

Authentic Assessment in Assessing Higher Order Thinking Skills

Razmawaty Mohamed¹ and Othman Lebar²

Faculty of Education and Human Development^{1,2},
Sultan Idris Education University, Perak, Malaysia

DOI: 10.6007/IJARBSS/v7-i2/2021 URL: <http://dx.doi.org/10.6007/IJARBSS/v7-i2/2021>

Abstract

Higher order thinking skills (HOTS) is an important aspect in Malaysian education system. This article is based on a literature review conducted on studies related to assessments specifically on authentic assessments and higher order thinking skills. This article covers concepts and definitions of authentic assessments and higher order thinking skills. Authentic assessments are a form of alternative assessment. The methodology used for this literature review was adapted from the Systematic Guide to Literature Review by Okoli & Schabram. The results of this literature review shows that authentic assessments have potential to measure higher-order thinking skills among students.

Keywords: Assessment, Higher order thinking skills, Authentic Assessment

Introduction

Achievement of students in international assessments such as the Programme for International Student Assessment (PISA) 2009 shows that Malaysia's ranking is at the 57th in Mathematics while 53rd in science. The results for 2012 show an overall ranking is at number 52, which is still below the average score of the Organization for Economic Co-operation and Development (OECD). For the assessment of Trends in International Mathematics and Science Study (TIMSS), the achievements of Malaysia has declined between 1999 and 2007 (Malaysia Education Blueprint, 2013). Achievement in PISA reflects students' ability in analytical reasoning, application, and continuous learning while TIMSS measures cognitive skills such as thinking, reasoning and apply. As a result from the less favorable achievement in the both international assessment (Abdullah & Peters, 2015), the assessment of higher order thinking skills (HOTS) has been addressed by entire community in Malaysia.

Malaysian Education Blueprint (MEB) 2013-2015 was designed to focus on higher order thinking skills (HOTS) (Malaysia Education Blueprint, 2013). To improve the quality of international education, examinations and assessments redesigned to increase the focus on HOTS. Among the factors that are said to affect the ability of students in HOTS is the effectiveness of assessment practice. This refers to the extent to which the assessment practiced is really measured HOTS. The assessment practices include forms of assessment, items used, questioning techniques, and scoring rubrics. A study carried out by Suah See Ling, Lan, and Osman (2014) shows that school teachers use a lot of multiple choice questions in the classroom assessment. Coombs, Milholland, & Womer (1956) argues that the weakness of

multiple choice item is the discriminating against students' answers whether the selected answers give the impression that students have partial information, incomplete information or incorrect information. Multiple choice item does not stimulate students' thinking when they guessed the answer without thinking. According to Coombs et al., the multiple choice item encourages guesswork. However, Goodson & Soul (1998) stated that HOTS can be measured using a variety of items including multiple-choice item, matching, and essay. Based on a literature review conducted, forms of assessment practiced and the items used play important role in improving students' ability in HOTS. Strengths and weaknesses of items used should be identified in accordance with the strategies of teaching and learning objectives and stimulate HOTS among students. Yin Peen & Yusof Arshad (2014) in his study also suggested that the items used in the assessment must be able to measure HOTS. Indirectly, it can be said that appropriate assessment methods should be practiced to increase HOTS.

Readiness and skills of teachers in teaching and learning (T&L) that implement elements of HOTS also affect the students' ability in HOTS as described in the study by Suhaili (2014). The study found that teachers assume that they have implemented HOTS in teaching but the teaching strategies practiced does not indicate an effective integration. Suhaili also suggested that professional development for teachers should be carried out to promote the implementation of HOTS. It is also aimed to enable teachers to better understand and implement the concept of HOTS in teaching and learning. Nagappan (2001) in his study also recognizes that the knowledge, skills, and behavior of teachers are the important factors that need to be addressed in the application of HOTS among students. Both of these studies indicate that the role and teachers' skills are also factors that affects the ability of students in HOTS.

This concept paper aims to explore about authentic assessment and how it can measure higher-order thinking skills (HOTS) among the students. Therefore, this paper will also discuss the literature review conducted to clarify the concept of both HOTS and authentic assessment by obtaining overall and specific views related to the focus of study.

