

Development and validation of an inventory to evaluate the implementation of main educational elements in promoting higher-order thinking skills

Nor Hasnida Che Md Ghazali*, Norfishah Mat Rabi and Nurulhuda Md Hassan
Sultan Idris Education University, Malaysia

Received: 9 June 2017; Accepted: 1 November 2; Published: 4 December 2017

A valid, reliable and practical instrument is needed to evaluate the implementation of higher order thinking skills (HOTs) amongst teachers in primary schools. The purpose of this study was to assess the validity and reliability of an inventory in promoting HOTs. The inventory was adapted from Nurasyikin (2016), Gulistan *et. al* (2016) and Gonzales and Fugan (2012). The instrument in the form of questionnaire was distributed to a sample of 220 primary school teachers teaching Islamic education subject. The content validity was assessed by a group of experts, and the construct validity was measured by an Exploratory Factor Analysis. The reliability of the instrument was measured by the alpha coefficient reliability or Cronbach Alpha. Results of Exploratory Factor Analysis suggested that 12 items needed to be removed due to their non-dimensionality as they had more or less equal loadings on several factors. The instrument developed yielded high values of internal consistency as reflected by the Cronbach alpha values. The final draft of the instrument contained 107 items which was considered valid and reliable. Even though the validity and reliability of the instrument were within the acceptable range, more data needed to be gathered with a bigger sample, and further analysis using an Item Response Theory (IRT) model could be used to explore deeper into the psychometric characteristics of the items.

Keywords: Curriculum, teaching strategies, assessment practises, HOTs, exploratory factor analysis

Introduction

Malaysia foresees the importance of education in producing the kind of manpower needed to achieve Vision 2020. By the year 2020, Malaysia is aiming to be a fully industrialized country in its own mould with highly knowledgeable and skilful workers. In order to achieve this, all we have to do is to prepare the country with an education system which provides the most appropriate and high quality curriculum, teachers, delivery system, infrastructure, teaching strategies and assessment system which could meet the present and future demands. The basis

* Email: hasnida@fpm.upsi.edu.my

for the success of an education system depends on quality of the National Curriculum which will affect the human capital quality we are aiming for (Bahagian Pembangunan Kurikulum, 2014). Hence, the Education Ministry is trying to come out with a curriculum which is at par with the international standards focusing in aspects such as creative skills, problem-solving and innovation. The objective of the National Curriculum is to produce a balanced future generations which master the 21st century skills. Students who master the 21st century skills with the ability to think creatively and critically are able to compete globally. These skills are in line with the six students' aspirations as stated in the Malaysian Education Blueprint, in which each student will have the knowledge, thinking skills, leadership skills, bilingual proficiency, ethics and spirituality and also national identity.

A research conducted by the Academy of Leadership for Higher Studies (AKEPT) found that fifty percent of the teachers observed failed to deliver their lessons effectively especially in their ability to inculcate HOTS (Ministry of Education, 2013). Teachers consequently make erroneous decisions while using various methods in their assessment practices (Shepard, 2000). In addition, teachers exhibit misconceptions about their assessment practices while evaluating student academic achievement. This might be due to lack of assessment knowledge and skills in assessment practices (McMillan, 2001). This situation is quite disturbing as classroom assessment is meant to support instruction and enhance students' learning (Shepard, 2000). Haris and Hofer (2009) believe that learning activities influence the daily instructional development, and the planning has to focus on students' standard and curriculum which is related to learning process, the outcome of learning and also the existing technology. So, to produce an effective teaching, a curriculum approach based on technology has to be produced which includes teachers' technique in planning instruction and also their knowledge in planning instruction. Furthermore most teachers, especially new teachers are having problems in planning teaching and fail to manage teaching effectively especially when it comes to teaching HOTS (Nurasyikin, 2016). Teachers feel that when planning teaching, it is difficult to transform it into words to write it in their daily lesson plan book. Some teachers do not understand the rationale in planning teaching and the use of teaching objectives (Orstein & Lasley, 2000). Most senior teacher focusses more on the content and instructional activities rather than planning for teaching objective. This is quite worrying because teaching objective is an important component in the curriculum as it will give a quite a big impact to the teaching in the classroom.

