

Framework of Assessment for the Evaluation of Thinking Skills of Tertiary Level Students

Chan Swee Heng (Corresponding author)

Faculty of Modern Languages and Communication, Universiti Putra Malaysia
43400 Serdang, Selangor Darul Ehsan, Malaysia
E-mail: shchan@upm.edu.my

Yan Ziguang

Faculty of Modern Languages and Communication, Universiti Putra Malaysia
43400 Serdang, Selangor Darul Ehsan, Malaysia
E-mail: lee194591@126.com

Doi:10.7575/aiac.all.v.6n.5p.67

Received: 09/05/2015

URL: <http://dx.doi.org/10.7575/aiac.all.v.6n.5p.67>

Accepted: 18/07/2015

Abstract

In the 21st century, students are required to master thinking skills in order to deal with many situations that arise in the tertiary environment which later would translate into the workplace. Nowadays, thinking skills play a vital role in tertiary education. To provide an approach for teachers, this paper identifies a 4-step model that can be implemented in test design with reference to Bloom's Taxonomy for thinking skills. This model illustrates a feasible procedure of test construction built upon existing resource. The procedures can easily be applied to different fields of study. It provides an appropriate way for teachers who may want to add more ideas to their repertoire skills in addressing the learning of HOT (higher-order thinking) skills.

Keywords: thinking skills, Bloom's Taxonomy, 4-step model

1. Introduction

Our rapidly changing and challenging world requires students to go beyond the building of their knowledge capacity; they need to develop their higher-order thinking skills, such as critical system thinking, decision making, and problem solving (Barak, Ben-Chaim, & Zoller, 2007). Over two thousand years ago, Socrates identified intellect as having the ability to argue, question and discuss the concepts of knowledge, virtue and be able to understand new ideas and uncover new thoughts (Govier, 1997). It's known that an intellect should think analytically and critically. Critical thinkers are not only mere accept knowledge but also organize, process and utilize. On contrary, passive thinkers suffer a limited view and merely focus on answering yes or no.

Over the decades, the aim of developing students' higher-order thinking (HOT) has been a major educational goal (Fisher, 1999; Marzano, 1993; Zohar & Schwartz, 2005). In Malaysia, the steady increasing influence of thinking skills in education is prominent (Yen & Halili, 2015). The Minister of Education has suggested that the education system will be revised to encourage rational and analytical thinking (Rajendran Nagappan, 2001). He also suggested that the capacity for logical rational thought, reasoning and critical thinking are essential for all students. In 2003, The National Higher Education Research Institute conducted a study among graduates, the findings showed that 561 unemployed respondents overrated themselves by believing that they are well qualified and met all requirements of the regular job market (National Higher Education Research Institute, 2003). Is the real situation ideal as reported? According to the Department of Statistics Malaysia's report (2011), many prospective employers complained that local graduates are viewed as technically proficient but lack of communication and analytical skills. From the aspect of employers, unemployment is the most obvious result that caused by lacking of thinking skills. And the trend of unemployment rate is increasing these years (The Malaysia Statistics Department, 2011). Back to school, a report of Malaysia education has noted that most lessons in schools haven't sufficiently engaged students in constructive thinking where teachers relied on lecture format, furthermore, the learning focus was still directed at recalling facts or achieving surface-level content understanding rather than cultivating higher-order thinking (Ministry of Malaysia Education, 2012). In short, lower-order thinking, instead of higher-order thinking, still leads teaching methods and learning outcomes in Malaysia that can cause serious issues for graduates. Due to this issue, students need a clear guide, equally, lecturers, especially those who are new to the profession, also need a guideline to direct their teaching. Therefore, a workable and effective classroom assessment could be an optional choice (Stiggins, 2005).

