Developing an Assessment Instrument of Higher Order Thinking Skills in Mathematics with in Islamic Context

To cite this article: Kusaeri et al 2018 J. Phys.: Conf. Ser. 1097 012151

View the article online for updates and enhancements.
Developing an Assessment Instrument of Higher Order Thinking Skills in Mathematics with in Islamic Context

Kusaeri, L USadieda, T Indayati, and M IFaizien
Department of Mathematics and Science Education, Universitas Islam Negeri Sunan Ampel Surabaya, Surabaya, East Java, Indonesia
kusaeri@uinsby.ac.id

Abstract. This study aims to develop an instrument to measure students’ Higher Order Thinking Skills (HOTS) in mathematics that integrates Islamic context and identify the Islamic context which becomes the student's constraint in answering HOT S test items. The study was conducted through 5 stages: developing test specifications, writing test, reviewing test items, administering the test try-out and analysing the test. The instruments included question in algebra, statistics and arithmetic, taking Islamic contexts of muamalah, fiqh and aqidah. The instruments were tried out to 165 grade-9 students in 4 Madrasah Tsanawiyah or junior secondary schools in East Java. The try-out results were analysed using Item Response Theory (IRT) with BILOG-MG software and triangulated using descriptive analysis to find items that are difficult. The results show that out of 8 test items developed, 1 item has to be taken out from the analysis because it has a negative point biserial correlation. Five items have good discrimination level and guessing level. The other 2 items have either low discrimination level or high difficulty level. The students find difficulties in answering questions with the context of muamalah and fiqh that require HOTS. However, students find it easy to answer questions using the context of aqidah.

1. Introduction
Since Indonesia has been involved in international assessments (TIMSS and PISA), Indonesian students' mathematical achievement in TIMSS and PISA has always been under those of students in Southeast Asian countries such as Thailand, Malaysia and Singapore [1, 2]. This fact signifies studies on the causes of such low achievement. Studies by Sembiring, Hadi&Dolk[3], Dewantara, Zulkardi&Darmawijoyo[4] and Rahmawati[5] found several common causes. First, the mathematical problems given by teachers in the learning process mainly measure students’ ability to apply formulas, procedures or algorithms. This type of problems, Hallman-Thrasher [6] states, may fail to encourage students to reason and construct arguments. Students may have good ability to solve problems that test their ability to remember or read facts but fail to do higher level problems that require them to combine various information provided [7].

Second, students are not accustomed to think divergently in doing mathematical problems or tasks. It is not common to find students’ different ways of answering questions in math classes. This result on students’ perception that the correct answers to mathematical problems are those that accord the teacher’s ways or steps [8]. This mindset arouses a single or convergent way or answer. Meanwhile, divergent way of thinking is crucial to train students to explore ideas to solve challenging mathematical problems or tasks. Because they are not accustomed to higher level of thinking and
divergent mathematical problems, students are often confused when they do test items in TIMSS and PISA models that require students to use all cognitive levels of Bloom’s taxonomy [5].

The differences between the problems that the teachers provide and TIMSS and PISA problems, according to Drollinger-Vetter et al.[9], leads to students’ different perspectives and ways of learning. Some researchers have also examined the relationship between the teacher’s perspectives and practices in the classroom with student achievement [10-12]. The results agree to a conclusion that teacher’s perspectives and classroom practices colour students’ achievement. Therefore, the important task of a mathematics teacher is to build students’ divergent way of thinking so they are not confined with a rigid convergent mindset. It is thus important to develop mathematical problems requiring high order thinking skills (HOTS) as stimuli to develop students’ cognitive abilities that enable them to have HOTS to perform mathematical thinking process.

There have been a number of studies on the development of instrument to measure students’ HOTS [13-15]. These studies strongly indicate that the development of students’ HOTS require learning strategies that can actively involve students and that can focus more on the learning process, not on the content. Despite the above findings from previous studies, so far there has not been any instrument developed to assess students' mathematical HOTS that integrates Islamic context.

In fact, the integration of Islamic context on mathematical problems can develop students' reasoning and build their critical awareness of the ultimate truth that comes from Islamic values and teachings. Moreover, the integration of HOTS mathematical problems and Islamic context can minimize students' assumption that it is difficult to associate mathematics to Islamic context [16]. Mathematics will also be a particularly interesting lesson for madrasah students because they feel that mathematics is indeed a part of their lives and religion.

