH AND EVALUATION (ICERE)

"Authentic Assessment for improving Teaching Quality"

November 8-9, 2014
Rectorate Hall and Graduate School
Yogyakarta State University
Indonesia

Organized by:
Doctoral and Master's Program in
Educational Research and Evaluation
Graduate School, Yogyakarta State University

Proceeding
International Conference on Educational Research and Evaluation (ICERE) 2014

Publishing Institute
Yogyakarta State University

Director of Publication
Prof. Djemari Mardapi, Ph.D.

Address
Yogyakarta State University
ISSN: 2407-1501

BACKGROUND

In its effort to improve the quality of education in Indonesia, the Indonesian government has imposed Curriculum 2013 on schools of all level in Indonesia. The main difference between Curriculum 2013 and the previous curriculum lies in its implementation which uses the scientific approach. For the reason, teachers need to develop teaching strategies different from those they used to apply in the implementation of the previous curriculum. Besides, teachers also need to develop the techniques of evaluating students' learning achievement, which are relevant to the scientific approach. The evaluation has to be able to show the students' learning achievement in observing, experimenting, social networking, etc.

Authentic assessment conducted in the classroom and focusing on complex and contextual tasks enables students to perform their competence in a more authentic arrangement. It is very relevant to the authentic approach integrated in their teaching process, especially at elementary schools, or for appropriate lessons. It must be able to show which attitude, skill, and knowledge have or have not been mastered by the students, how they use their knowledge, what aspect they have or have not been able to apply, and so on.

On the basis of the above consideration, teachers can identify what materials the students can study further and for what material they need to have a remedial program. Authentic assessment, however, is not that easy!

FOREWORD

In the academic year of 2014, the government in this case the Ministry of Education and Culture has established the policy to run the curriculum of 2013 for the all levels of elementary and intermediate education in Indonesia. It means the schools have to be ready to implement the Curriculum of 2013. Basically, the implementation of the 2013 curriculum is an effort from the government to enhance the quality of education.

One of the characteristics of the 2013 curriculum is make use the scientific approach in the learning process. This approach is to improve the students' creativity in learning. In general, this approach seems to be a new thing for the teachers in which several problems and obstacles appear in its practice. The teachers are required to develop the learning strategies and the assessment systems which are relevant and appropriate in order to nurture the students' creativity. One of the assessment methods that can support the concept of scientific approach is by sing the authentic assessment. Authentic assessment can give the description of the knowledge, the attitudes, and the skills as well as what has or has not owned by the students and the way they apply their knowledge. Also, in what case they have or have not been able to implement the learning acquisition.

Based to the above circumstances, the Study Program of Educational Research and Evaluation, Graduate School of Yogyakarta State University (Universitas Negeri Yogyakarta) conduct the international seminar on the theme "Classroom Assessment for Improving Teaching Quality". There will be three sub-themes on this seminar, i.e. Issues of Classroom Assessment Implementation, Implementation of Authentic Assessment, and Developing a Strategy of Creative Teaching. By having this seminar, the participants are expected to possess the knowledge and the skills to develop and to apply the authentic assessment.

Yogyakarta, November 8, 2014
Head of Committee

Prof. Dr. Sudji Munadi

CONTENTS:

Theme 2:
Implementation of Authentic Assessment

Theme 3:
Developing a Strategy of Creative Teaching

EFFECTIVENESS OF REASONED OBJECTIVE CHOICE TEST TO MEASURE HIGHER ORDER THINKING SKILLS IN PHYSICS IMPLEMENTING OF CURRICULUM 2013

