INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2222-6990

The Development, Validity and Reliability of Higher-Order Thinking Skills - Peer Appraisal Inventory (HOTS-PA)

Mohammad Aziz Shah Mohamed Arip, Norakmar Abd Rashid, Siti Hawa Ahmad, Nurul Syahirah Husin

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v8-i3/3836 DOI: 10.6007/IJARBSS/v8-i3/3836

Received: *23 Feb 2018,* ***Revised:*** *21 Mar 2018,* ***Accepted:*** *23 Mar 2018*

Published Online: 26 Mar 2018

In-Text Citation: (Arip, Rashid, Ahmad, & Husin, 2018)

To Cite this Article: Arip, M. A. S. M., Rashid, N. A., Ahmad, S. H., & Husin, N. S. (2018). The Development, Validity and Reliability of Higher-Order Thinking Skills - Peer Appraisal Inventory (HOTS-PA). *International Journal of Academic Research in Business and Social Sciencest*, *8*(3), 395–409.

Vol. 8, No. 3, March 2018, Pg. 395 - 409

http://hrmars.com/index.php/pages/detail/IJARBSS | **JOURNAL HOMEPAGE**

The Development, Validity and Reliability of Higher-Order Thinking Skills - Peer Appraisal Inventory (HOTS-PA)

Mohammad Aziz Shah Mohamed Arip
Associated Professor, Faculty of Human Development, Sultan Idris Education University, 35900 Tanjong Malim, Perak, Malaysia
Email: aziz.shah@fpm.upsi.edu.my

Norakmar Abd Rashid, Siti Hawa Ahmad
Postgraduate student, Faculty of Human Development, Sultan Idris Education University, 35900 Tanjong Malim, Perak, Malaysia
Email: shah.aziz@ymail.com, shah.aziz@ymail.com

Nurul Syahirah Husin
Research Assistant, Faculty of Human Development, Sultan Idris Education University, 35900 Tanjong Malim, Perak, Malaysia
Email: nurulsyahirahhusin@gmail.com

Abstract
The purpose of this study is to develop and obtain the content validity and reliability of the Higher-Order Thinking Skills Peer Appraisal Inventory based on the Bloom Taxonomy Model reviewed by Anderson and Krathwohl (2001) which was modified in line with the Ministry of Education requirement. Based on the model, there are two divisions namely Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) which consist of six (6) sub-scales namely remembering, understanding, applying, analyzing, evaluating and creating. A total of nine (9) panel of experts with knowledge, skills and practice of HOTS were involved in validating the HOTS-PA. Overall, the experts agreed that the items of HOTS-PA can measure higher-order thinking studies. The expert panel gave the overall content value of HOTS-PA of 0.876 (87.56%). The reliability value was obtained from 40 secondary school students. Overall, the reliability analysis for the HOTS-PA inventory shows a high Cronbach Alpha 's coefficient value of .927. This suggests that HOTS-PA have good reliability and proven that HOTS-PA is suitable to be used to know the level of Higher-Order Thinking Skills among high school students aged 16 to 19 years old. Overall, 30 items have high reliability value and these items are developed at a good and

acceptable level. The modifications to some of the items were made after obtaining the experts assessment in making the reliability value conform to the standards set in the development of the items. With regards to the results obtained, several recommendations have been submitted.
Keywords: Validity, Reliability, Higher-Order Thinking Skills-Peer Appraisal Inventory

Introduction

In order to achieve a developed world status, several issues need to be addressed by the government, especially in education. Education is an important effort in preparing human resources in all areas. An important statement in the education philosophy in Malaysia is the effort to develop a comprehensive and integrated individual potential to produce happy and balanced person who are intellectual, spiritual, emotional human beings based on trust and obedience to God (Ministry of Education, 1990). Higher-Order Thinking Skills (HOTS) was also introduced in the Malaysian education with the aim of creating a knowledgeable and intellectual generation in order to compete internationally.

Thinking process is a mental process or mental activity to find meaning. Teaching and Learning (T&L) that emphasize on thinking skills is the core of learning (McGregor, 2007). In the Malaysian education system, the application of thinking skills is associated with a process of using the mind whether to seek meaning and understanding of things, to make judgments and decisions, or to solve problems. The T&L process based on thinking skills and thinking strategies, requires the use of planned T&L strategies and methods to comprehensively develop a student's minds (Ministry of Education, 1999). Furthermore, in the transformation of national education through the implementation of the Primary School Standard Curriculum (KSSR) starting in 2011 and the Secondary School Standard Curriculum in 2017, critical thinking and creative thinking along with innovative thinking ability become the focus in the efforts to create balanced human capital to address the current and future challenges (Ministry of Education, 1999).