Authentic Assessment

Assessment can basically be classified into traditional assessment, performance-based assessment and authentic assessment (Whitlock & Nanavati, 2013). Traditional assessments emphasize what students know and lead to the measurement of knowledge to gain skills. In contrast to the performance-based assessment, it is more focused on what can be demonstrated by students and measure their skills in a particular scenario. Whitlock & Nanavati also said that authentic assessments emphasize what students can do in the context of real life, and tend to measure students' skills when it's needed in real life. They also concluded that based on the educational needs of the current focus on measuring student competency, performance-based assessment and authentic assessment seeks to provide the best information about what can be done by a student.

According to Gulikers, Bastiaens, & Kirschner (2004), authentic assessments are assessments that require students to use the competence, or a combination of knowledge, skills and behaviors that should be applied in their professional life situations. Authentic nature is subjective, therefore it is important for the perception of authentic assessment that will

influence learning. A simple definition of authentic assessment described by Frey (2014) is the assessment that in line with tasks and requirements in real life. Frey also believes that an assessment is said to be authentic when assignments, content, requirements and evaluation methods for the assessment is the same with the real world needs. According to the study, an assessment is said to be more authentic when it becomes similar to real life situations. Authentic assessment involves realistic dimension, student involvement and diversity scoring component. Gulikers (2006) stated that authentic assessments are able to affect students' learning. Indirectly, authentic assessment is formative assessment and assessment for learning. Additionally, the right skills for authentic assessments involve procedural knowledge, problem solving, collaboration and motivation (Borich & Tombari, 2004). According to Mueller (2005), the authentic assessment is a form of assessment that occurs when students are asked to perform tasks that demonstrate the use of knowledge and skills that are meaningful. Through authentic assessments, knowledge, understanding, problem solving skills, social skills and behavior are used or practiced in real life (Aitken & Pungner, 2005). Aitken & Pungner also said that the role of authentic assessment in the classroom can be enhanced when the belief that education is not only involved skills but also critical thinking and application of knowledge. Authentic assessments were defined through various approaches. Nine characteristics of authentic assessments described by Frey is classified into three main categories: assessment context, the role of students and scoring as shown in Figure 1.

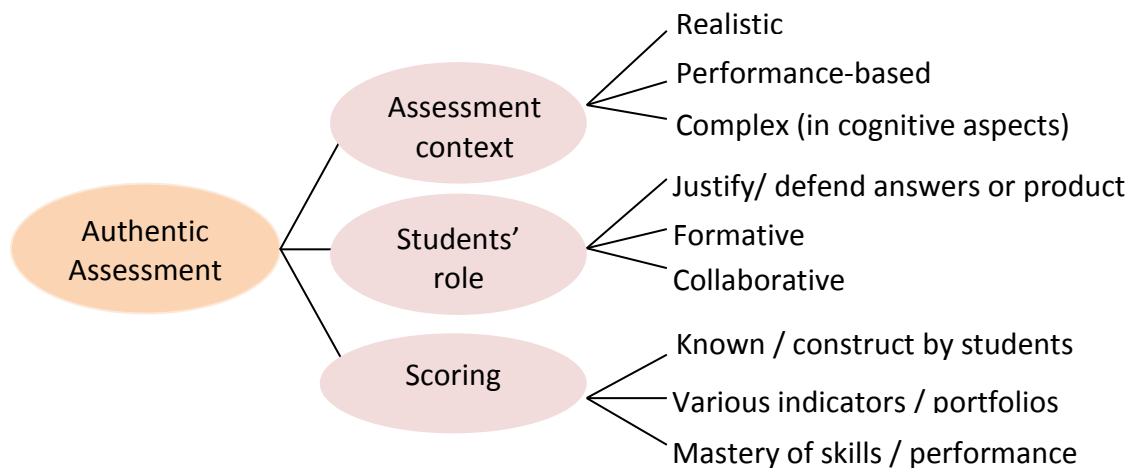


Figure 1. Characteristics of Authentic Assessment

For an authentic assessment, activity or context involved is realistic which is referred to the actual situation in life. Tasks in authentic assessments are based on performance and complex in cognitive aspect. Through authentic assessments, students play their role as they are given opportunities to justify and defend their answer or product. Assessment carried out is a form of formative assessment. In other words, the assessment takes place continuously throughout the

learning period. Formative assessment requires students to collaborate with colleagues and teachers. In terms of scoring, the criteria that must be achieved in an authentic assessments are known by the students and they also allowed to develop scoring criteria based on the assignment. Various instructions and portfolio can be used for scoring. Scoring in authentic assessments enables students to master the desired performance.