2. Research Methods
This research is a developmental study adapted from Mardapi development model [17], emphasizing the development of HOTS test items as assessment instruments that integrate Islamic context in each of the stems and options. Figure 1 describes the development stages. The development initially began with determining the test specifications. At this stage, a grid is developed as a guide for writing the test items. The thinking dimensions developed in this domain are directed to three aspects of ability: analyzing (C4), evaluating (C5) and creating (C6). The items consist of 5 items of analyzing (C4) domain, 2 items of evaluating (C5) domain and 1 item of creating (6) domain.

![Figure 1. The Instrument Development Stages.](image-url)
empirical and theoretical chances of embark schedules for pilgrimage on hajj while question item number 8 is related to the pattern in seating arrangement in an iftar.

From the grid, the HOTS mathematical problems were constructed, emphasizing the ability of deep thinking and the context of real cases in muamalah, fighandaqidah. The deep thinking ability is related to the reasoning built on the context of the questions, the stems and the homogeneity of the options. Under these consideration, 8 (eight) HOTS items of mathematical problems were developed. Although the use of multiple-choice items is generally considered as low-level thinking measurement, a careful construction of multiple-choice items in this study can potentially stimulate students’ HOTS process [18]. This brings the challenge in the development of question items and the uniqueness of the items developed. The options constructed also present contexts relevant to the question items. The students must analyze and evaluate the options in order to best determine and select the correct option relevant to the contexts of the questions.

A review to determine the quality of test items was conducted before the test items were tried out. MTs mathematics teachers, lecturers and an expert in Islamic jurisprudence reviewed the items. The involvement of jurisprudence experts was intended to evaluate the Islamic context used in the question items. The try out was conducted in four State Islamic Junior Secondary Schools (MTsN) in East Java, involving 165 students of grade IX. The schools for piloting the items were selected based on the level and quality of the result of National Examination in 2016 and were categorized as high, moderate and low. The results of the try out were analyzed using Three-Parameter Logistic (3PL) model of BILOG-MG software in order to know the characteristics of the items including discrimination level, difficulty level, and guessing level of the developed items. In addition to software analysis, descriptive analysis was also conducted to identify students’ problems in solving the mathematical problems that integrate Islamic context in the question items.

3. Result and Discussion

Hulin et al. [19] suggests that the good quality of the result of item analysis is indicated by the matching level of the item with the model with the index of difficulty of -2 to 2, the index of discrimination level between 0 to 2, and the guessing level not exceeding 0.2. Figure 2 presents result of instrument analysis and different biserial correlation values.

| ITEM | NAME | # TRIED | # RIGHT | PCT | LOGIT/1.7 | PEARSON BISERIAL |
|------|------|---------|---------|------|------------|------------------|
| 1    | MATH1| 165.0   | 84.0    | 50.9 | -0.02      | 0.519            | 0.650            |
| 2    | MATH2| 165.0   | 49.0    | 29.7 | 0.53       | 0.329            | 0.474            |
| 3    | MATH3| 165.0   | 65.0    | 39.4 | 0.25       | 0.042            | 0.016            |
| 4    | MATH4| 165.0   | 65.0    | 39.4 | 0.25       | 0.196            | 0.249            |
| 5    | MATH5| 165.0   | 47.0    | 28.5 | 0.54       | 0.081            | 0.207            |
| 6    | MATH6| 165.0   | 72.0    | 43.6 | 0.15       | 0.187            | 0.236            |
| 7    | MATH7| 165.0   | 65.0    | 39.4 | 0.25       | 0.388            | 0.493            |
| 8    | MATH8| 165.0   | 110.0   | 66.7 | 0.41       | 0.408            | 0.590            |

Figure 2. The Statistical Calculation of the Result of Instrument Development Analysis.

Item number 3 has a negative biserial correlation of -0.016. This negative biserial value is because the item has a poor discrimination level and falls into the category of difficult item [20]. This question item number 3 take the material of number, asking students to determine the number of chapters (surahs) memorized by female students. To determine the number of the chapters, the students must understand every statement in the problem because they are related. It seems that the students find this problem difficult because of the following possible causes: (1) the stimulus on the item is too complicated and long so it is confusing for the students to understand the context of figh and (2) the students are lack of knowledge related to the context of figh. Thus, item number 3 was excluded or was not used in the next analysis phase. Figure 3 presents the result of instrument analysis after item number 3 was excluded.
The results of the second stage analysis bring two interesting items of mathematical problems worth further analysis; those are item number 1 and item number 5. Item number 1 has low discrimination level (5.407) while item number 5 has high difficulty level (2.448). The other 5 items share the characteristics of good question items for mathematical problems.