Background of the Study

The Higher-Order Thinking Skill is the highest level in the cognitive process hierarchy. HOTS occurs when a person gets new information, keeps it in his memory and compiles, associates it with existing knowledge and generates this information to accomplish something or solve complex situations. Thus, HOTS is defined as the use of potential minds to deal with new challenges (Onosko & Newmann, 1994). In this situation, one needs to understand, translate, analyze, and manipulate information. HOTS requires students to critically evaluate information, create inferences, and make generalization. They will also produce genuine communication, make predictions, propose solutions, create and solve non-routine problems related to everyday life, express new opinions and ideas, evaluate ideas, make choices and make decisions (Onosko & Newmann, 1994).

The Malaysian Education Blueprint (MEB) 2013-2025 has outlined thinking skills as one of the six aspirations to be achieved. Thinking skills are needed by every student to be able to compete globally. Mastery of sound thinking skills plays a significant role in producing creative and innovative students. The thinking skills emphasized in education are the Higher-Order Thinking Skills (HOTS). According to the Ministry of Education Malaysia in the Higher-Order Thinking Skills Initiative (2013), the Higher-Order Thinking Skills (HOTS) is defined as the ability to apply

knowledge, skills and values in making reasoning and reflections to solve problems, make decisions, being innovative and able to create things.

The implementation of HOTS in schools comprises of seven elements: Curriculum, Pedagogy, Assessment, Co-curriculum, Community and Private Support, Capacity and Resource Building. Higher-Order Thinking Skills Inventory -Peer Appraisal (HOTS-PA) is developed based on HOTS elements in pedagogy. It is used to measure the students level of thinking throughout their learning process. In Malaysian education system, the thinking skills of students in manipulating ideas and understanding of the feelings contained in the text are given priority (Indramalar, 1997). The education system has been updated from time to time to improve the teaching and learning of analytical and rational thinking (Indramalar, 1997). Therefore, in order to achieve the objective of producing students with thinking skills, Ministry of Education has formally introduced Teaching Thinking Skills Programs in schools and teachers training colleges since 1992 (Rajendran, 1998).

Elements of thinking skills embedded in value-added skills are taught to students through questions and activities that require critical and creative thinking in various activities. Thinking skills whether in conceptualizing ideas, solving problems, or making decision are important in the daily life and future career of the students (Center for Curriculum Development, 2003).

According to Rajendran (1997), some of the higher-order thinking test instruments that have been developed and published in line with the level and focus of student assessments are as follows:

i. The California Critical Thinking Skills Test: College Level (1990),
ii. The Cornell Critical Thinking Test, Level x (1985)
iii. The Ennis-Weir Critical Thinking Essay Test (1985)
iv. The Watson-Glaser Critical Thinking Appraisal Test (1980)
v. The Cornell Class Reasoning Test (1964)
vi. Observation Appraisal Test (1983)

Numerous studies and inventories were conducted to measure the HOTS yet there has not been any truly successful inventory available to be used among Malaysian secondary school students, specifically in peer appraisals. Hence, the researchers have developed HOTS-PA for that purpose.

Literature Review

According to Phillips (1997), cognition is the thinking skills used in procurement, processing and recalling information obtained through the senses of sight, hearing, touch, taste and smell. The received information is matched, compared to and classified with information stored in short term memory and long-term memory. Metacognition is an executive function that manages and controls how one uses his mind and is the highest and sophisticated level of thinking.

Abd. Rahim Abd. Rashid (1999) argued that learning using concepts is an essential part of the thinking process. According to him, thinking skill can be developed when students try to understand a concept by classifying or categorizing. Maimunah (2004) also stated that thinking and thinking skills are two different things. Thinking is an abstract activity and often occurs at semi-conscious state. Thinking is considered as a natural ability similar to the ability to run and breathe, while thinking skill is not a natural skill. Thinking skill is a knowledge discipline that can be followed and practiced in forming habits or experiences (Maimunah, 2004). Thinking exercises cause individuals to make less mistakes, confusions or mistakes in thinking.

The education curriculum transformation in the Malaysia Education Blueprint (MEB) 2013-2025 emphasizes on the concept of higher-order thinking skills (HOTS) which is capable of producing generations with critical and creative thinking capabilities. This approach was introduced to achieve the ultimate goal of education which is to produce more students with high level cognitive ability through active learning pedagogy in teaching and learning. However, the goal has not yet been fully achieved; thus, various approaches have been introduced to produce intelligent, creative and innovative human capital to meet the challenges of the 21st century so that the country can compete in today's world.

Hence this understanding coincides with the emphasis given by Islam that intellectual capacity is a key element in the formation of a holistic student that is the balance between academic excellence and strong character. Therefore, the process of Islamic Education transformation has been designed based on the implementation of teaching in schools through modifications involving components of curriculum capacity, school culture formation, knowledge enhancement of teachers and the students ability to apply each learning content acquired so that the goal of HOTS-oriented teaching and learning introduction can be implemented effectively and thus forming a strong appreciation of faith among students (Noor Hisham, 2011; Zuraidah, 2013). Higher-Order Thinking Skills according to Brookhart (2010) involve questioning, being capable, understanding and analyzing things to understand their own and others' thoughts. Among the activities that can be carried out for the Higher-Order Thinking Skills (HOTS) are through active thinking, viewing the environmental context based on different perspectives and putting each idea in order (Saad et al., 2013).