Studies conducted by Azmah et al. (2014) also proposed the practices of authentic assessments in schools. Definition and characteristics of authentic assessments indicated in previous studies clearly show that authentic assessment is a process that not only conducted to obtain information on students' knowledge, but it also involves skills acquired and subsequently applied by the students themselves in real life. Indirectly, through authentic tasks, cognitive ability and higher order thinking skills in students can be improved.

Authentic assessment is not a new form of assessment but it has been widely adopted in education, including in other countries. A study was conducted to investigate the level of development in the use of the teacher education curriculum and assessment strategies in order to prepare teachers to meet the challenges of education. The study showed increases in teacher education programs using authentic assessments as a tool to help novices (Darling-Hammond & Snyder, 2000). According to Darling-Hammond & Snyder, authentic assessment practices in teacher education not only affects the learning of teachers themselves but also affect student learning. This means that authentic assessments directly act as an assessment for learning because of its impact not only on the students but also the teachers' teaching itself.

Higher Order Thinking Skills

Thinking skills have been given attention since the days of Socrates and known as critical thinking that includes creative thinking (Ariffin et al., 1989). According to Ariffin et al., thinking skills are defined as the intellectual process that involves the formation of the concept, application, analysis or evaluating information gathered through observation, experience, reflection, reasoning, or communication. Thinking skills is the process of using the mind to make decisions and solve problems. There are two levels of thinking; lower order thinking skills (LOTS) and higher order thinking skills (HOTS).

Various definitions of HOTS have been provided in the field of education. HOTS cover all intellectual activity that is not just recall information (Baker, 1989). According to Goodson & Rohani (1998), higher-order thinking involves the application of various thought processes in complex situations. Higher-order thinking for an individual depends on the individual's ability to apply, develop and enhance knowledge in the context of thinking. HOTS includes cognitive levels of analysis, synthesis and evaluation and mastery in applying the routine things in new and different situations (McDavitt, 1994). According to Goodson & Rohani also, higher order thinking skills involve critical thinking, logical, reflective, metacognitive and creative which these skills results when students at various ages resolve questions and problems that are not routine, and overcome a dilemma. These studies demonstrate higher-order thinking can be achieved when an individual find solution and subsequently make decisions regarding a situation that is uncommon. In other words, HOTS is stimulated when an individual applies the knowledge or skills to solve problems in the context of everyday life.

HOTS is also defined by three categories, namely transfer (information or knowledge), critical thinking and problem-solving (Brookhart, 2010). The first category is when the HOTS is defined on the basis of the transfer or exchange of knowledge. Through this concept, higher-order thinking occurs when students not only remember but justify and apply what they have learned (Anderson & Krathwohl, 2001). Based on aspects of critical thinking, HOTS was defined as a reasonable idea, and reflective thinking that is focused on deciding what to believe or do (Norris & Ennis, 1989). Higher-order thinking based on the concept of problem solving is defined as one or more of a thought process that is used by a student to achieve a goal or make a solution for a problem (Nitko & Brookhart, 2007). Based on the three definitions, it can be concluded that the overall higher-order thinking involves thinking processes to solve problems, applying the knowledge and reflective thinking.

The Ministry of Education has defined HOTS as the ability to apply knowledge, skills and values to make reasoning and reflection to solve problems, make decisions, innovate and strive to create something (Curriculum Development Centre, Ministry of Education, 2013). Based on the definition by the Curriculum Development Division, HOTS is focused to apply, evaluate, analyze and create. In the assessment aspect, HOTS items are the items that assess cognitive skills to analyze, evaluate and create. However, HOTS item can also measure cognitive level of application but involving new situation (Malaysian Examination Syndicate, 2013). Figure 2 shows the elements of assessment items to measure HOTS by Malaysian Examination Syndicate.