Background and rationale

Previously, Critical and Creative Thinking Skills (CCTS) have been introduced in 1994 through Primary School New Curriculum (KBSR) and Secondary School New Curriculum (KBSP). Primary School Standard Curriculum (KSSR), which is introduced in 2011 is an effort to strengthen thinking skills with more emphasis on reasoning skills. Various thinking skills have been combined for the purpose of making judgments and assessments in problem solving processes. In facing global competition in the economic environment which is driven by innovation, MOE has emphasized HOTS in the school system. By definition, HOTS is an abstract thinking which integrates informational systems and follow rules of logic and judgment (Ivie, 1998). Norman (2009) defines HOTS as a non-algorithmic and complex mode of thinking which could generate various solutions to the proposed problem. In general, there are two types of thinking skills or cognitive skills which are higher and lower-order of thinking skills. If the thinking skills involve acquiring and understanding knowledge, it is called LOTS whereas if it involves applying and evaluating knowledge, it is called HOTS (Ozgelen, 2012).

According to *Bahagian Pembangunan Kurikulum* (Curriculum Development Centre) (BPK) (2014), basically, the implementation of HOTS in the Malaysian education system

context is using a comprehensive and systematic approach which includes three main elements (curriculum, pedagogy and assessment) together with the supporting elements (co-curriculum, community and private support and also resources and capacity building factors). All these elements are supposed to support each other. HOTS is applied in the school system so that students do not just memorize, but also understand and know what they are learning, and they are using common sense at a high level by mastering skills like evaluating, application, analysing and innovating. HOTS enable students to apply knowledge, skills and values in making reflection to solve problems, innovating and able to invent something (BPK, 2014).

In improving education, there are many aspects which have to be considered because all aspects are important in improving students learning. It includes assessment, well-prepared teachers, well-designed and coherent curriculum and also a skilful instruction which is adapted to students' needs and personalized learning environments (Darling-Hammond, 2000). Providing students' with these key features of a sound education in instilling HOTS in students is a major foundation of an accountability system. In general, curriculum is seen an active process (Rajendran, 2016). However, curriculum to teach thinking skills has to have certain characteristics which have to be different from traditional curriculum. A thinking curriculum has to be flexible so that teachers can plan, implement and assess learning based on the uniqueness of each student. Then, it integrates subject matter and thinking skills, promotes in-depth learning of subject matter and processes related to it, the content and process objectives are situated in real-world tasks by using prior knowledge, has to be dynamic and gain findings from various research and lastly it should be taught as part of the core curriculum. Even, Costa (1999) also suggested that in order to produce a thinking curriculum, it has to have at least seven main shift which are shifting from innate intelligence to effort-based learning, from transmitting meaning to constructing meaning, from compartmentalized subjects to trans-disciplinary learning, from knowing right answers to knowing how to behave with wrong answers, from uniformity to diversity, from external evaluation to self-evaluation, from motivating others to learn to liberating the human innate passion for learning. In addition, a clear specific and measurable learning outcome could produce a quality and valid assessment (Lim et al, 2016). In general, in order to produce quality teachers for the future, Malaysia really needs a well-designed curriculum which encompasses content, pedagogy and evaluation (Nagendralingan et al, 2014).

Teaching strategies play a vital role in enhancing students' acquisition of HOTS (Constantinou & Kuys, 2013). According to Rajendran (2016), there are six strategies which could be used in enhancing thinking; questioning, metacognitive approaches, componential approaches, heuristic based approaches, critical thinking approaches and creative thinking approaches. Some of the characteristics of effective strategies for developing HOTS are activating students' prior knowledge, using classroom activities such as hand-on inquiry, grouping approach or using different forms of assessment such as alternative assessment or evaluation approaches (Gulistan *et al.*, 2015). Miri *et al.* (2007) suggest that teachers could create an environment to give opportunities to students to explore more about the complex problems or conduct group activities to promote metacognition. Next is on assessment practices related to HOTS. Classroom assessment should support instruction and also increase students' learning (Shepard, 2000). Thus, teachers need to be able to assess the acquisition of thinking skills by their students. However, this is not an easy task (Rajendran, 2016). Why? Because there is no single definition of thinking and the fact that there is no multiple choice tests that effectively tests for higher-order thinking disposition. However, the study by Suah *et al.* (2009) found that the form of assessment frequently used by school teachers was multiple-choice objective test. This is something to think about seriously by the educationist. Few characteristics have been listed out if teachers were to assess the acquisition of thinking skills in students. First, students are required to develop responses rather than just select

predetermined options, to elicit HOTs in addition to basic skills, to evaluate holistic projects directly, to synthesize with classroom instruction, to do portfolios over an extended period of time, to allow for the possibility of multiple peer judgments.