For Rajendran (2001) and Normah (2013), the factor of readiness is the effort to develop Higher-Order Thinking Skills (HOTS) which is an essential and must-have element for a teacher. The absence of readiness factor among the teachers can cause the Higher-Order Thinking Skills (HOTS) in the teaching process to be put on halt although the teacher has the awareness and knowledge about the importance of HOTS (Fathiyah, 2010). In the study conducted by Rajendran (2008) on Bahasa Melayu (BM) and English teachers, it shows that Bahasa Melayu and English teachers are more confident in terms of knowledge they have to teach Bahasa Melayu or English compared to knowledge associated to HOTS. Vijayaletcmy and Selvam (2015) obtained teachers' degree of knowledge on the difference between HOTS and CCTS (Critical and Creative Thinking Skills) from 200 teachers in a survey. The findings show that teachers' knowledge of HOTS is still at a moderate level.

Attitude can also reflect positive or negative feelings toward something. In the context of this study, the positive attitude of teachers in accepting HOTS encourages and motivates students to learn and understand something well. According to Habib (2005), the attitude of teachers in relation to the curriculum change to HOTS show unfavorable attitude. Sukiman et al. (2012) study shows a moderate result on the practice of various forms of thinking skills conducted on 144 Mathematics teachers. Teachers incapability in teaching will affect teachers' teaching, where the final goal of teaching cannot be properly achieved. Determining the right method will generate a student's critical and creative thinking. Teacher-centered teaching and learning sessions are shown to be ineffective and the overall curriculum delivery, particularly those related to HOTS, is less successful.

Theoretical Foundation of Hots-Pa Inventory Development

The Higher-Order Thinking Skills Inventory - Peer Appraisal (HOTS-PA) is developed based on the Bloom Taxonomy as revised by Anderson and Krathwohl. Benjamin S. Bloom, in 1949, proposed his idea of a classification based on the order of thinking ability in an increasingly high level of process. In 1990, Lorin W. Anderson and David Krathwohl researched and made improvements to the Bloom's taxonomy, and it was published in 2001 as Anderson and Krathwohl Revision of Bloom Taxonomy.

The thinking level categories are still ordered hierarchically from the lowest to the highest. In the cognitive thinking capabilities domain, analysis and synthesis are integrated into analysis. The six categories in the previous concepts remain unchanged because Lorin W. Anderson and David Krathwohl included the new category which is *creating* that was previously unavailable. Lorin W. Anderson and David Karthwohl have also changed the Bloom's taxonomy from noun to verb.

The thinking levels found in the Anderson and Krathwohl Revised Bloom Taxonomy were arranged hierarchically as shown in Figure 1. However, they are interconnected between one level of thinking to another level of thinking.

a) Remembering

Retrieve, reproduce and recall relevant knowledge from long-term memory. The verbs used to measure the Remembering level of thinking are to select, describe, define, indicate, label, register, place, match, memorize, name, eliminate, collect, recognize, specify and state.

b) Understanding

Developing meaning from verbal, written and graphical information through interpreting, giving examples, classifying, summarizing, making inferences, comparing and explaining.

The verbs used to measure the understanding level of thinking are classifying, defending, demonstrating, differentiating, explaining, expressing, presenting, expanding, giving examples, describing, displaying, linking, interpreting, evaluating, considering, matching, making expression, representing, restating, rewriting, defining, summarizing, stating, translating and outlining.

c) Applying

Using procedure to do or implement something.

The verbs used to measure applying thinking level are to apply, define, dramatize, explain, generalize, estimate, manage, organize, prepare, generate, produce, select, demonstrate, sketch, complete and use.

d) Analyzing

Breaking materials into small parts, determining how the small parts are related to each other and towards common structures or uses through comparison, arrangement and attributes.

The verbs used to measure the analyzing level of thinking are analyzing, categorizing, classifying, comparing, distinguishing, scrutinizing identifying, concluding, dividing, detailing, selecting, defining, demonstrating and conducting surveys.

e) Evaluating

Make judgments based on criteria and standards through reviews and criticisms.
The verbs used to measure the level of Evaluating thinking are valuing, considering, criticizing, defending, and comparing.

f) Creating

Unifying elements to form a coherent or functional general idea; arranging elements into forms or to new structures through generating, planning and producing.
The verbs used to measure the thinking level of creating are selecting, specifying, combining, composing, constructing, developing, creating, designing, planning, expanding, performing, formulating, hypothesizing, finding, creating, managing, planning, producing, role-playing and presenting.