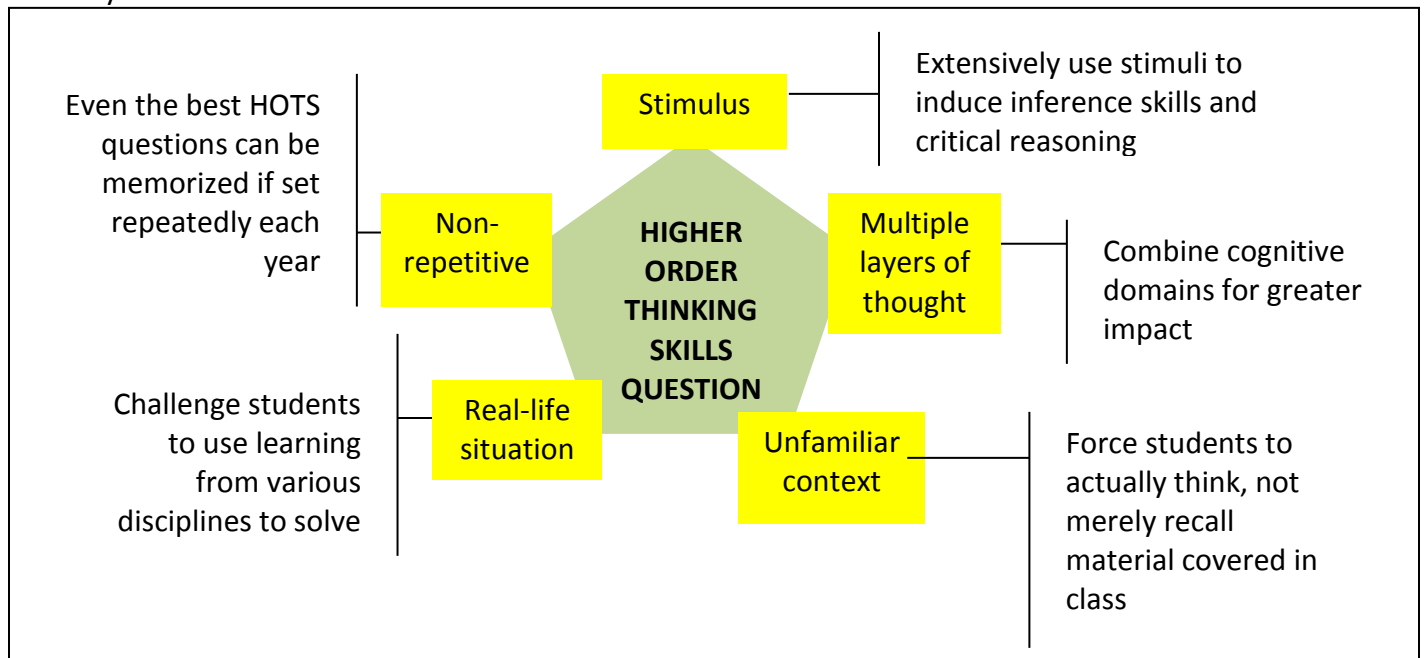


Figure 2 . The elements of assessment items to measure HOTS

Elements that are applied in assessing higher order thinking skills including the extensively use of stimulus, measure various levels of cognitive thinking, not a common context, based on the

real-life situation, and the non-repetitive item. The stimulus provided for a HOTS item must be functional and give enough information to generate inference skills and reasoning critically. HOTS items should also be able to obtain various and relevant response. The item also involves uncommon context which refers to new situation outside the classroom or out of the textbook to encourage students to think more deeply. The real situation in HOTS item is intended to allow students to apply a variety of disciplines to solve the problems in everyday life. HOTS item must not be repeated. This means the item is different each year and the material uses is beyond the textbook, workbook and etc. With reference to the definition and concept of HOTS from the previous studies, it can be said that the elements used by the Malaysian Examination Syndicate in item building to measure HOTS meet and match the concept of HOTS from the educational experts.

Methodology

According to Randolph (2009), the literature review allows the framework of a survey to be identified to link the old with the new research study. In this study, a literature review is conducted to identify the characteristics of authentic assessments and the elements to assess higher order thinking skills (HOTS). From the results of the literature review, a mapping is generated to identify the similarity or the suitability of authentic assessment with the elements of HOTS to formulate assessment tasks. In the early stage of research literature, reference materials are reviewed holistically based on the problem of the study, the conceptual framework, methodology, variables and findings of other studies. According to Idris (2013), the review of the literature can be used to justify the need for the study, rationalize the study, problem, purpose of the study and research questions. Therefore, the information obtained is used to strengthen the focus of the study.