If we are to compare teaching of thinking skills between Malaysia and United Kingdom in general, there are some similarities and some differences as well (Rajendran, 2016). The education system in Malaysia today is actually rooted from the British education system so it still maintains the centralized system and conducting examinations. However, an explicit attempt to teach thinking skills in Malaysia was started in schools in 1993 whereas in UK it was developed a long time ago inspired by the work of Feuerstein, Lipman and Edward de Bono. Finally, the need to have a valid and reliable instrument to evaluate the implementation of teacher practices in instilling HOTs in Islamic education teaching is becoming increasingly important. When talking about HOTs people use to relate it to science and mathematics subject only. Validity (measuring what it is supposed to measure) and reliability (the extent to which scores are free of measurement error) of the questionnaire are the most important things to consider when dealing with measurement (Barroon and Abd Rahman, 2015; Muijs, 2011). In this study, content validity is checked by the experts in this field. Construct validity is the extent to which a set of items actually reflect the theoretical latent construct those items are designed to measure (Hair *et al.*, 2006) and in this study it is measured using EFA. And, internal reliability is a concept referring to the degree to which all of the items are measuring the same underlying construct (Pallant, 2007) whereby it is measured using Cronbach Alpha value. When a questionnaire is valid and reliable, a researcher will have confidence in the results obtained using those questionnaires during data collection. Hence, the purpose of this study is to develop an instrument to evaluate teachers' practices including in curriculum, teaching pedagogy and assessment practices in instilling HOTs and LOTs.

The implementation of the curriculum, teaching strategies and assessment practices should facilitate the transition of students' knowledge and skills into responsible actions. So, this study aims at investigating all of the main elements used by teachers to improve higher cognitive skills in an Islamic education subject for upper primary schools. This research is important as it produces instruments on the three main elements in influencing HOTs in teaching. To date, no study has developed and validated instruments on HOTs implementation for Islamic education teachings. In future research, these instruments will be used to look at the linkages in the following sets of elements: curriculum, pedagogy, assessment and students' HOTs and LOTs achievement in Islamic education subject. Recent study found that there is quite a lot of research that has conducted concerning HOTs in science and mathematics subjects but, not many research have been carried out on HOTs in Islamic education among primary school children (Wan Ismail *et. al.*, 2016).

Research objectives

The purpose of this study is to develop and assess the validity and reliability of an inventory for assessing teacher practises in the context of teaching HOTs in an Islamic education subject for upper primary schools. It explores teachers' main elements of practices in implementing HOTs in teaching and learning. The three main elements involved are curriculum, pedagogy and assessment practises, as suggested by the Malaysian government. Specifically, this study seeks to:

- a) Develop an inventory for evaluating teacher practises in instilling HOTs in an Islamic education subject;
- b) Establish the validity of the inventory; and
- c) Establish the reliability of the inventory.

Methodology

This survey was piloted with 220 primary school teachers teaching Islamic education in Perak. Thirty primary schools in Perak were selected through purposive sampling. I have used purposive sampling due to practical reasons. Initially, the instrument developed was trialled by 10 teachers to check for the appropriateness of the language and content. Analysis showed that all the items were appropriate. However, some minor changes have been made to phrases which were a bit ambiguous.

Development of the instrument

The instrument is developed based on the three main elements (curriculum, pedagogy and assessment) which support each other in HOTs implementation in schools as suggested by the Malaysian government (BPK, 2014). Curriculum is adapted from the Daily Lesson Plan Preparation Inventory (K-RPH) (Nurasyikin, 2016), pedagogy is adapted from Strategies Used Survey Questionnaire (SUS-Q) developed by Gulistan *et. al* (2016) and assessment practices inventory is adapted from the CAP-Q inventory developed by Gonzales and Fugan (2012).