2. Higher-Order Thinking Skills

For decades, the promotion of students' thinking skills has been the focus of educational studies and programs (Boddy, Watson, & Aubusson, 2003). Higher-order thinking (HOT) skills is an important aspect in teaching and learning especially at higher education institutions (Heong, Yunos, & Hassan, 2011). HOT is defined as the expanded use of the

mind to meet new challenges (Rajendran, 2008). It requires someone to apply new information or prior knowledge and use the information to get the possible answer in new situation (Lewis & Smith, 1993). Successful applications of HOT skills can be automatic. They are activated when individuals encounter unfamiliar problems, uncertainties, questions, or dilemmas (Goodson & Rohani, 1998). In higher education, students are required to utilize HOT skills thinking to make decisions, solve problems, and know how to learn. All students are capable to think, but most of them need to be encouraged and assist to think logically and analytically. The goal of teaching is to enable students to make well-reasoned, thoughtful, and healthy lifelong decisions (Florida Department of Education, 1997).

In Malaysia, the importance of HOT in tertiary education has been noted (Indramalar, 1997). The former Director-General of Education Datuk Matnor Daim stressed that lecturers should make their responsibility to shape students into thinking leaders (Rajendran, 2001). Therefore, lecturers should be fully-prepared to understand both higher and lower-order thinking (LOT) skills. Although there are many definitions of thinking skills based on different programs, either focus on “cognitive skills” or “thinking”, they all distinguish between higher-order and lower-order skills (Barak et al., 2007). Framed in more traditional terms, HOT corresponds with the taxonomy of Bloom (1956), which is typically designed lower-order thinking (LOT) and higher-order thinking (HOT) skills. HOT skills is defined including critical, logical, reflective, metacognitive, and creative thinking (Goodson & Rohani, 1998). According to Bloom’s (1956) taxonomy, “thinking” as a “roof”, which encompasses various different forms of thinking skills, such as critical, systemic, and creative thinking. Specifically, Bloom’s (1956) taxonomy includes six thinking levels and/or skills: knowledge, comprehension, application, analysis, synthesis and evaluation. In this taxonomy, skills involving analysis, evaluation and synthesis (creation of new knowledge) are thought to be a higher order involving the learning of complex judgmental skills such as critical thinking and problem solving (“Higher-order thinking,” n.d.). And the rests skills : knowledge, comprehension, application are defined as LOT skills (“Higher-order thinking,” n.d.). Bloom’s (1956) taxonomy provides a quite clear guide for lecturers.

3. Methodology

Traditional teaching can not meet the ever-changing world’s need. The 21st century explicitly requires teachers try to change their teaching strategies by shifting the emphasis from the traditional textbook-based, rote learning, to exploration, inquiry-based learning situated in real-world phenomena. Content delivery is a venerable and popular lecture format in higher education, but it seldom encourage students to study actively or think critically (Duron et al., 2006). Lecturers need a practical approach to make study actively. To reassign student’s enthusiasm of study, classroom assessment information is the essential fuel (Stiggins, 2005).

This paper identified a 4-step interdisciplinary model of assessment that can be conducted in any teaching or training setting to increase students’ thinking skills.