Edi Istiyono[1)], DjemariMardapi[2)], and Suparno[3)]
[1), 3)]Physics Education Department, Mathematics dan Natural Science Faculty,
Yogyakarta State University
[2)]Electrical EngineeringEducation Department,EngineeringFaculty,
Yogyakarta State University

email: edi_istiyono@uny.ac.id; edi_istiyono_uny@yahoo.co.id

Abstract

This research has been done to studyusefulnessand effectiveness of the reasoned objective choice test to measure higher order thinking skills (HOTS) in Physics (PhysROCTHOTS) implementing Curriculum 2013. Research subject was 1001 students of ten senior high school in Yogyakarta. Usefulnessof the test was indicated by validity, reliability, total information function,and SEM. The test validity was obtained by experts judgment and goodness of fit to model. Goodness of fit of the test scoring in polytomousaccording to the partial credit model (PCM) to measure HOTS based on average of INFIT Mean of Squareof 1.0. The effectiveness of the test was indicated by practicalityandefficiencybased on response of senior high schools physics teachers. The results show that:the PhysROCTHOTSis valid and reliable so the test is can be used to measure HOTS of senior high school students and the test scoring in polytomous according to PCM can measure HOTS of senior high school students implementing Curriculum 2013 effectively.

Keywords*: Reasoned Objective Test*, *Hots*, *Physics*, *Curriculum 2013*

Introduction

To improve the competitiveness of countries in the world trying to improve the quality of human resources.To improve the quality of human resource begins with improving the quality of national education. To improve the education quality in Indonesi ais done by reviewing the curriculum every 10 years periodically. The curriculum has been approved by the government in 2013 further tried out its implementation in July 2013. Since July 2014 curriculum2013 implemented nationally.

Curriculum 2013 was developed in three domains, namely attitudes (affective), knowledge (cognitive), and skills (psychomotor). The knowledge gained through the activities: remembering, understanding, applying, analyzing, evaluating, and creating (Lampiran PeraturanMenteri Pendidikan dan Kebudayaan RI No.66 Tahun 2013). The sixlevels of proficiency level abilities according to revised loom's taxonomy. Three low levels ability, namely: remembering, understanding, applying are alower order thinking skills (LOTS), while

the next three abilities, namely: analyzing, evaluating, and creating are a higher order thinking skills (HOTS) (Anderson & Krathwohl, 2001:30). Thus, curriculum 2013 mandated the learning process activities based on Bloom's taxonomy revised, evenmore emphasis on higher order thinking skills. This implies that the assessment of course, also includes the six abilities, not only LOTS but more emphasized HOTS.

Bloom's Taxonomy has long been applied in education and the present taxonomy is still used in many curricula and teaching material (Brookhart, 2010:39; Schrawand Robinson, 2011:158-159). Thus it was appropriate to adopt the learning ability in Bloom's taxonomy.

Based on Piaget's development theory, formal operational stageis the stage of the child began eleven years old. At this stage the child has begun to develop the ability to manipulate eabstract concept sthrough the use of propositions and hypotheses, whichis HOTS (Piaget, 2005:122 and Reedal, 2010:7). Senior high school students age between15 to 18years, so the high order thinking skills of high school students are already well established.

The curriculum 2013 aims to prepare Indonesian man to have the ability to live as a private citizen and a faithful, productive, creative, innovative, and affective and able to contribute to society, nation, state, and world civilization (Lampiran Peraturan Menteri Pendidikan dan Kebudayaan RI No. 69 Tahun 2013). This is what indicates the importance of HOTS. Similarly, in a physics lesson, students arenot only required lower order thingking skills but also mastered the Physics high reorder thingking, which includes physics abilities in analyzing, evaluating, and creating.

To monitor the process, progress, and improvement of student learning outcomes continuously, the necessary assessment. Educational assessment is the process of collecting and processing information to measure the achievement of student learning outcomes (Peraturan Menteri Pendidikan dan Kebudayaan RI No. 66 Tahun 2013). In addition, an assessment of individual data collection activities that result illustrates characteristics (Djemari Mardapi, 2012: 7-12). Thus, assessment of learning outcomes in physics is the rating of the outcome of the learning process in physics that forms numbers describing the characteristics of individual students. Assessment can be done orally or in writing. A written appraisal performed by a written test. There are two forms of the written test questions, namely: selecting answers and supplying answers. Answer written test questions by selecting answers include: multiple-choice, two-choice (true-false, yes-no), matching, and causation.