a) Development of instrument on Daily Lesson Plan

Nurasyikin (2016) developed an inventory (K-RPH inventory) to determine the level of student teachers' understanding on daily Lesson Plan. Data from 388 student teachers from Universiti Pendidikan Sultan Idris on the instrument was analysed descriptively. The reliability and validity of the instrument was analysed using Rasch measurement model. It was found that the student teachers' have a high level of understanding on daily Lesson Plan. The inventory which was a double-layer instrument contains 32 items with 160 rubrics (which make its total items as 160 items) that addressed issues of understanding towards daily Lesson Plan. The instrument used a 5-point scale following a total of 5 rubrics. Each rubric had 1 to 2 scores. Maximum score was 10 point and minimum is 0. Scoring scale consisted of 3 score (0= none; 1=some; and 2=all). With this score, we transferred it into an ordinal scale as in Table 1. The 160 items Nurasyikin's K-RPH inventory covered a wide range of daily Lesson Plan planning as shown in Table 2. Items in Nurasyikin's instrument were organized into several components; i) Analyzing Daily Lesson Plan; ii) Determining Objectives in daily Lesson Plan; iii) Organizing Learning experiences; and iv) Assessment and reflection.

Table 1. Transfer from rubric score to ordinal score.

Rubric Score	Ordinal Score
0 - 2	Very Low (1)
3 - 4	Low (2)
5 - 6	Middle (3)
7 - 8	High (4)
9 - 10	Very high (5)

Table 2. Constructs and variables of K-RPH instrument.

Construct	Number of Items	Sub-construct	Variables
Analysing daily Lesson Plan	11	Daily Lesson Plan Date and day Time component Class component Subject component Title or topic component Teaching aids (TA) and learning aids (LA) component Existing knowledge	The use of daily lesson plan Rationale for writing date and day Rationale for writing time Students' background Rationale for writing subject Rationale for writing title or topic Rationale for using TA and LA Rationale for choosing TA and LA Rationale to know the existing knowledge
Determining Objectives in daily Lesson Plan	5	Learning objectives TA and LA	The function of learning objective Reference used to form objective Characteristics of learning objective Reference in choosing TA Things to consider when using TA
Organizing Learning Experiences	10	Rational of Induction set planning Teaching strategy Teaching content layout Learning activity	Induction set function Induction set activity Teaching strategy used Teaching strategy characteristics Learning strategy Factors affecting teaching content layout How to lay out classroom? Learning activity characteristics Factors affecting learning activity planning? Students' assignment characteristics
Assessment and Reflection	6	Assessment Reflection	Why writing assessment component? The use of assessment results The use of evaluation instrument Grading Why do we need to record reflection? The importance of writing students' comments in reflection section

b) Development of instrument on Teaching Strategy

The instrument for teaching strategy is adapted from Gulistan *et. al* (2016). The inventory, known as Strategies Used Survey Questionnaire (SUS-Q) developed by Gulistan is to determine strategies used by 7th grade secondary science teachers in teaching science, consisting of 34 items in the form of 5-point Likert scale (1 = never to 5 = always) based on the constructs of cognitive development. After validity and reliability processes, there are only 31 items.

The sample for the study was 212 7th grade science teachers in the Iraqi-Kurdistan region. Data were analysed by adopting descriptive and inferential statistics such as t-test and one-way ANOVA. Findings of the study indicated that the most popular strategy among 7th grade science teachers was the strategy for acquiring knowledge which focused more on memorizing basic concepts in science, while the least used strategy by science teachers was

the strategy for applying knowledge such as problem solving and hands-on activities. Items in Gulistan's instrument were organized into three main constructs; i) Strategies used for acquiring the knowledge (14 items); ii) Strategies used for applying knowledge (8 items); and iii) Reflection on knowledge strategies (9 items).

Why is this instrument chosen? Firstly, it is an instrument to determine strategies used by science teachers to teach HOTS in science education so it is suitable with this study in determining teaching strategies concerning HOTS in Islamic education subject. Secondly, this instrument is developed to suit the schools in Iraqi Kurdistan region which is not a developed country and still in the process of developing its HOTS. What changes have been done? The instrument is adapted to suits this study by referring to the Curriculum and Assessment Standard Document (DSKP) for Islamic education subject for primary Year 4, 5 and 6 and also the Administration Guidelines for Islamic education subjects produced by Ministry of Education in 2015. Changes are made to suit the teaching of Islamic education subject at the primary level.

c) Development of inventory on Classroom Assessment Practises

The inventory for classroom assessment practises (CAP-Q) of teachers is adapted from Gonzales and Fugan (2012). Initially, the inventory developed by Gonzales and Fugan consists of 89-items altogether in the form of 5-point Likert scale (1 = never to 5 = always). The inventory consists of five main constructs which are assessment planning, assessment item preparation, assessment administration and scoring, reporting of scores and grading and assessment data utilization and evaluation.