In order to categorize higher-order thinking skills, it is referred to the Ministry of Education Malaysia (2013), which states that the element of the Higher-Order Thinking Skills (HOTS) is embodied in the curriculum standard document which is the foundation of the teaching and learning activities in the form of standard content and learning standard in KSSR and KSSM. The standard content and standard learning are outlined in the standard curriculum document which emphasize on the ability of the students to apply, analyze, evaluate, and create in their learning activities.

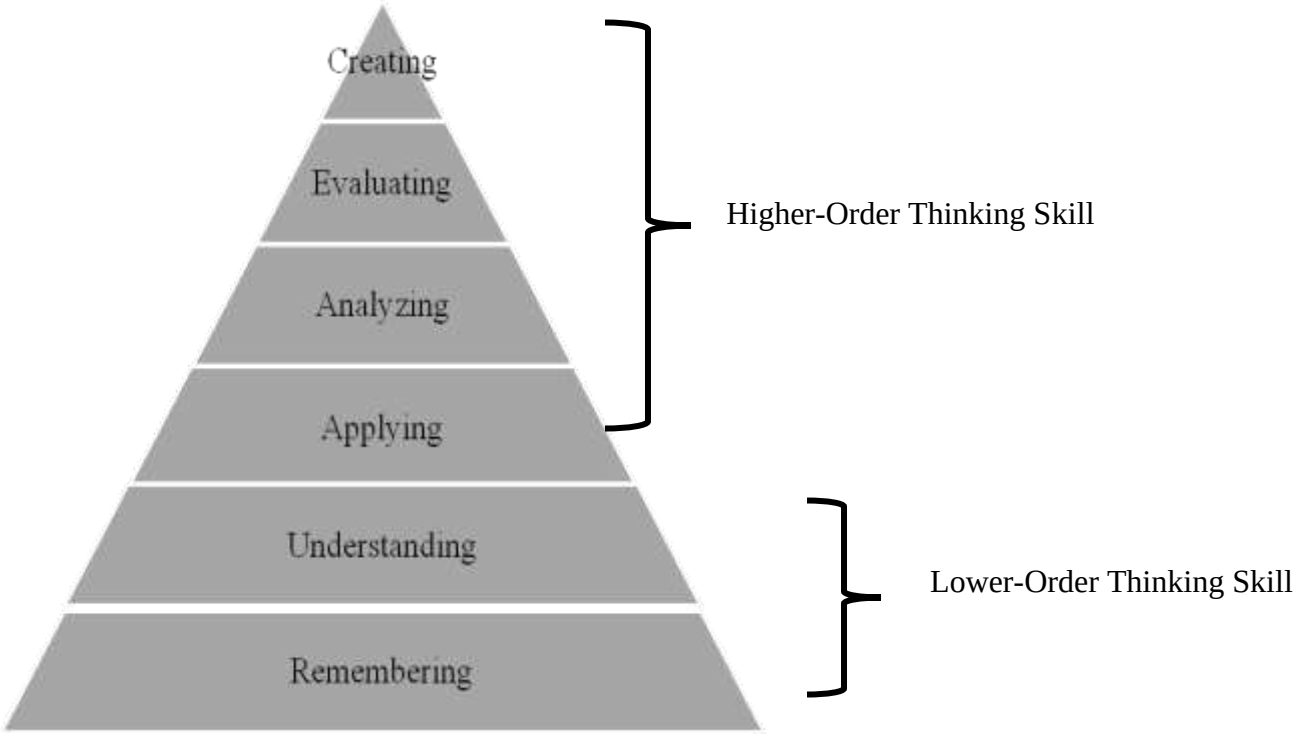

Figure 1:

Bloom Taxonomy Model (2001) revised by Anderson and Krathwohl is modified accordance with the Ministry of Education Malaysia

Purpose of the Study

The purpose of the study was to determine the reliability and content and face validity of the Higher-Order Thinking Skill Peer Appraisal Inventory (HOTS-PA). The study specifically aimed to:

a. Develop HOTS-PA based on library research.
b. Determine the overall content validity value of the HOTS-PA through expert assessment.
c. Determine the overall reliability value of HOTS-PA through alpha coefficient analysis.
d. Determine the reliability value of all HOTS-PA sub-scales through alpha coefficient analysis.