The literature review begins with the specialization of topics to research keywords to find the locations of the study materials either through internet or library. The resources used for these studies including database ProQuest, Eric, British Education Index, Google Scholar and etc. Method of literature review conducted in this study was adapted from the Guidelines Systematic Literature Review by Okoli & Schabram (2010) which consists of eight steps. However, the adjustment was carried out in which the review of the literature has been devoted to the four main stages: planning, selection, extraction and execution. In the planning stage, the purpose and goals of the study are identified. In this context, the goal of the study is to identify the characteristics of authentic assessments and HOTS, and then identify the suitability of authentic assessment as a method of measuring HOTS.

In the first stage, the method and procedure for conducting the literature review is identified. Information about the previous study is obtained from the library. In the second stage, the journals, articles or information from the previous study are reviewed to select those related with this study. In this literature review, information, journals and articles related to authentic assessments and HOTS are selected with reference to the findings, methodology, variables and framework used. Rowley & Stack (2004) also stated that the evaluation of the source of information is an important aspect in the study of literature. Information related to the concept of authentic assessments and HOTS are then extracted. The information obtained is then

organized and combined to be used in this study. Figure 3 shows the method in the study of literature adapted from Guidelines Systematic Literature Review by Okoli & Schabram.

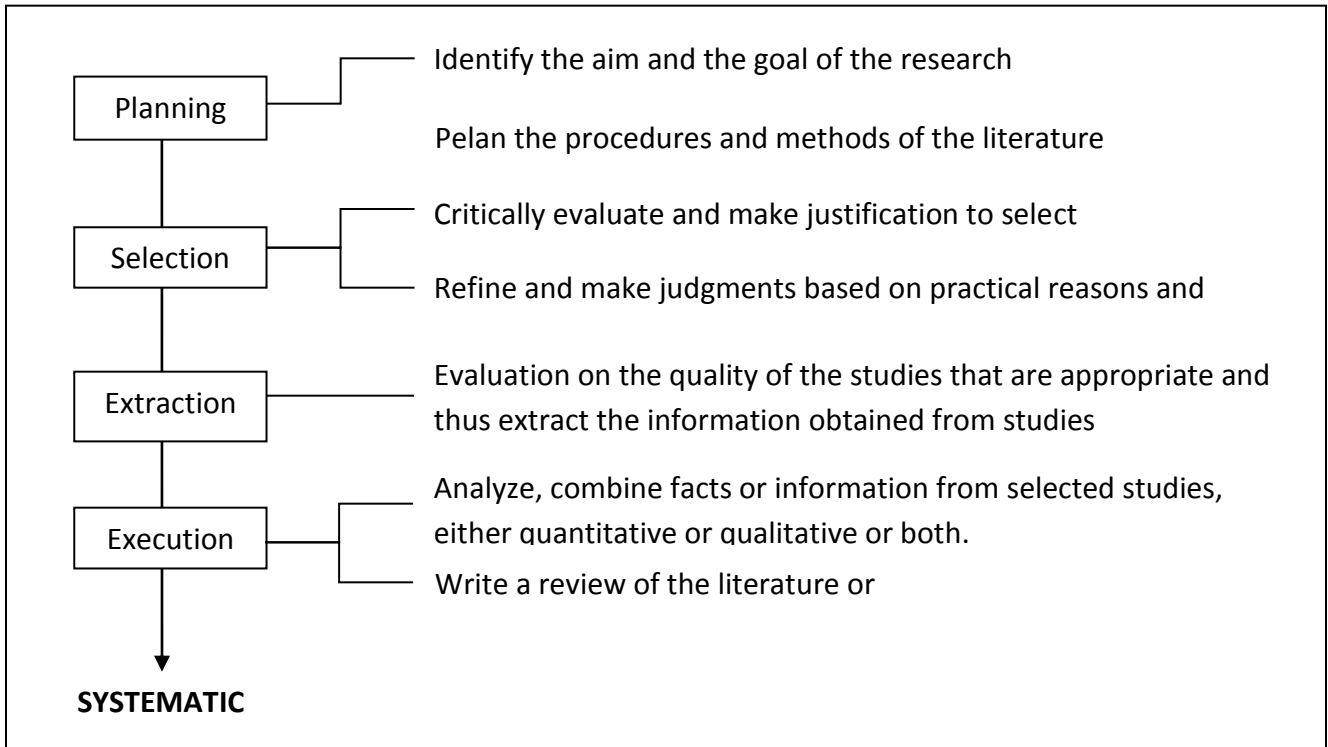


Figure 3. Method in the study of literature adapted from Guidelines Systematic Literature Review by Okoli & Schabram.