The sample for the study was a group of primary school teacher from the Philippines, Nepal and the Kyrgyz Republic. Data from the study were analysed by performing exploratory factor analyses and reliability test. The results showed that the factor analysis supported a five-factor structure accounting for 70.71% of the variance of the questionnaire. The items also showed satisfactory internal consistency. Finally, 56 items were retained. Finally, items were organized into several components: i) Assessment planning (14 items); ii) Assessment item preparation (20 items); iii) Assessment administration and scoring (12 items); iv) Reporting of scores and grading (5 items) and; v) Assessment data utilization and evaluation (5 items).

Why is this instrument chosen? Firstly, it is an instrument to determine assessment practises by primary school teachers in few Asian countries so it is suitable with this study. Secondly, the inventory focusses on activities that primary school teachers do in relation to conducting classroom assessment including formative and summative assessments. It is not focusing on concept, belief or knowledge of primary school teachers as most of the inventory has but more to teacher practises. Since assessment in Malaysia also includes both types of assessment so this inventory seems to suit best. Next, what changes have been done? The instrument is adapted to suits this study by referring to the Curriculum and Assessment Standard Document (DSKP) for Islamic education subject for Year 4, 5 and 6 and also Administration Guidelines for Islamic education subjects produced by MOE in 2015. Changes are made to suit the assessment for Islamic education subject at primary level.

The newly developed teaching practises in the implementation of HOTS inventory of this study consists of 120 items (excluding demographic information that were organized in four sections as below: Section A: Demographic Profile; Section B: Daily Lesson Plan (32 Items); Section C: Teaching Strategies (31 items); and Section D: Classroom Assessment Practises (56 items). Respondents are given sufficient time to complete the questionnaire since it is not a test, but their perception on the issues only. The questionnaires are then gathered,

and then data are analysed concerning characteristics of the items using factor analysis and internal consistency measure in order to validate the inventory.

Findings and discussions

The findings discussed in this paper are organized around two important aspects, namely validity and reliability of the instruments.

Validity of the instrument

Firstly, the validity of the instrument is established. The adequacy of the data is checked using KMO Bartlett test. A statistic value of 0.947 is obtained which indicates that 94.7 percent of the variables properties are explained by the data thus, factor analysis would be meaningful. Then, exploratory factor analysis (EFA) is performed according to section (except for demography section) to identify the number of constructs and to group the items for each construct. EFA on Section B (daily Lesson Plan) has yielded four constructs (Table 3). Three items are removed as they are grouped into two constructs with more or less equal loadings. Factor loadings for the remaining items are greater than 0.6.

Table 3. Results of EFA on Section B (Daily Lesson Plan) – 32 items

Pattern Matrix ^a				
Section B				
Item	Component			
	1	2	3	4
Daily LP helps me in ...				
B1) organizing teaching in classroom	.867			
B2) plan teaching activities	.862			
B3) plan teaching strategy	.767			
B4) plan learning experience	.779			
B5) control class	.669			
B6) Manage time wisely	.722			
B7) Prepare teaching aids	.655			
B8) Prepare assessment tools	.886			
B9) Prepare reference for future	.832			
B10) Determine strength and weakness in teaching	.656			
B11) determine the best teaching aids to use	0.23	0.34		
The objective characteristics that I have planned ...				
B12) Could measure behaviour		.722		
B13) Able to predict behaviour		.883		
B14) Have criteria in explaining skills		.682		
B15) Are specific		.604		
B16) Could explain behaviour		.781		
Teaching strategy that I have planned....				
B17) Involve students actively			.657	
B18) Trigger inquiry learning	.275		.299	
B19) Relate theory with actual practises		.301	.322	
B20) Involve higher-order thinking			.781	
B21) Involve group work			.698	
B22) Involve all individual in the group			.657	
B23) Enjoyable learning			.735	
B24) Improve leadership skills in students			.798	
B25) Involve idea development			.781	
B26) Encourage self-learning			.698	
The result of assessment and reflection is used in ...				