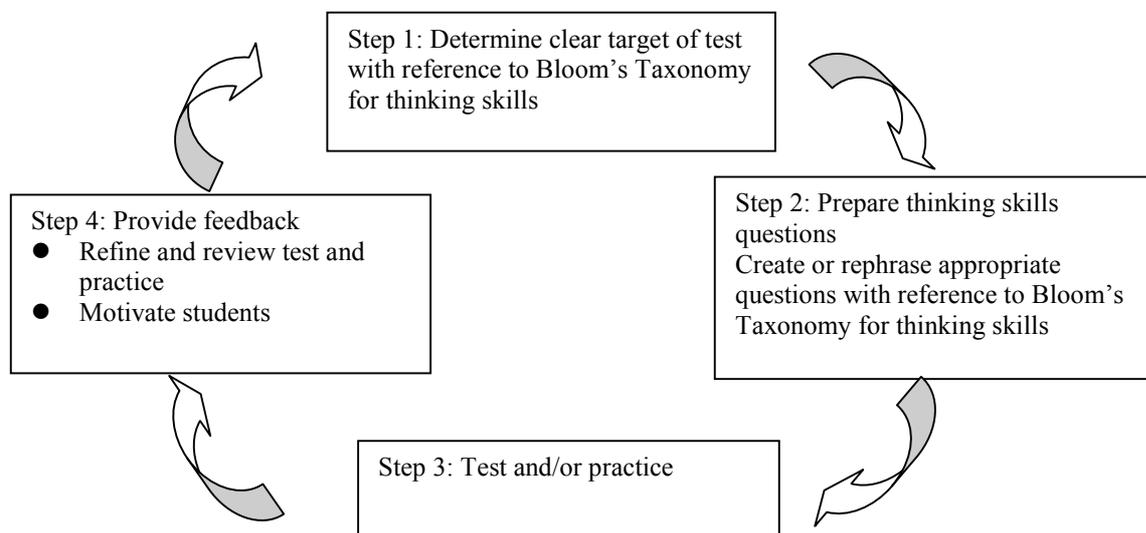


Figure 1. 4-Step Model of Thinking Skills Assessment

3.1 Step 1. Determine clear target of test

Elder and Paul (1997) proposed that the vital of questioning is the art of learning: if the question failed to set clear target, students wouldn't take the content seriously. A genuine question should set a clear goal. To test thinking skills, Bloom's(1956) taxonomy have been by educators to define or guide the development of assessments: tests and other evaluations of student learning, curriculum: units, lessons, projects, and other learning activities, and instructional methods such as questioning strategies (The Glossary of Education Reform, 2015).

Bloom's (1956) taxonomy straightforward defined different level's objectives that are appropriate for test and teaching. This taxonomy can be useful and helpful for course and assessments design because the different levels can help teachers move students through the process of learning—from the most basic level of *Knowledge* to the more complex

Evaluating (Northern Illinois University, 2010). In the lowest level, *Knowledge* level requires basic recall of facts and data. Students could be asked who, when, where. *Comprehension* level requires students to demonstrate understanding of the information. This level could ask students to summarize, explain, paraphrase, compare and contrast. *Application* require students to take a concept and apply it in a new or hypothetical situation. Students may be asked to apply, construct and show. The upper level, *Analysis* requires students to show ability of breaking something into component parts, seeking patterns, classifying, categorization and analysis. *Synthesis*, students may be asked to exhibit an ability of presenting items or thoughts together to create new or in new ways. In this level, students may be tested to combine, create, role-play and suppose. The last level is *Evaluation*, it requires students to judge/argue evidence based on established criteria/argument. Students may be tested to assess, criticize, predict and evaluate. Bloom's (1956) Taxonomy can provide an accessible guide for lecturers to scaffold well-written test items which access both LOT and HOT skills.

3.2 Step 2. Illustrative example of rephrasing appropriate questions

In testing thinking skills, Goodson and Rohani (1998) proposed that short answers are one of more effective ways to uncover students' thinking skills in tests. Revision existing resources is one of approaches to get short answers questions. In an effort to demonstrate the revision work of test items, a mid-2010 MUET (Malaysian University English Test) reading test paper was randomly selected among the MUET papers in the last five years. The paper includes five articles and 43 multiple-choice questions. MUET is an English language proficiency test, largely for university admissions (Malaysian Examinations Council, 2008). In this illustration, there are three ways that can be used to modify the MUET reading test paper into one that will show the elicitation of thinking skills that are over the different domains of the Bloom's Taxonomy. The first approach involves the examination of the question to evaluate its potential of use. From there, it is then matched to a level of thinking skill that is identified. For example in figure 2, Q2 is a LOT MCQ (multiple choice question) from the test which can be turned into a HOT synthesis question which demonstrates the compilation of information. The approaches are presented in Figure 2.