Administration, Scoring and Interpretation of HOTS-PA Scores

The Higher-Order Thinking Skills Inventory -Peer Appraisal (HOTS-PA) was developed using the framework of Anderson and Krathwohl Revised Bloom Taxonomy Model (2001). This inventory aims at measuring the level in the process of thinking whether to remember, understand, apply, analyze, evaluate and create. HOTS-PA has 30 items which are divided into 2 divisions namely Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS). For the LOTS section, there are 2 sub-scales i.e., HOTS 1-Remembering sub-scale and HOTS 2-Understanding. Meanwhile for the HOTS section, there are four sub-scales, namely HOTS 3-Applying, HOTS 4-Analyzing, HOTS 5-Evaluating and HOTS 6-Creating. There are five items for each subscale. Each item is measured with the five-point Likert scale of Never, 1-3 times, 4-6 times, 7-9 times and more than 10 times. It is designed to measure the level in the thinking process of students aged 16 to 19 years. This inventory is governed by a 'third party report' method that uses peer observation methods towards the respondents. HOTS-PA is also easy to be administered as it takes only 30 to 45 minutes. This inventory is ideal for administration in pairs. Respondents need to indicate the frequency of activities performed by their peers in the last 60 days based on observations. Situation in responding to this inventory should be conducive, comfortable and respondent should answer honestly. For the scoring, HOTS-PA gave a value of 0 for Never, 1 for 1-3 times, 2 for 4-6 times, 3 for 7-9 times and 4 for more than 10 times, for answers marked based on peer observation. The answers need to be added according to columns. Afterwards, it is multiplied by 0 for the scale of never, multiplied by 1 for the frequency of 1-3 times, multiplied by 2 for the frequency of 4 - 6 times, multiplied by 3 for the frequency of 7-9 times and multiplied by 4 for the frequency of more than 10 times. The scores are then sum up for each HOTS sub-scale of 1 to 6**.** The HOTS-PA interpretation analysis is divided into 3 levels namely High, Medium and Low. For the LOTS section, the high level represents a score of 27 to 40, a medium level represents a score of 14 to 26 and a low level represents a score of 0 to 13. Meanwhile for the high-level HOTS section represents a score of 54 to 80, medium levels represent the score of 27 to 53 and the low level represent the score of 0 to 26. This score leads to the thinking level of respondent's peer whether it is HOTS or LOTS which has been practiced in the classroom.

Methodology

The design of this study is a descriptive study. The purpose of the descriptive study is to obtain the content validity and the reliability value of the HOTS-PA developed through libraries

reference and past studies. This study involved three study phases namely; Phase 1: Developing HOTS-PA, Phase 2: Obtaining Face Validity and Content Validity and Phase 3: Reliability Value Analysis.

Phase 1: Developing HOTS-PA

The HOTS-PA was developed based on a thorough library study through the reading of books, past studies and related journals. The HOTS-PA was built based on the Revised Bloom Taxonomy Model of Anderson and Krathwohl (2001) framework which was modified in line with the Ministry of Education of Malaysia. The collection and reinterpretation of relevant information were done in this phase.

Phase 2: Obtaining Face Validity and Content Validity

In this second phase, the researcher has performed the content validity of the measuring instruments which involved the 9 experts panel consisting of academics namely lecturers in the Faculty of Education, University of Malaya, Faculty of Education, Universiti Teknologi MARA and School Improvement Specialist Coaches (SISC +), District Education Office. The rest are practitioners comprising of high school teachers. The group of practitioners will contribute to the assessment of sentences, clarity of sentences or terms and suggest more important and relevant items. The opinion of Rubio et al. (2003) is parallel to Miller et al. (2013) which emphasize that for studies in the context of education, the expert may be friends in the faculty, college administrators or students themselves. Rubio et al. (2003) added that practitioners at a studied institution are capable of being the source in developing and reviewing the measurements. This is because they are working and involved in the forefront of the population which is the subject of the actual study sample. However, the face validity by previous researchers were done by language teachers before making the content validity. Othman (2004) pointed out that the appropriate amount to obtain validity value is between six to nine panel experts. Murphy and Davidshofer (1998) believe there are three steps in predicting the validity of test instruments that describe the content, determining parts of the content, determining parts of the content to be measured by each measuring item and comparing the structure of the measurement with the content structure.

Phase 3: Reliability Analysis

The third phase is aimed at achieving the reliability value of the developed HOTS-PA. Davidshofer and Murphy (1998) state that reliability is used to measure the effects and inconsistencies of psychological measurements. The PA was administered to 40 secondary school students. Sample was selected using a simple random method. Data findings were analyzed using SPSS to obtain the value of Cronbach Alpha to assess the reliability level of the HOTS-PA.

Subjects and Locations of Study

The subjects were only involved in the 2nd and 3rd phases of the study. In the second phase of the study, nine panel members gave the content validity value of the HOTS-PA. Seven academics, a lecturer at the Faculty of Education, University of Malaya, a lecturer at Faculty of Education, Universiti Teknologi MARA and five School Improvement Specialist Coaches (SISC +) District Education Office were involved. In addition, two secondary school teachers have been involved

among practitioners. In order to gain a legitimacy, two high school Malay language teachers were involved. Meanwhile, for the third phase, 40 high school students were selected to obtain the reliability data of HOTS-PA.