Assessment should: (1) be designed to measure the knowledge and concepts, science process skills, and high order reasoning; (2) adopt the question type of PISA and TIMSS to encourage teaching and learningthat contribute to the improvement of students'scientific literacy and explore the scientific thinking skills, critical, creative, and innovative; (3) emphasizes the mastery of lower and higher order concepts with various forms of assessment (multiple choice, multiple choice reasoned, limited description); and (4) introducing the type of questions that tested both nationally and internationally to students and science teachers (PusatKurikulum, 2007: 23-24).

Assessmen tin education using two kinds of measurement theory, namely: classical measurement and modern measurement theory. Classical test theory is also called the Classical True-Score Theory, Classical Test Theory named because elements of this theory has been developed and applied for a long time, but still valid today (Suryabrata, 2000:21). According to classical measurement theory that test scoring is usually done partially based on the steps that must be taken to correctly answer an item. Scoringis done every step and scores obtained by participants per item sums core of learners every step, and the ability estimated by raw scores. The scoring modelis not necessarily appropriate, because the difficulty level of each step is ignored.

PCM is also appropriate to analyze the response to the measurement of critical thinking and conceptual understanding in science (Linden & Hambleton, 1997:101-102). PCM was developed to analyze the test items that require several steps to resolve. PCM can be given to measures that can be under taken by individuals. Thus HOTS in physics test scoring would be more suited to this PCM.

The fact tha tmultiple choice tests are more widely used than other forms oftesting. This is because the usual multiple-choice test has certain advantages, among others: (1) the material being tested can covermost of the learning materials, (2) the student answers can be corrected easily and quickly, (3) answer to each question is definitely true or one, so that an objective assessment (Sudjana, 1990:49). While there are also drawbacks of this test, namely: (1) the possibility of students to guess the answer is still quite large and (2) the thinking of student scan not be seen with the real (Sujana, 1990:49). Therefore we need a model that can reduce the weakness of such tests, reasoned objective multiple choice test as an alternative solution.

Assessment model was also influential onstudents' thinking skills. Van den Berg(2008: 15) states that the assessment can be implemented to assist students in improving their ability

to think critically. This is supported by another opinion, that the higher order thinking questions to encourage students to think deeply about the subject matter (Barnett & Francis, 2012:209). Based on these two opinion scan be concluded that the high order thinking skills tests can provide a stimulus for students to develop high ordert hinking skills as well.

Based on the results of preliminary surveys by conducting interviews with senior high school physics teachers in the Special Region of Yogyakarta (DIY), that the first, most schools use multiple-choice objective test that is common in the midterm test and final test of the semester. Second, multiple-choice tests that are used in most senior high school physics measure LOTS ability to: remember, understand, and apply. Third, multiple-choice test scoring usual dichotomous model, which means that if the item response is correct as givena score of 1, and ifonewas givena score of0.The scoring models have not been fair because of the different error rate to get the same score namely 0. Thus, a multiple-choice test main stay test models used in senior high school and still measure the lower order thinking skills do not measure Physics higher order thinking skills (Phys HOTS). In addition, the scoring models used unfair because not using a model that considers politomous steps to resolve the matter.

Based on the above, to measure the learning outcomes of students not only lower order thinking skills, but also higher order thinking skills in Physics.This required a model of the instrument in the form of an enhanced objective choice hereinafter called reasoned objective test. Thus it is necessary to research on the application of the test to the test and scoring objective reasoned with PCM. The main issue to be raised in this study is the usefulness and effectiveness of reasoned choice objective test to measure higher order thinking skills (HOTS) in Physics (PhysROCTHOTS) implementing Curriculum 2013.

In line with the problem solved, then the purpose of this research was to test the usefulness and effectiveness of reasoned choice objective test to measure higher order thinking skills (HOTS) in Physics (Phys ROCTHOTS) implementing Curriculum 2013.