Discussion

The literature review conducted based on the Guidelines Systematic Literature Review allows the concept of authentic assessment and higher order thinking skills (HOTS) to be explored and identified clearly. Research on the articles and other resources such as journals, books, and etc. are focused on the characteristics of authentic assessment and its role in the assessment. The focus in this literature review is to identify the extent to which authentic assessment and HOTS have in common and suitable for assessing higher-order thinking skills. A deeper study of the articles and researches related to HOTS are focused on the concept and the extent to which it can be measured through the construction of an appropriate item. The definition of authentic assessments discussed in the previous studies and articles enable the identification of characteristics of authentic assessments . Table 1 presents a summary derived from the literature review based the definition of authentic assessments expressed by several researchers.

Study / Research	Context	Sudents' Role	Scoring
Frey, 2014	/	/	/
Whitlock & Nanavati, 2013	/	/	
Torrance, 2009	/		
Abernethie, 2006	/	/	/
Gulikers, 2006	/	/	
Mueller, 2005	/		/
Engel, Pulley, & Rybinski, 2003	/		
French, 2003	/		/
Maden & Taylor, 2001	/	/	
Ratcliff, 2001	/		
Jolly & Kettler, 2000	/		/
Lawton, 2000		/	/
Moorcroft, Desmarais, Hogan, & Bekowitz, 2000	/		/

Table 1 Summary Definition of Authentic Assessments.

Based on Table 1, almost all researchers define authentic assessments with reference to the context of the assessment. They believe that authentic assessments can measure various cognitive skills in a realistic situation. Indirectly, they agree that authentic assessments involves performance-based tasks. This means that authentic assessment is a method that allows students to demonstrate the knowledge and skills that have been achieved. In other words, authentic assessments measure how well students can apply what they know in real life. Most of the researchers define authentic assessments based on the role of students. Through authentic assessments, students are able to give reasons or justification for their answers or the products and thus encourage them in reasoning and problem solving. Through authentic assessments, students play a role in creating cooperation and relationship among friends and teachers. Most researchers define authentic assessments based on scoring elements produced. This indicates their agreement that the criteria provided in scoring are known by students and students have the opportunities to provide criteria appropriate to the tasks assigned. The using of various indicators and portfolios in authentic assessments shows the holistically method in scoring which is not limited to one form of the indicator only.

Based on the literature review, the meeting point of authentic assessments with elements to measure higher-order thinking skills (HOTS) is identified. Various cognitive levels for an item encourage students to use HOTS. It meets with the authentic assessments that involves complex task in terms of cognitive level. Various forms of assessment tasks will invite various response from the students and thus involves holistic scoring. Students should also be able to master a variety of skills and knowledge through various stages of cognitive assessments for an item. Items consist of stimulus which are used extensively will involve students' thinking on a range of cognitive levels and thus encourage HOTS. To measure HOTS, real-life situations and new contexts should be used in item building. Both of these features are also available in

authentic assessments. Basically authentic assessment is a comprehensive assessment of the various tasks that are appropriate and can measure multiple constructs. Elements of HOTS are subset of authentic assessments. This means, there is a characteristics or elements which are in common for the two concepts in this study. In other words, based on the literature review conducted, the common characteristics of authentic assessments and HOTS can be identified. Therefore, it can be said that authentic assessment has high potential to measure higher-order thinking skills.

Conclusion

The literature review conducted enables an analysis of previous studies and other sources about the concept in general and specific assessments. Through this research, more specific knowledge acquired and thus comparisons between form of authentic assessment and assessment of higher order thinking skills can be made. Based on the concept and previous studies of authentic assessments, it can be concluded that it is an assessment for learning and able to measure higher-order thinking skills. The meeting point that indicates the suitability of authentic assessment and higher order thinking skills can be used to formulate assessment tasks for the next study.

Acknowledgement

Acknowledgement to the biggest education grant of the Sultan Idris Educational University (UPSI), Niche Research Grant Scheme (NRGS): 2014-0001-107-82-3 for project 3: Assessment
Special acknowledgement to the Assessment Project Leader: Profesor Emeritus Dr. Othman Lebar

Corresponding Author

Razmawaty Mohamed; razma_raz@yahoo.com

Othman Lebar; othman@fppm.upsi.edu.my

Faculty of Education and Human Development,
Sultan Idris Educational University, Perak, Malaysia.