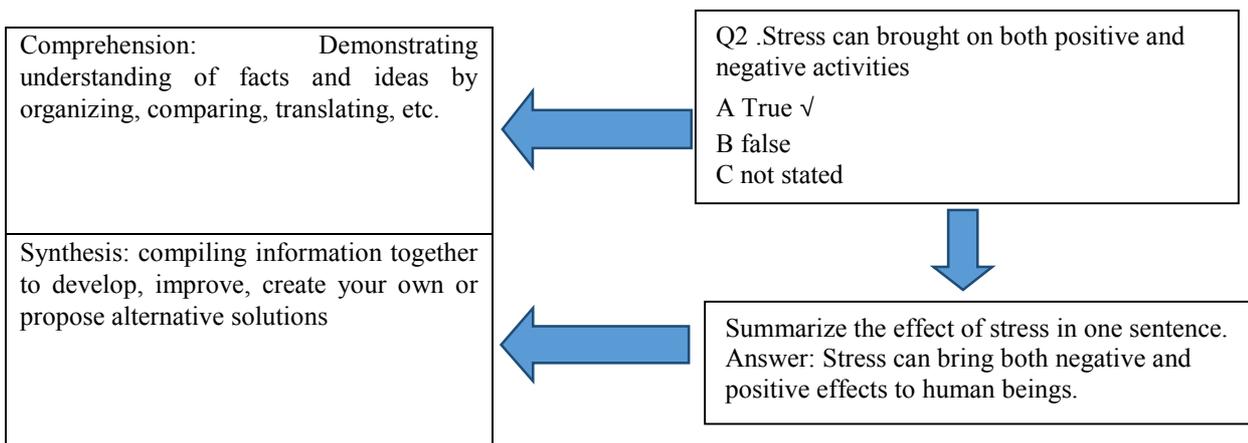


Figure 2. Revision of LOT MCQ into HOT Written-response Questions Based on Bloom's Taxonomy

As can be seen, the students need to put information together to develop their own answer. Lecturers can adjust the ratio of LOT and HOT questions based on their testing specification. The next possibility is to turn another MCQ which can fall into LOT level which requires students to use more language in their answers.

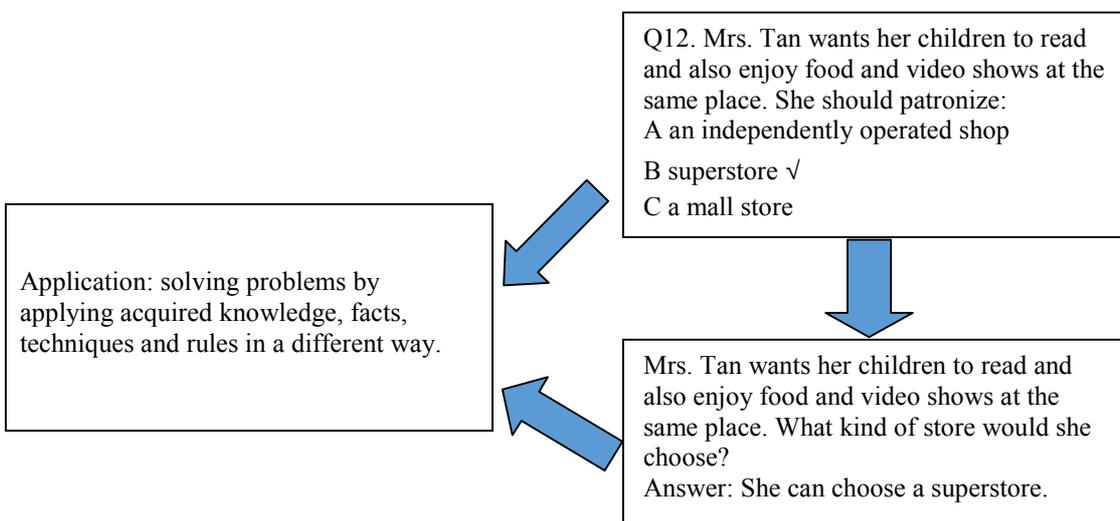


Figure 3. Revision LOT MCQ into LOT Written-response Questions Based on Bloom's Taxonomy

In the revised question, the word “patronize” was replaced by “choose” to enable students to understand the task clearly. After completing the questions, an expert panel can moderate them and provide comments. In this way, quality items are ensured.