Research Method

Subjects of this research were all students of grade XI ofthe ten state senior high school in Yogyakarta. The number of research subjects as much as 1001 students. Determination of senior high school that is used as a subject for research based ranking of schools based on the value of the national examination in Physics of 2012 (low, medium, and high). The schools were used as research subjects, among others: SMAN 5 Yogyakarta, SMAN 11 Yogyakarta, SMAN 1 Bambanglipuro, SMAN 1 Sedayu, SMAN 1 Wates, SMAN 1 Pengasih, SMAN 1 Gamping, SMAN 1 Minggir, SMAN 1 Wonosari and SMAN 1 Patuk.

Data analysis of this research using the partial credit model 1 PL (PCM 1-PL) for testing the goodness of fit of learning outcomes test in Physics higher order thinking skills of senior high schools. Basic considerations are used, the first that the PCM as an extension of the Rasch model is a 1-PL models, can use the not large samples (Keeves & Masters, 1999: 12-13). Second, that the characteristic response to high order thinking skills item follow PCM.

Testing of the test usefulness used validity, reliability, total information function,and SEM. The test validity was obtained by experts judgment and goodness of fit to model. Goodness of fit of the test scoring in polytomousaccording to the partial credit model (PCM) to measure HOTS based on average of INFIT Mean of Square. Ifthe averageINFITMNSQof 1.0 and standard deviation of 0, then the test fit with the model (Adams & Khoo , 1996:30). The effectiveness of the test was indicated by practicality and efficiency based on response of senior high schools physics teachers.

Result and Discussion

1. Usefulness of PhysROCHOTS

Structure of PhysROCHOTS consists of stem, the option of stem, reason, and the option of reason. The number of options and choices stem reason each of the five, as Figure 1.

Results of an expert judgment eclared that PhysROCHOTSis valid to measure PhysHOTS. Testing of goodness of fit the overall test on the basis if the average INFITMNSQ about 1.0 and standard deviation of 0.0, then the test fit with the model PCM1PL. Based on Table1, the value of mean INFITMNSQ of 1.01 (about 1) and a standard deviation of 0.02 (about 0.0), then the PhysROCHOTS sharped reasoned choice objective test fit with the PCM1PL.

(stem)…….. ……
(option of stem) A….
B….
C…..
D. …
E …
(reason) ……
(option of reason) A….
B….
C…..
D. …
E …