Item number 1 deals with the concept of numbers and is associated with the context of *fiqh*. Students must understand the context of Islamic jurisprudence presented on the problem because there is important information that is useful to determine the correct answer. Students, regardless of their ability (either high or low), seem to have difficulty in understanding the problem and determining the correct answer. Item number 1, thus, cannot discriminate students with high ability from those with low ability. It is highly possible that the students who understand the material in item number 1 thought that there is incorrect information in the item. In terms of the stem and options, question item number 1 has long stems and options that bring students’ confusion in understanding the context of Islamic jurisprudence. This finding confirmed previous findings [3-5] in that students are not accustomed to answering questions that require reasoning and to the presentation of problems that begins with long elaboration of context as stimulus. This is because they are used to answering questions from their teachers that mostly require the application of formula and procedures.

Question item number 5 discusses the concept of arithmetic with the context of *muamalah* as the stem. The mathematical problem presented requires the students to understand the context of *muamalah*. The results of the analysis show that this question item can be categorized as difficult, indicated by students’ difficulty in answering the question, even for students with high mathematical ability. The correct answer given by students with low ability indicates the guessing factor. Students’ lack of understanding toward the context of *muamalah* seems to be the contributing factor toward students’ difficulty in providing the correct answer to this question item number 5.

In this context, it is true that students tend to have difficulty in solving a complex mathematics problem that involves varied information [7]. This happens since teachers focus on items that demand the ability in applying formula, procedure and algorithm [3-5]. As a result, student will be confused when they have an item that requires critical thinking and HOTS from taxonomy Bloom [6-7].

Actually, teachers have willingness to implement HOTS but they don’t have knowledge and experience about HOTS and how to implement HOTS in the context of mathematics learning [22]. Also, most of teachers still have thinking to follow the teaching based on the textbook when they teach in the class. Therefore, they have to finish the materials without necessarily considering how to give mathematics problems according to the current context.

The other five items share the good characteristics of question items in terms of their difficulty, discrimination and guessing levels. The five items has the discrimination level (parameter $a$) in the range of $0 \leq a \leq 2$, the difficulty level (parameter $b$) in the range of $0 \leq b \leq 2$, and the guessing level (parameter $c$) at the intervals of $0.124 \leq c \leq 0.276$. In terms of the Islamic context used, 2 out of these 5 items use the context of *aqidah*. This context has been studied by students since their Islamic primary education (MI), so the students are already familiar with the terms used in the context. This brings ease to the students in answering the questions.

Examined from Item Characteristic Curve (ICC), the above 7 items can be put into 2 categories. Item number 1 has low discrimination level, item number 3 has the highest difficulty level and the rest
five items are presented in Figure 4. The ICC graphic clearly show the differences among the two categories of the question items.

The shape of ICC in Figure 4 depends on the magnitude of the discrimination and difficulty levels. The discrimination level is shown by the slope in the graphic, the more vertical the curve of the corresponding item, the better the item in discriminating students with high ability from students with lower level of ability. Figure 4 (a) represents ICC item number 1 which has low discrimination level and item number 5 with high difficulty level. This is because the slope in Figure 4 (a) goes along with the difficulty level of the item, the more sloping the higher the ability is. This results in the higher probability of correct answer and the minute difference between students with high and lower level abilities. Figure 4 (b) represents the 5 with have good item characteristics shown in an S-shape ICC [21].

![Figure 4. (a)ICC of Item with Low Discrimination Level and High Difficulty Level.](image1)

![Figure 4. (b)ICC of Item with Good Item Characteristics.](image2)

Figure 4 indicates that a number of students are able to solve complex items that are different from their real ability, and it happens because of guessing [23]. This description doesn't occur in Figure 4 (b). Relating to ICC, only items that have good characteristic can be used in the context of learning.

4. Conclusion
Based on the previous analysis and discussion, it can be concluded that of the seven question items of HOTS mathematical problems developed, 5 items have good discrimination, difficulty, and guessing levels while the other two items has either low discrimination level (item number 1) or high difficulty level (item number 5). In regards to the Islamic context used as the stimulus in the HOTS mathematical problems developed, most students find difficulty with the contexts of fiqh and muamalah but find it easy with the context of aqidah. It can be assumed that this is because aqidah has been instilled since they were born and is always practiced in their daily lives.

5. Acknowledgments
The authors would like to thank the Research and Community Outreach Unit of UIN Sunan Ampel Surabaya for the research grand, the school principals, teachers and students of MTsN 1 Gresik, MTsN 1 Sidoarjo, MTsN 2 Surabaya and MTsN 3 Surabaya for participating in the research. May Allah give good reward for all of their kindness. Amin....YRA
6. References

[1] A Jupri, P Drijvers, and M van den Heuvel-Panhuizen, "Difficulties in initial algebra learning in Indonesia," Mathematics Education Research Journal, vol. 26, pp. 683-710, 2014.