Study Findings

Phase 1 Study Findings: Development of Scale, Sub Scale and Inventory Items

The development of HOTS-PA is based on library research and information from reference materials such as books, articles, journals and previous research from within and outside the country to gain in-depth information relating to the main definitions and concepts of Higher-Order Thinking Skills. From the sources, the Higher-Order Thinking Skills Peer Appraisal Inventory was developed which comprised of 30 items and is divided into six sub-scale fractions namely; HOTS 1-Remembering and HOTS 2-Understanding, HOTS 3- Applying, HOTS 4-Analyzing, HOTS 5-Evaluating and HOTS 6-Creating. These 6 sub-scales are divided into two major sub-scales namely Higher-Order Thinking Skills (HOTS) and Lower-Order Thinking Skills (LOTS).

Phase 2 Study Findings: Content Validity Value of HOTS-PA

The study findings show the comments on improvements made by the expert panel of assessors. Overall, nine (9) selected experts gave positive agreement to the items submitted. All the expert panel views are referred to as the basis for the researcher to make improvements to the items to be measured in terms of reliability.

Overall, the experts agree that the HOTS-PA can measure higher-order thinking studies. Majid Konting (1998) stated that the validity of the measurement instrument was referred to as how far the instrument is used to measure the desired data to achieve the objective of the study.

Table 1: Overall and HOTS-PA Sub-Scale Content Validity Value (n = 9)

Scale / Sub-scale	Item No.	Value (%)	Experts Evaluation
Overall HOTS-PA	30	.876 (87.56%)	Accepted
HOTS 1 Sub-scale: Remembering	5	.900 (90.00%)	Accepted
HOTS 2 Sub-scale: Understanding	5	.900 (90.00%)	Accepted
HOTS 3 Sub-scale: Applying	5	.858 (85.78%)	Accepted
HOTS 4 Sub-scale: Analyzing	5	.867 (86.67%)	Accepted
HOTS 5 Sub-scale: Evaluating	5	.884 (88.44%)	Accepted
HOTS 6 Sub-scale: Creating	5	.844 (84.44%)	Accepted

Table 1 shows the overall HOTS-PA content validity value of .876 (87.56%). The highest content validity is .900 (90.00%) which is the HOTS 1 sub-scale: Remembering and sub-scale HOTS 2: Understanding and the lowest sub-scale is HOTS 6: Creating sub-scale of .844 (84.44%). Overall, this suggests that HOTS-PA has a high expert content validity value.

Phase 3 Study Findings: Reliability Value of HOTS-PA

The third phase was carried out to obtain the reliability value of the HOTS-PA. Data obtained from pilot study were processed using SPSS version 20. According to Creswell (2010), reliability is often

referred to obtain the stability and internal consistency of a construct. The reliability value was obtained by using the subject of 40 high school students to acquire reliability data. Reliability findings are shown in the following table 2:

Table 2: Overall Reliability and HOTS-PA Sub-Scales Values (n = 40)

Scale/Sub-scale	Item No.	Value	Interpretation
HOTS 1 Sub-scale: Remembering	5	.889	High
HOTS 2 Sub-scale: Understanding	5	.891	High
HOTS 3 Sub-scale: Applying	5	.886	High
HOTS 4 Sub-scale: Analyzing	5	.901	High
HOTS 5 Sub-scale: Evaluating	5	.896	High
HOTS 6 Sub-scale: Creating	5	.891	High
Overall: LOTS	10	.874	High
Overall: HOTS	20	.925	High
Overall HOTS-PA	30	.927	High

Significance level 0.5

The Cronbach Alpha's reliability value interpretation of the entire item and each sub-scale is based on the theory proposed by Valette (1997) which states the minimum reliability value of 0.50. According to Kerlinger (1973) and Majid Konting (1998), the reliability coefficient greater than 0.60 is often used, i.e. Cronbach Alpha's 0.60-0.80 value is considered to be medium high and values greater than 0.80 are high. Thus, overall reliability analysis of the HOTS-PA inventory demonstrates the high value of Cronbach Alpha's value of .927. This suggests that HOTS-PA has good reliability and proven that HOTS-PA is appropriate to determine the level of Thinking Skills among high school students aged 16 to 19 years.

Based on sub-scale, all sub-scale HOTS-PA have high reliability value of HOTS 1 sub-scale Remembering .889, HOTS 2 Understanding .891, HOTS 3 Applying .886, HOTS 4 Analyzing .901, HOTS 5 Evaluating .896 and HOTS 6 Creating .891. For sub-scale Lower-Order-Thinking Skills (LOTS) the reliability value is .874, while the Higher-Order Thinking Skills sub-scale (HOTS), the reliability value is .925. It can be concluded that sub-scale of Lower-Order-Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) have high reliability. Next, Table 3 shows the reliability value analysis to test the quality level of the items being developed.