Figure 1. Diagram of PhysROCHOTS Structure

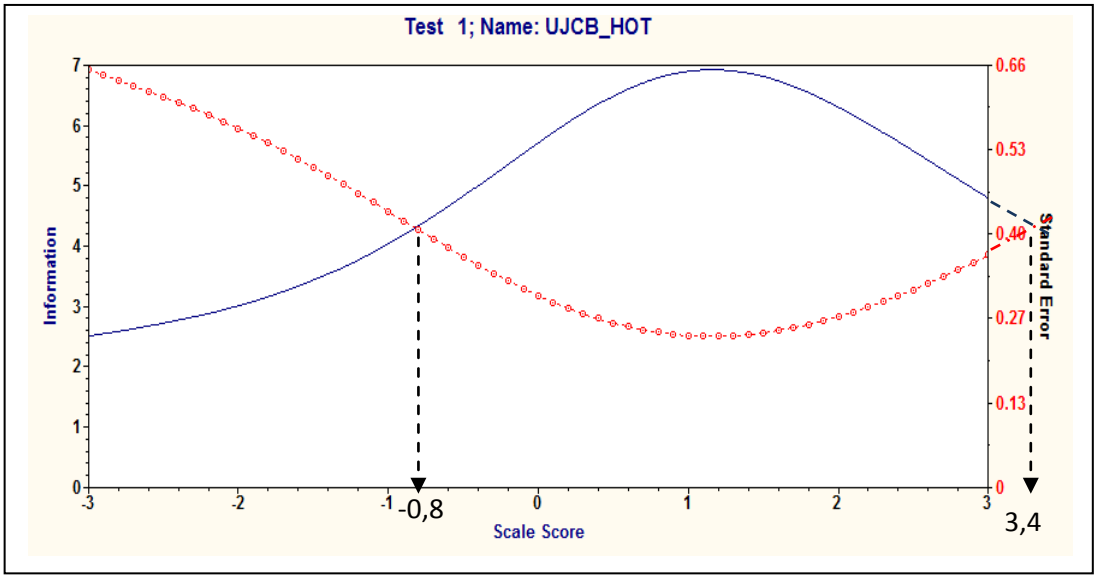

Figure 2. Total Information Functionand SEM

Based on the analysis it is estimated the test reliability of 0.95. Reliability coefficient is quite high category. The test has a reliability coefficient of at least 0.90 with the test results of the testing can be used to make decisions about individuals (Suryabrata, 2002:39-40). In addition, based on Figure 2 that the test is appropriate for measuring students' physics higher order thinking skills that value from -0.8 to 3.4. This is supported by Nitko & Brookhart (2011: 223) that multiple-choice tests can be to measure higher order thinking skills. Based on the experts judgment, the goodness of fit, the coefficient of reliability, graphs total information functionand SEM that the reasoned choice objective test can be used to measure Physics higher order thinking skills.

2. Effectiveness of PhysROCHOTS

Response of senior high school Physics teachers related to there as one objective choice test and the application of PCMin senior high school are presented in Table1.

Table1.
Response of senior high school physics teacher to test models and scoring models

Aspect	Statement	Percentage of teachers' response "Yes"
Suit ability of model and scoring	1. The use of model reasoned choice objective test to measure Physics higher order thinking abilities (PhysROCTHOTS) in senior high schoolis very precise	75
	2. Scoring model of PhysROCTHOTSin senior high school should be the politomous (score1, 2 ,3, 4)	85
	3. PhysROCTHOTSin senior high school scored according to the partial credit model (PCM) is right	70
	4. Reasoned choice objective test scored according to thePCMcan be applied to test the learning outcomes in physics	80
	Average	77.5
Practicality	5. To correct senior high school students' response of PhysROCTHOTS are flexible time because they do not have to finish each item for all students	85
	6. Scoring of PhysROCTHOTSinsenior high schoolis easy to implement	80
	7. Model of reasoned choice objective test for measuring PhysHOTSin senior high school can explore information of student's knowledge	75
	Average	80
Efficiency	8. PhysROCTHOTS can be applied to a large number of senior high school students	90
	9. To measure Physics higher order thinking reasoned choice objective test require one rater	85
	10. The cost of the correction model of reasoned choice objective test to measure senior high school students'PhysHOTS is cheap	75
	11. The correction time of reasoned choice objective test for measuring physics higher order thinking in senior high school is short	80
	Average	82.5

Based on Table1, reasoned choice objective test with scoring model of politomous four categories according to the PCM can be effectively implemented in senior high school. First, reasoned choice objective test is practical in using an scoring supported by 80 % of senior high school physics teachers. Second, the test is efficien tbecause it can be used on a large number of students with a relatively low cost. In addition,the correction time of response of the reasoned objective test is relatively short. That is supported by 82.5‘% of senior high school physics teachers. Third, results of students’ the exam assessment based on the stages can be completed students. Although only just completed the initial phase, participants had test scores. The highest value of course obtained when the examinee has completed all phases of the examin that clause. This is consistent with the statement Widhiarsa (2010: 6) that solve a problemonly to the first stage is analogous to thecategory of 'never', while when it comes to the final stage, analogous to the category of 'always on'. Thus, the objective reasoned choice physics test can measure higher order thinking skills effectively.

Conclusion and Recommendations

Conclusion

Based on the analysis, the conclusions are as follows:

1. The reasoned choice objective test (PhysROCTHOTS) is valid and reliable so the test is can be used to measure senior high school students’ Physics higher order thinking skills.
2. The PhysROCTHOTS scoring in polytomous according to PCM can measure senior high school students’ Physics higher order thinking skills implementing curriculum 2013 effectively.

Recommendations

Based on the analysis, it is suggested that Physics teachers of senior high school should measure higher order thinking skill implementing curriculum 2013 with the reasoned choice objective test.