Pattern Matrix ^a				
Section B		Component		
Item	1	2	3	4
B27) Determining improvement in students' performance				.740
B28) Giving written reports to parents				.677
B29) Determining strength and weakness of students				.702
B30) Determining formality of schools			.332	.240
B31) Reflection in recording students' performance				.685
B32) Reflection in recording teaching effectiveness				.702

For the 31 items in Section C (Teaching Strategies) EFA has been yield three constructs. Items related to strategies used for acquiring the knowledge are grouped together as one factor, strategies used for applying the knowledge fall into the second factor and items for reflection on knowledge strategies fall into the third factor. However, five items (item C7, C10, C14, C15 and C19) are grouped into both constructs with more or less equal loadings and thus are removed from the final instrument. Factor loadings for the remaining items are greater than 0.6. Details of the results are presented in Table 4.

Table 4. Results of EFA on Section C (Teaching Strategy) – 31 items

Pattern Matrix ^a				
Section C		Component		
Item	1	2	3	
C1) Organize students to read selected chapters of the Quran with correct recitation	.867			
C2) Organize students to memorize selected chapters of the Quran with correct recitation	.862			
C3) Focus on learning students basic concepts	.767			
C4) Explain the 'process of the practice' to students	.699			
C5) Ask students to explain certain concept to peers	.669			
C6) Give an assignment which needs exploration	.722			
C7) Encourage students to generate their own questions	.655	.521		
C8) Pose a problem and encourage students to form hypothesis	.819			
C9) Give assignments which require students to use the methods taught	.732			
C10) Encourage students to answer questions that need reference	.656	.544		
C11) Involve the entire class in the search for the solution to a problem	.873			
C12) Observe students, and ask them in their small group	.722			
C13) Conduct a pre-assessment to determine early understanding of students	.883			
C14) Allow students to complete their homework in the classrooms	.701	.682		
C15) Boost students to do research	.565	.604		
C16) Encourage students to determine their basic beliefs		.781		
C17) State the problem and ask students to solve it		.717		
C18) State certain skills and ask them to explain		.684		
C19) State certain action and ask students to justify	.575	.599		
C20) Encourage students to do formulation of the basic concepts of worship		.762		
C21) Encourage students to formulate 'moral values' in everyday life		.801		
C22) Encourage students to solve the problem referring to the selected hadith		.742		
C23) Give assignments so students could identify terms in forming <i>Jawi</i> text				.657
C24) Encourage students to explain the rationale behind their ideas				.735
C25) Ask students to consider alternative explanations				.798
C26) Observe students, and ask questions while they work individually				.781

Pattern Matrix ^a			
Section C Item	Component		
	1	2	3
C27) Review students' notebooks			.698
C28) Ask students while they are discussing in large groups			.697
C29) Allow students to make oral presentations			.740
C30) Use cooperative learning approach			.675
C31) Assess students by using open ended response test			.702

Results of Section D (Assessment Practises) shows that three items are found to have poor psychometric characteristic, either they have more or less equal loadings on several factors or they have poor loadings (<0.4). Thus they would be removed in the final draft of the instrument. Details of the results are shown in Table 5.

Table 5. Results of EFA Results on Section D (Assessment Practises) – 56 items

Pattern Matrix					
Section D Item	Component				
	1	2	3	4	5
D1) I prepare at least 3 learning objectives for each subject I teach	.745				
D2) I refer to the curriculum when I organize my learning objectives	.699				
D3) I follow a taxonomy in preparing learning objectives	.723				
D4) I prepare learning objectives for each theme that I intend to assess	.721				
D5) I prepare a test plan according to the learning of my lessons	.651				
D6) I ensure that every topic that I cover in class is included in the assessment plan	.733				
D7) I relate to the instructional process with the assessment process	.729				
D8) I try to include a variety of questions to measure different levels of cognitive skills	.652				
D9) I ensure that learning objectives are clear before I plan for assessment	.699				
D10) I ensure that appropriate assessment strategies are employed	.782				
D11) I ensure that every objective is given a proportionate importance in the assessment plan	.344	-.321			
D12) I determine the category of every objective to be assessed	.722				
D13) I prepare a table of specifications for all subjects that I teach	.511				
D14) I write clear learning objectives so that students are aware of what is to be assessed	.677				
D15) I use textbooks as reference when I write test items		.675			
D16) I review the guidelines for item writing whenever I construct my test		.543			
D17) I include a variety of questions in a single test		.771			
D18) I group all items of similar format		.723			
D19) I go over all questions and revise them to suit the specifications of my new test		.699			
D20) I make sure I give clear directions for every type of question I include in a test		.713			