[2] S Lailiyah, T Nusantara, C Sa’dijah, E Irawan, K Kusaeri and A H Asyhar, "Structuring students’ analogical reasoning in solving algebra problem," in IOP Conference Series: Materials Science and Engineering, p. 012029, 2018.

[3] R K Sembiring, S Hadi, and M Dolk, "Reforming mathematics learning in Indonesian classrooms through RME," ZDM, vol. 40, pp. 927-939, 2008.

[4] AH Dewantara, Z Zulkardi, and D Darmawijoyo, "Assessing seventh graders’ mathematical literacy in solving PISA-LIKE task," Journal on Mathematics Education, vol. 6, pp. 117-128, 2015.

[5] R Rahmawati, "Menggali lebih dalam kelemahan siswa Indonesia berdasarkan hasil analisis TIMSS 2015," Jakarta: Pusat Penilaian Pendidikan Kemendikbud, 2015.

[6] A Hallman-Thrasher, "Prospective elementary teachers’ responses to unanticipated incorrect solutions to problem-solving tasks," Journal of Mathematics Teacher Education, vol. 20, pp. 519-555, 2017.

[7] NAida, KKusaeri, and SHamdani, "Karacteristik instrumen penilaian hasil belajar matematika ranah kognitif yang dikembangkan mengacu pada model PISA," Suska Journal of Mathematics Education, vol. 3, pp. 130-139, 2017.

[8] H Haryanto, "Pengembangan cara berpikir divergen-konvergen sebagai isu kritis dalam proses pembelajaran," Majalah Ilmiah Pembelajaran, vol. 2, 2005.

[9] B Drollering-Vetter, F Lipowsky, C Pauli, K Reusser, and E Klieme, "Cognitive level in problem segments and theory segments," ZDM Mathematics Education, vol. 38, pp. 399-412, 2006.

[10] L Ingvarson, A Beavis, A Bishop, R Peck, and G Elsworth, "Investigation of effective mathematics teaching and learning in Australian secondary schools," Melbourne: Australian Council for Educational Research, 2004.

[11] J Hiebert and D A Grouws, "The effects of classroom mathematics teaching on students’ learning," Second handbook of research on mathematics teaching and learning, vol. 1, pp. 371-404, 2007.

[12] P Andrews, A Ryve, K Hemmi, and J Sayers, "PISA, TIMSS and Finnish mathematics teaching: An enigma in search of an explanation," Educational Studies in Mathematics, vol. 87, pp. 7-26, 2014.

[13] BMiri, BC David, and Z Uri, "Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking," Research in science education, vol. 37, pp. 353-369, 2007.

[14] L G Snyder and M J Snyder, "Teaching critical thinking and problem solving skills," The Delta Pi Epsilon Journal, vol. 50, p. 90-99, 2008.

[15] S Samritin and SSuryanto, "Developing an assessment instrument of junior high school students’ higher order thinking skills in mathematics," REiD (Research and Evaluation in Education), vol. 2, pp. 92-107, 2016.

[16] A Kurniati, "Mengenalkan matematika terintegrasi Islam kepada anak sejak dini," Suska Journal of Mathematics Education, vol. 1, 2015.

[17] D Mardapi, "Teknik penyusunan instrumen tes dan nontes," Yogyakarta: Mitra Cendekia, 2008.

[18] KKusaeri, "Pengembangan tes diagnostik dengan menggunakan model DINA untuk mendapatkan informasi salah konsepsi dalam aljabar," Diss. Universitas Negeri Yogyakarta, 2012.

[19] C Hulin, F Drasgow, and C Parsons, "Item response theory: Application to psychological measurement," Homewood, Illinois: Dow Jones-Irvin, 1983.

[20] R L Ebel and D A Frisbie, Essentials of educational measurement: Prentice-Hall of India Private Limited, 1991.
[21] R K Hambleton, H Swaminathan, and H J Rogers, "Fundamentals of item response theory," Newbury Park, CA: Sage Publication, 1991.

[22] B Tanujaya,"Development of an instrument to measure higher order thinking skills in senior high school mathematics instruction," Journal of Education and Practice, vol.7, pp. 144-148, 2016.

[23] H Retnawati, “Teori respon butir dan penerapanya,” Yogyakarta: Parama Publishing, 2014.