References

Adams, R. J. & Khoo, S. T. (1996). *Quest: The interactive test analysis system version 2.1*. Victoria: The Australian Council for Educational Research.

Anderson, L.W., &Krathwohl, D.R. (2001).*A Taxonomy of Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives*. New York: Longman.

Barnett, J. E & Francis, A.L. (2012).*Using higher order thinking questionsto foster critical thinking: a classroom study.* Educational Psychology: An International Journal of

Experimental Educational Psychology. http://www.tandfonline.com/loi/cedp20. Diakses tanggal 10 Desember 2012

Brookhart, S. M. (2010). *How to Assess Higher Order Thinking Skills in Your Classroom.* Alexandria: ASCD

Kemendikbud. (2013). *Peraturan Menteri Pendidikan dan Kebudayaan No 66 Tahun 2013 tentang Standar Penilaian*

----------. (2013). *Peraturan Menteri Pendidikan dan Kebudayaan No 69 Tahun 2013 tentang Kurikulum SMA-MA*

Djemari, Mardapi. (2012). *Pengukuran, Penialaian, dan Evaluasi Pendidikan*. Yogyakarta: Nuha Litera.

Keeves, J.P. & Masters, G.N. (1999). Introduction. Dalam: Masters, G.N. & Keeves, J.P. (eds.). *Advances in Measurement in Educational Research and Assessment* (pp.1-22). Amsterdam: Pergamon, An imprint of Elsevier Science

Sudjana, N. (1990). *Penilaian Hasil Blajar Mengajar.* Bandung: PT Remaja Rosdakarya

Nitko, A.J & Brookhart, S. M. (2011). *Educational Assessment of Students.* (6thed). Boston*:* Pearson Education, Inc.

Piaget, J. (2005). *The Psychology of Intellegence* [Versielektronik].Taylor & Francis e-Library.

PusatKurikulum. (2007). *NaskahAkademikKajianKebijakanKurikulum Mata Pelajaran IPA.* Jakarta: BalitbangDepdiknas

Reedal, K.E. (2010).*Jean Piaget's Cognitive Development Theory in Mathematics Educatio*n. Department of Mathematics and Computer Science – Ripon College. Summation, May 2010, pp. 16-20 http://ripon.edu/macs/summation.

Schraw , G, &Robinson, D.H. (2011). *Assessment of Higher Order Thinking Skills*. New York: Information Age Publishing, Inc

Suryabrata, S. (2002).*Pengembangan alat ukur psikologis.* Yogyakarta: Andi Offset

Van den Berg, G. 2008. The use of assessment in the development of higher-order thinking skills. *Africa Education Review*, 1:2, 279-294. Diambil Tanggal 24Oktober 2012 dari http://dx.doi.org/10.1080/18146620408566285.

Van der Linden, W. J &Hambleton, R. K. (1997). *Handbook of Modern Item Response Theory* . New York: Springer-Verlag New York, Inc

Widhiarso, W. (2010). *Model politomi dalam teori respons butir.* Yogyakarta: Psikologi UGM

CERTIFICATE

Ref : 8247/UN34.17/TU/2014

This is to certify that

Dr. Edi Istiyono, M.Si.

has participated in

INTERNATIONAL CONFERENCE ON EDUCATIONAL RESEARCH AND EVALUATION (ICERE) 2014

organized by Graduate School, Yogyakarta State University on November 8 - 9, 2014

as a: **PRESENTER**

with the paper entitled

EFFECTIVENESS OF REASONED OBJECTIVE CHOICE TEST TO MEASURE HIGHER ORDER THINKING SKILLS IN PHYSICS IMPLEMENTING OF CURRICULUM 2013

Yogyakarta, November 9, 2014

Director

Prof. Dr. Zuhdan Kun Prasetyo, M.Ed
NIP.19550415 198502 1 001

Chairperson

PANITIA ICERE
PEP PASCASARJANA UNY
YOGYAKARTA

Prof. Dr. Sudji Munadi
NIP. 19530310 197803 1 003