Providing short answers requires the use of a scoring method that needs to be devised to include appropriated descriptors to show the level of the attainment. To do this, an analytic scoring method is used as it has obvious advantages (refer to Wang, 2009 for more details on analytical scoring). It helps to analyze attainment according to specific descriptors such as Clarity, Depth of reasoning, Grammar and Vocabulary.

Table 1. The example of scoring scheme of reading comprehension test

Thinking skills		A. Clarity	B. Depth of reasoning	C. Grammar	D. Vocabulary
		3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0
1.K	Define, label, list, name, memorize, order, recognize, recall, repeat, reproduce, state.				
2.C	Classify, describe, discuss, explain, express, indicate, locate, restate, select.				
3.App	Apply, choose, demonstrate, employ, illustrate, interpret, operate, use.				
4.Ana	Analyze, appraise, calculate, categorize, compare, contrast, and criticize.				
5.S	Arrange, assemble, collect, compose, construct, create, design, develop, plan, propose, set up.				
6.E	Appraise, argue, assess, attach, defend estimate, judge, predict, rate, core, support, value, evaluate.				

*K=Knowledge; C=Comprehension; App=Application; Ana=Analysis; S=Synthesis and E=Evaluation.

The descriptions of levels in A. Clarity, B. Depth of reasoning, C. Grammar and D. Vocabulary, see Appendix A.

Table 1 shows an example of the analytical scoring scheme use in this revision of reading test items. The scale is arrived at from recommendations about universal intellectual standards of thinking, advocated by Elder and Richard (2010). They stated that clarity (logic, consistency) and depth of reasoning (precision and depth of thought) are necessary to evaluate students' performance of thinking skills. Universal intellectual standards define the quality of reasoning about a problem, issue, or situation. This test also requires students to provide short answers and therefore grammar and vocabulary are included in this scoring scheme.

In line with testing procedures, a pilot test would be recommended. After the pilot test, the test items and /or scoring scheme can further be improved. The illustration of the application of the thinking skills framework for the construction of test items points towards how teachers can adapt and modify existing resources so that thinking skills can be brought into the classroom to match curriculum requirements. In the illustration, the passages would be aligned to those that would be of the require reading level of pre-tertiary students in Malaysia. As such, an accessible HOTS test paper can be easily constructed as a classroom resource. Although it may have some potential disadvantages, such as being the time consuming, it is believed that it is a feasible approach for teachers who may need more ideas to add to their repertoire skills in addressing the learning of HOTS.

3.3 Step 3. Conduct test and/or practice

These thinking skill questions can be arranged into different ways to test students. Besides paper-pencil test, short answer questions can be classroom practices. Fink et al (2003) provided an example: in reflective writing, students should address the following questions: What am I learning? What is the value of what I am learning? How am I learning? What else do I need to learn? These allow students' experiment with ideas, to develop concepts and to integrate concepts into systems. As such, even it is not a formal paper-pencil test, short answer questions also positively affects students to develop their own evaluation or interpretation by using thinking skills.

3.4 Step 4. Provide feedback

Immediate feedback is an effective way to develop thinking skills after assessment. Dihoff, Brosvic and Epstein (2003) quoted that in the early 1960's, Brackbill and her associates noted that delayed feedback across brief intervals promoted the retention of meaningful material (Brackbill, Bravos, & Starr, 1962). This outcome uncovered that when feedback was delayed for one to two days and retention intervals were lengthened to seven days (Dihoff, Brosvic, & Epstein,