Pattern Matrix					
Section D	Component				
Item	1	2	3	4	5
D21) I arrange test questions from easy to difficult		.745			
D22) I ensure that questions and options are on the same page		.623			
D23) I avoid including items that suggest racial, ethnic or gender biases		.714			
D24) I ensure that answers are arranged in random in multiple-choice test		.826			
D25) I only include essay questions when necessary		.811			
D26) I try to prepare questions that minimize guessing		.544			
D27) I explain the basis for scoring open-ended response items		.714			
D28) I try to balance easy and difficult questions		.826			
D29) I include on the same page the diagrams or maps needed in a particular question		.811			
D30) I write direction clearly		.714			
D31) I provide blank space for writing their names and date of testing		.826			
D32) I proofread all test questions and directions before printing them		.667			
D33) I provide enough space for each test question		.564			
D34) I indicate whether a separate answer sheet is used	-.322	.377			
D35) I give proper motivation to students before testing		-.322	.355		
D36) I ensure that the place is conducive for testing activities			.767		
D37) I see to it that cheating is not encouraged in the classroom			.565		
D38) I prepare rubrics before I start marking test papers			.656		
D39) I check whether students have enough papers and pens before starting a test			.642		
D40) I score test papers at random			.646		
D41) I ensure that I have enough test materials before I administer a test			.776		
D42) I provide enough space for all students to do their testing activities			.666		
D43) I score essay questions objectively by using rubrics			.578		
D44) I re-check test paper when necessary			.651		
D45) I follow scoring criteria strictly when marking test papers			.712		
D46) I ensure I have enough time to score test papers			.723		
D47) I provide feedback to students after every test				.887	
D48) I give a grade equivalent to the total score in a test				.776	
D49) I explain to the students how scores are derived				.656	
D50) Share the test results with teachers and headmasters, if necessary				.665	
D51) I inform parents on the students' results				.712	
D52) I determine the difficulty level of each item after exam finished					.465
D53) Conduct item analysis to differentiate students' capacity					.498
D54) Construct a simple bank items for each subject					.677
D55) Display the names of good students to motivate their peers					.665
D56) Return all examination papers which have been examined					.776

Reliability of the instrument

Analysis of validity using EFA by constructing for all sections in the instrument has yielded quite a high reliability measures. After all the 12 items have been removed, the remaining items are checked for reliability. The reliability values are greater than 0.67 for all constructs. Details of reliability index by construct are shown in Table 6.

Table 6: Values of Overall Cronbach's Alpha for Each Construct

Constructs	Numbers of Item Deleted	Numbers of Item Remained	Overall Cronbach's Alpha Value
Daily Lesson Plan			
i) Analysing daily Lesson Plan	1	10	.863
ii) Determining Objectives in daily Lesson Plan	0	5	.786
iii) Organizing Learning Experiences	2	8	.670
iv) Assessment and Reflection	1	5	.900
Teaching Strategies			
i) Strategies used for acquiring the knowledge	3	11	.929
ii) Strategies used for applying knowledge	2	6	.866
iii) Reflection on knowledge strategies	0	9	.890
Assessment Practises			
i) Assessment planning	1	13	.723
ii) Assessment item preparation	1	19	.757
iii) Assessment administration and scoring	1	11	.801
iv) Reporting of scores and grading	0	5	.822
v) Assessment data utilization and evaluation	0	5	.767
TOTAL ITEMS	12	107	

Conclusions and recommendations

Analysis of validity by EFA and internal consistency on the data show that the instrument seem to be sound and could be used to measure the implementation of curriculum, teaching strategies and assessment practices in promoting HOTS. However, results of the analyses suggest that 12 items should be removed from the instrument. There are four items from section B, five items from section C and three items from section D are removed. Thus, out of 119 items, 107 items are retained in the final draft of the instrument. The analyses yielded evidence that all the three instruments could be a useful scale to measure curriculum, teaching strategies and assessment practises amongst Islamic education teachers. The instruments from this study can be a starting point for further research. Even though the reliability and validity of the final draft of the instrument are within the acceptable range, some of the items had to be removed. Therefore, a more detailed analysis with a larger sample (>1000) using Item Response Theory (IRT) model may need to be conducted before the instrument can be finalized. The use of IRT will allow us to explore deeper into the psychometric characteristics of each item, and thus provide us with a higher level of confidence to keep only important items in the instrument.