Table 3: Reliability Value of Each HOTS-PA item (n = 40)

No.	Item	Cronbach Alpha's	Interpretation
1	Can define basic concepts of subjects.	.924	High
2	Can rewrite short notes.	.925	High
3	Have demonstrated the techniques learned.	.925	High
4	Can state the intent of a quotation.	.928	High
5	Able to defend own idea during discussion.	.925	High
6	Can create a simple note with own method.	.925	High
7	Can relate the content of the lessons learned.	.924	High
8	Able to interpret a diagram.	.924	High
9	Able to complete group assignment.	.924	High
10	Able to relate subjects with life.	.927	High
11	Can criticize current issues.	.930	High
12	Have developed ideas in creative writing.	.924	High
13	Able to list anything that was understood.	.924	High
14	Can specify important content in quotes.	.924	High
15	Can produce good essays.	.926	High
16	Have made a comparison towards something.	.927	High
17	Able to make reasonable judgments.	.924	High
18	Can plan group assignment distribution.	.922	High
19	Can mention important points about the topic studied.	.924	High
20	Can give an example other than that given by a teacher.	.926	High
21	Have pointed out Mathematical problems to friends.	.922	High
22	Can analyze current issues well.	.926	High
23	Can compare the advantages and disadvantages of an issue.	.924	High
24	Able to generate new ideas in discussion.	.923	High
25	Can remember the formula in any subject accurately.	.924	High
26	Can explain the cause of a problem.	.924	High
27	May show pure value in the classroom.	.925	High
28	Can determine the degree of importance of a thing according to priority.	.924	High
29	Can make an assessment to appreciate something good.	.924	High
30	Have managed new activities in the classroom.	.930	High

Significance level 0.5

As shown in Table 3, the lowest reliability value is .922 for item 18 under HOTS 6 sub-scale of Creating and item 21 under the HOTS 3 sub-scale of Applying. Whereas the highest reliability

value is .930 for item 11 under the HOTS 5 sub-scale of Evaluating and item 30 that is under the HOTS 6 of Creating sub-scale. Overall, all 30 items obtained high reliability value which indicates that the developed items are at a good and satisfactory level. The modifications to some items were made after getting the expert's view to make the value of reliability conform to the standards set in the development of the item. This is in line with the opinion expressed by Majid Konting (1998), which states that the reliability coefficient value of 0.60 or greater is better and acceptable.

Discussion and Recommendations

This study has a significant impact on education in Malaysia. In addition, the instruments and studies related to the measurement of the Higher-Order Thinking Skill in Malaysia are still unclear and limited. To date, there has been no inventory of the Higher-Order Thinking Skills that have been established or developed to see the level of thinking skills among Malaysian students specifically using the 'third party report' method. Previous studies which have been carried out were more on the development of questions based on subjects in schools which use thinking skills. Measurements were made based on the ability of the students to answer the given questions according to the right level of thinking. Hence, with the development of HOTS-PA it can assist teachers, officers in the District Education Office, officers in the Department of Education as well as officials in the Ministry of Education to measure the students thinking skills throughout the learning and teaching process.

Based on the analysis conducted, HOTS-PA has a high validity and reliability value. The results of this pilot study also show that HOTS-PA is able to measure the level of thinking skills among upper secondary school students. Based on the result of content validity that has been done by 9 experts panel, HOTS-PA items incorporate the content to be measured. The findings of the reliability analysis of HOTS-PA also proved that all HOTS-PA sub-scales are at an appropriate level. According to Majid Konting (1988), the reliability coefficient of 0.60 or greater is good and accepted as well as Cohen Kappa which sets the coefficient value of 0.70-0.89 as high.

Furthermore, although this inventory is developed to help identify the students' thinking skills, the usage of the inventory is still limited. This is because, this inventory is not suitable to be administered to individuals under 15 years of age. This means that it is unsuitable to be administered to primary school and lower secondary school students. Several items in this HOTS-PA use language that require a high level of understanding. Additionally, this item is developed using a 'third party report' method of peer appraisal and requires peer observation on in-class activities in the last 60 days. It may take a while to respond to this inventory. Therefore, researchers suggest simple language use for the items to be understood and should be based on self-appraisal rather than peer appraisal, so that it can be understood by all.

Further studies are proposed to reconstruct the Higher-Order Thinking Skills - Peer Appraisal Inventory in the future to produce better quality and broader usage of HOTS-PA so that it can be adopted by all parties. The researcher also suggested further research to get a deeper analysis of the HOTS-PA items such as the use of factor analysis and analysis using the RASH model and also in obtaining validity analysis. According to Frankal & Watlen (2011), validity can be proven by means of content validity, predictive validity, concurrent validity and construct validity methods. The future researcher can use other methods to test the validity of HOTS-PA. Additionally, the researcher suggested that further research on higher-order thinking skills needs to be done even

more so as the HOTS element is one of the key focus in the development of Malaysian education in this century.

Conclusion

Overall, this study succeeded in developing a Higher-Order Thinking Skills - Peer Appraisal Inventory (HOTS-PA) with good and acceptable content validity and reliability values. Thus, this HOTS-PA can measure students' thinking skills levels that comprises of two major sub-scales namely Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) with six other sub-scales in the thinking level of remembering, understanding, applying, analyzing, evaluating and creating.