2003). Therefore, it is necessary for lecturers to return an effective feedback in time. In paper-pencil test, marks and reflective comments can be provided instantly after grading. To give an effective feedback, lecturers should aware that the purpose of feedback is to enhance the quality of student learning and performance, rather than to get the marks. Feedback allows the lecturers and student(s) to conduct a dialogue about what distinguishes successful performance from unsuccessful performance as they discuss criteria and standards (Fink, 2003). Therefore, lecturers should be thoughtful and purposeful when providing feedback after assessments. They should provide informational feedback rather than controlling, based on agreed-upon standards, specific and constructive, quantitative, prompt, frequent, positive, personal, and differential (Wlodkowski & Ginsberg, 1995). The feedback can indicate students' further academic aim and drive students' learning motivation. Besides lecturers, student peers also can provide feedback and evaluation. Involving in providing feedback can mobilize the students' enthusiasm. Even students play a small role in setting the (study achievement) target, the considerable motivation, achievement, and benefits can be gained (Stiggins, 2005). For example, the question is to evaluate author's opinion in an article, the lecturers can ask students to check each other's answers and evaluate. This method practices student to consider others' ideas and apply HOT skills to evaluate.

In variety classroom activities or quiz, lecturers should monitor these closely. To track student's participation, Duron et al (2006) proposed that a teaching diary can be kept that identifies the students that participated, describes the main class activities, and provides an assessment of their success. This diary also can provide initial information when revising or updating instructional activities and questions based on thinking skills. On the other hand, students' feedback should be considered as a reference to improve assessment and teaching. Duron et al (2006) agreed that: students are asked to identify the most important point learned. Lecturers can review the comments and use them in future classes to emphasize issues identified.

4. Conclusion

Today's rapidly changing world increasingly requires cognitive skills to be mastered. As such, thinking skills are vital in order to deal with many situations that arise in the tertiary environment which later would translate into the workplace. Mastering HOT skills at the tertiary level is essential to meet employers' continual demands for independent, problem-solving employees (Allen & Wern, 2011). To meet the demands of the new century, lecturers and students got to take on the challenges, which require them to handle new and unexpected situations. The 4-step model illustrates a simple procedure of test construction that focuses on LOT and HOT skills.

In the illustration, MUET is demonstrated to show a manner of modification that rests on existing resources. The procedures can easily be applied to other fields of study. Thinking skills will motivate students to seek out answers on their own, particularly when they are required to provide short or extended answers. The effort on devising a scoring method also shows how assessment must encompass a reasonable and appropriate marking scheme that can offer feedback to students so that there is motivating wash back effects.

References

- Allen, S. H. L., & Wern, L. E. (2011). Involvement Of Higher Order Thinking Skills Within A Preparatory Course For The Malaysia University English Test. *The English Teacher*, *XL*, 95–112.
- Barak, M., Ben-Chaim, D., & Zoller, U. (2007). Purposely Teaching for the Promotion of Higher-order Thinking Skills: A Case of Critical Thinking. *Research in Science Education*, *37*(4), 353–369. doi:10.1007/s11165-006-9029-2
- Brackbill, Y., Bravos, A., & Starr, R. H. (1962). Delay-improved retention of a difficult task. *Journal of Comparative and Physiological Psychology*, *55*, 947–952.
- Dihoff, R. E., Brosvic, G. M., & Epstein, M. L. (2003). The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited. *The Psychological Record*, *53*, 533–548.
- Duron, R., Limbach, B., & Waugh, W. (2006). Critical Thinking Framework For Any Discipline, *17*(2), 160–166.
- Education Blueprint 2013-2025. (2012).
- Fink, L. D. (2003). *A self-directed guide to designing courses for significant learning*. Retrieved March 12, 2014, from http://www.byu.edu/fc/pages/tchlrnpages/Fink/Fink_Article.doc
- Fisher, R. (1999). Thinking skills to thinking schools: Ways to develop children's thinking and learning. *Early Child Development and Care*, *153*, 51–63. Retrieved from <http://ezproxy.um.edu.my:2077/doi/pdf/10.1080/0300443991530104>
- Florida Department of Education. (1997). *The basics of school improvement and accountability in Florida*. Tallahassee.
- Goodson, L., & Rohani, F. (1998). Higher Order Thinking Skills • Definition • Teaching Strategies • Assessment. *Thinking*, *18*, 458. Retrieved from http://www.cala.fsu.edu/files/higher_order_thinking_skills.pdf
- Govier, T. (1997). *Socrates' children*. 1997. Peterborough, Ontario, Canada: Broadview Press.
- Heong, Y. M., Yunos, J. Bin, & Hassan, R. Bin. (2011). The perception of the level of higher order thinking skills among technical education students. *2011 International Conference on Social Science and Humanity*, *5*, 281–285.
- Higher-order thinking*. (n.d.). Retrieved July 21, 2015, from http://en.wikipedia.org/wiki/Higher-order_thinking