References

- Barroon, I. A., & Abd Rahman, A. (2015). Reliability and validity of a questionnaire to evaluate diabetic patients' intention to adopt health information technology: A pilot study, *Journal of Theoretical and Applied Information Technology*, 2(72), 253-258.
- Constantinou, M., & Kuys, S. S. (2013). Physiotherapy students find guided journals useful to develop reflective thinking and practice during their first clinical placement: A qualitative study. *Physiotherapy*, 99(1), 49-55. doi: 10.1016/j.physio.2011.12.002
- Costa, A. L. (1999). *Changing Curriculum means changing your mind*. Palatine, IL: Skylight Publishing Inc.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1).
- Gonzales, R., & Fugan, C. (2012). Exploring the conceptual and psychometric properties of classroom assessment. *The International Journal of Educational and Psychological Assessment*, 9(2).
- Gulistan Ahmed, M. A. (2016). Development of a Higher Order Thinking Teaching Models for basic education students in Science (Doctoral thesis). University Malaya, Kuala Lumpur, Malaysia.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis* (6th ed.). New Jersey: Prentice-Hall International, Inc.
- Haris, J., & Hofer, M. (2009). Instructional planning activity types as vehicles for curriculum-based TPACK development introduction: TPACK. *Information Technology in Teacher Education*, 2009, 99-108.
- Ivie, S. D. (1998). Ausubel's learning theory: An approach to teaching higher order thinking skills. *The High School Journal*, 82(1), 35-42. Retrieved from <http://ezproxy.um.edu.my:2057/stable/pdfplus/10.2307/40364708.pdf?acceptTC=true>
- Lian, L. H., Yew, W. T., & Meng, C. C. (2016). Preservice teachers' assessment knowledge: Do teaching experiences make a difference? *Teacher Education*, 6(2), 16-22.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practises*, 20(1), 20-32.
- Ministry of Education Malaysia (2013). *Malaysia Education Blueprint 2013-2025 (Preschool to Post-Secondary Education)*. Putrajaya: Kementerian Pendidikan Malaysia.
- Miri, B., David, B.C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education*, 37(4), 353-369.
- Muijs, D. (2011). *Doing quantitative research in education with SPSS*. London: SAGE Publications Ltd.
- Norman, G. (2009). Problem-solving skills, solving problems and problem-based learning. *Medical Education*, 22(4), 279-286.
- Ratnavadivel, N., Hoon, C. L., Salih, M., Low, J., & Karuppiah, N., Omar, A., Md. Yasin, S., Dawi, A. H, Saad, N.S., & Hashim, A. T. M. (2014). Curriculum framework for preparing quality teachers for the future: Developing guiding principles. *Journal of Research, Policy & Practice of Teachers & Teacher Education*, 4(2), 32-44.
- Nurasyikin Abdul Rahman (2016). *Kajian penilaian tahap kefahaman RPH di kalangan pelajar* (Master thesis). Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak.
- Ornstein, A. C., & Lasley, T. J. (2000). *Strategies for effective teaching*. USA: The McGraw-Hill Co.
- Ozgelen, S. (2012). Students' Science process skills within a cognitive domain framework. *Eurasia Journal of Mathematics, Science & Technology Education*, 8(4), 283-292.

- Pallant, J. (2007). *SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows* (3rd ed.). New South Wales: Allen and Unwin.
- Rajendran, N. S. (2016). *Teaching and acquiring HOTS, theory and practice*. Tanjong Malim: Penerbit UPSI.
- Shepard, L. A. (2000). *The role of classroom in teaching and learning (CSE Tech. Report 517)*. Los Angeles, CA: University of California.
- Suah, S. L., Ong, S. L. & Osman, S. (2010). Pentaksiran pembelajaran pelajar: Amalan guru-guru di Malaysia. *Malaysian Education Dean's Council Journal*, 5, 68-83.
- Wan Ismail, W. A., Muhammad, W. I, Lubis, M. A., & Hamzah, M. I. (2016). Kesiediaan guru PI sekolah rendah di Selangor. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 3(1), 79-92.
- Yaacob, A., Walters, L. M., Md. Ali, R., Shaik Abdullah, S., & Walters, T. (2014). Reflecting on Malaysian teacher trainees' journals. *Malaysian Journal of Learning and Instruction*, 11, 1-21.