Corresponding Author

Name : Associate Professor Dr Mohammad Aziz Shah Mohamed Arip
Affiliation : Sultan Idris Education University.
Country : Malaysia
Email ID : aziz.shah@fpm.upsi.edu.my
Address: : Faculty of Human Development, Sultan Idris Education University 35900 Tanjong Malim, Perak

References

Rashid, A., & Rahim, A. (1999). *History Education: Philosophy, Theory and Practice.* Kuala Lumpur. Utusan Publications & Distributors.

Anderson, L. W., Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Blueprint, M. E. (2016). Blueprint 2013-2025.(2013). *Ministry of Education, Malaysia*.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

Creswell, J. W. (2010). *Educational Research - Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th Ed.). New Jersey: Pearson Merril Prentice Hall.

Fakhruddin, F. (2010). *Application of the Element of Islamic Education Philosophy in Islamic Education Teaching.* Doctor of Philosophy Thesis, University Putra Malaysia.

Fraenkal, J. R., & Wallen, N. E. (2011). *How to Design and Evaluate Research in Education*. (8 ed.). U.S.A: Mcgraw-Hill College.

Som, H. M. (2005). *Profiling Profile of Secondary School Teachers on Implementation of Curriculum Change*. Doctor of Philosophy Thesis, UM.

Indramalar, S. (1997). Develop thinking skills, students urged. *The Star*, 3.

Ministry of Education Malaysia. (2014). *Elementary Thinking Book (Pedagogy).* Putrajaya: Curriculum Development Division

Kementerian Pendidikan Malaysia. (2014). *Elementary Thinking Book (Assessment).* Putrajaya: Curriculum Development Division

Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach*. Harcourt School.

Osman, M. (2004). Thinking skills. Kuala Lumpur: National Institute of Public Administration (INTAN).McGregor, D. (2007). *Developing thinking; developing learning*. McGraw-Hill Education (UK).

Miller, W. R. (2000). *Integrating Spirituality into Treatment*. Washington: American Psychological Association

Konting, M. M. (1998). *Research method (Fourth edition)*. Kuala Lumpur : Dewan Bahasa dan Pustaka.

Murphy, K. R., & Dan Davidshover, C. O. (1998), *Psychology Testing Principles and Application* (Edisi ke-4). Eaglewood Clifts, New Jersey : Prentice Hall

Nawi, N. H. Md. (2011). *Teaching and Learning: The Original Research Concepts of Principle According to the Idea of Islamization of Modern Science*. Presentations in the Teaching and Learning Congress, Universiti Kebangsaan Malaysia.

Yamin, N. Binti. (2013). *Teaching of Thinking Skills in Secondary School Geography students*. . Doctor of Philosophy Thesis. Faculty of Education, Universiti Malaya, Kuala Lumpur.

Onosko, J. J., & Newmann, F. M. (1994). Creating more thoughtful learning environments. *Creating powerful thinking in teachers and students: Diverse perspectives*, 27-49.

Othman, M. (2004). *Thesis Writing in Applied Social Science*. Kuala Lumpur: Published by University Putra Malaysia.

Phillips, J. M., & Gully, S. M. (1997). Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal--setting process. *Journal of applied psychology*, *82*(5), 792.

Rajendran, N. (2008). *Teaching & Acquiring Higher-Order Thinking Skills: Theory & Practice*. Malaysia: Published by Sultan Idris Education University.

Rajendran, N. (2001). *High Skills Teaching (KBAT): The readiness of teachers to handle the teaching and learning process.* The papers are presented at the Critical and Creative Thinking Skills Project Seminar (KBKK): KPM Curriculum Development Center.

Rajendran, N. (2001). *Teaching Higher-Order Thinking Skills in Language Classrooms: The Need for Transformation of Teaching Practice*. Paper presented at the 9th International Conference on Thinking, Auckland, New Zealand.

Rubio, D. M., Berg-Weger, Tebb, S. S., Lee, E. S., Rauch, S. (2003). Objectifiying Content Validity : Conducting a Content Validity Study in Social Work Search. *ProQuest Psychological Journal. Pg.94*.

Saad, S., Saad, N. S., & Dollah, M. U. (2012) Thinking Skills Teaching: Perceptions and Practice of Mathematics Teachers during Teaching and Learning in the Classroom. Journal of Science & Mathematics Education Malaysia, Vol.2, No.1 ISSN 2232-0393.

Valette, R. M. (1977). Modern language testing.

Muniyandy V., & Narayanasamy, S. (2015). *Teaching and Learning Of Higher-Order Thinking Skills(HOTS) in Secondary Schools: A Review of Teachers Knowledge HOTS and Their Implementation*. National Education Research Seminar, 2015. SPPK2015/037, 352-360.

Ramdzan@Ramban, Z. (2013). *Readiness, Practice and Strategy of Islamic Education Teaching Curriculum for First Year Junior High School*. Johor Bharu. Universiti Teknologi Malaysia.