- Indramalar, S. (1997). Bangsa Malaysia “the only way.” *The Star*, 4.
- Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131–137.
- Marzano, R. J. (1993). How classroom teachers approach the teaching of thinking. *Theory into Practice*, 32(3), 154–156. Retrieved from <http://ezproxy.um.edu.my:2057/stable/pdfplus/10.2307/1476696.pdf>
- Ministry of Malaysia Education. (2012). *Malaysia Education Blueprint 2013 - 2025. Education*. doi:10.1016/j.tate.2010.08.007
- Northern Illinois University. (2010). Bloom ’s Taxonomy.
- Rajendran, N. S. (2001). The Teaching of Higher-Order Thinking Skills in Malaysia. *Journal of Southeast Asian Education*, 2(1).
- Rajendran, N. S. (2008). *Teaching and Acquiring Higher-Order Thinking Skills Theory & Practice*. Penerbit Universiti Pendidikan Sultan Idris: Tanjong Malim Perak.
- Stiggins, R. J. (2005). *Student-Involved Assessment FOR Learning*. New Jersey: Pearson Prentice Hall.
- The Glossary of Education Reform. (2015). *BLOOM’S TAXONOMY*. Retrieved April 13, 2015, from <http://edglossary.org/blooms-taxonomy/>
- The Malaysia Statistics Department. (2011). Statistics Labour Force.
- Wlodkowski, R., & Ginsberg, M. (1995). *Diversity and motivation*. San Francisco: Jossey-Bass.
- Zohar, A., & Schwartz, N. (2005). Assessing teachers’ pedagogical knowledge in the context of teaching higher-order thinking. *International Journal of Science Education*, 27(13), 1595–1620.

Appendix A

A. Clarity (logic, consistency)

- 0- Unscorable: Response is blank, voided, off-topic.
- 1- Poor: Response is illogical, ineffective, shows a lack of consistency of thought.
- 2- Satisfactory: Response is quite logical, quite effective, shows adequate consistency of thought.
- 3- Good: Response is marked by good consistency in thought, ideas are relevant and logical.

B. Depth of reasoning (precision and depth of thought)

- 0- Unscorable: Response is blank, voided and off-topic.
- 1- Poor: Response is unclear, superficial, imprecise and inaccurate.
- 2- Satisfactory: Response shows some precision of thought, somewhat accurate and clear but lacks exemplary depth.
- 3- Good: Response is exemplary, accurate, clear and marked by high precision of thought.

C. Grammar

- 0- Unscorable: Response is blank.
- 1- Poor: Response has serious and persistent errors in grammar that severely interfere with meaning.
- 2- Satisfactory: Response may has some errors, but generally demonstrates control of grammar.
- 3- Good: Response is generally free from errors in grammar. No more than three errors.

D. Vocabulary

- 0- Unscorable: Response is blank.
- 1- Poor: Response has very limited range of vocabulary use.
- 2- Satisfactory: Response shows some reasonable use and control of owned vocabulary.
- 3- Good: Response is generally characterized by reasonable and appropriate use and control of owned vocabulary.

*If the response has more than three grammatical errors or vocabulary errors, the response would automatically fall into unscorable category in A and B.