References

- Abdullah, A., & Peters, R. F. Malaysia's Post-PISA2012 Direction.
- Aitken, E. N., & Pungur, L. (2005). Literature synopsis: Authentic assessment. *Online at http://education.alberta.ca/apps/aisi/literature/pdfs/Authentic_Assessment_UofAb_UofL.PDF*.
- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon.
- Ariffin, S. R., Abd Majid, R., Haron, Z., Idris, R., Alias, A., Samsuri, S., ... Abdul Hamid, N. A. (1989). Kemahiran pemikiran kritikal dan penyelesaian masalah pelajar- pelajar sains di malaysia.
- Azmah, N., Yusuff, N., Rahaman, N. M., Abdullah, N., Ridzuan, P. D., Kebangsaan, S., ... Selangor, K. (2014). Primary school pupils' acquisition of science process skills via hands-on activities and authentic assessment 1, 3, 4(1), 15–28.
- Baker, E. L. (1989). Higher Order Assessment and Indicators of Learning.
- Blueprint, M. E. (2013). Blueprint 2013–2025. *Ministry of Education Malaysia*.
- Borich, G. D., & Tombari, M. L. (2004). *Educational assessment for the elementary and middle school classroom*. Prentice Hall.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- Coombs, C. H., Milholland, J. E., & Womer, F. B. (1956). The Assessment of Partial Knowledge 1. *Educational and Psychological Measurement*, 16(1), 13-37.
- Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and teacher education*, 16(5), 523-545.
- Frey, B. B. (2013). *Modern classroom assessment*. SAGE Publications.
- Goodson, L., & Rohani, F. (1998). Higher Order Thinking Skills• Definition• Teaching Strategies• Assessment. *Thinking*, 18, 458.
- Gulikers, J. (2006). Authenticity is in the eye of the beholder. Beliefs and perceptions of authentic assessment and the influence on student learning.
- Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational technology research and development*, 52(3), 67.
- Idris, N. (2010). *Penyelidikan dalam pendidikan*. McGraw Hill (Malaysia).
- Malaysia, K. P. (2013). Pentaksiran Kemahiran Berfikir Aras Tinggi. *Kementerian Pendidikan Malaysia*.
- McDavitt, D. S. (1994). Teaching for Understanding: Attaining Higher Order Learning and Increased Achievement through Experiential Instruction.
- Mueller, J. (2005). The authentic assessment toolbox: enhancing student learning through online faculty development. *Journal of Online Learning and Teaching*, 1(1), 1-7.
- Nagappan, R. (2001). Language teaching and the enhancement of higher-order thinking skills. *ANTHOLOGY SERIES-SEAMEO REGIONAL LANGUAGE CENTRE*, 190-223.
- Nitko, A. J. (2001). *Educational assessment of students*. Prentice-Hall, Inc., PO Box 11071, Des Moines, IA 50336-1071.
- Norris, S., & Ennis, R. H. (1989). Evaluating critical thinking: The practitioners' guide to teaching

- thinking. *Pacific Grove, California: Midwest Publication.*
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst, 10(26).*
- Peen, T. Y., & Arshad, M. Y. (2014). Teacher and Student Questions: A Case Study in Malaysian Secondary School Problem-Based Learning. *Asian Social Science, 10(4), 174.*
- Randolph, J. J. (2009). A guide to writing the dissertation literature review. *Practical Assessment, Research & Evaluation, 14(13), 1-13.*
- Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News, 27(6), 31-39.*
- Suah, S. L., & Ong, S. L. Shuki Osman.(2010). Pentaksiran pembelajaran pelajar: Amalan guru-guru di Malaysia. *Malaysian Education Dean's Council Journal, 5, 68-83.*
- Suhaili, A. B. H. (2014). *Exploring Teachers' Experiences on Integration of Higher Order Thinking Skills (hots) in Teaching of Science* (Doctoral dissertation, Universiti Malaysia Sarawak).
- Whitlock, B., & Nanavati, J. (2013). A systematic approach to performative and authentic assessment. *Reference Services Review, 41(1), 32-48.*