AN ANALYSIS OF TESTS DESIGNED BY ISLAMIC EDUCATION TEACHERS BASED ON HIGHER-ORDER THINKING SKILLS

Wahyu Sakban[1], Nurmawati[2], Haidir[3]
[1,2,3] Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
wahyu.sakban23@gmail.com

ARTICLE INFO

ABSTRACT

This study aims to analyze the quality and quantity of the PAI (Islamic Education) tests based on Bloom's taxonomy. The research was conducted at SMP Tarbiyah Islamiyah in Hamparan Perak. This research is qualitative research with a descriptive approach. Primary data sources are Islamic Education teachers and Mid-term tests for Islamic Education subjects, while secondary data sources are from literature studies and previous studies. Data collection instruments are documentation and interviews. The research findings indicated that the analysis of the quality of the items on the reliability aspect shows a reliability coefficient of 0.959 in the reliable category. Distractor analysis showed a functioning distractor at 25 questions or 71.43% and a malfunctioning distractor at 10 questions or 28.57% and analysis of the quality of the items according to Bloom's taxonomy in the cognitive domain. Therefore, from 35 multiple-choice items, it was found that 51.43% were in the HOTS category and 48.57% were included in the LOTS item category.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis kualitas dan kuantitas tes PAI (Pendidikan Islam) berdasarkan taksonomi Bloom. Penelitian ini dilakukan di SMP Tarbiyah Islamiyah di Hamparan Perak. Penelitian ini merupakan penelitian kualitatif dengan pendekatan deskriptif. Sumber data primer adalah guru Pendidikan Agama Islam dan Ujian Tengah Semester mata pelajaran Pendidikan Agama Islam, sedangkan sumber data sekunder adalah dari studi literatur dan penelitian terdahulu. Instrumen pengumpulan data adalah dokumentasi dan wawancara. Hasil penelitian menunjukkan bahwa analisis kualitas butir soal pada aspek reliabilitas menunjukkan koefisien reliabilitas sebesar 0,959 dengan kategori reliabel. Analisis distraktor menunjukkan distraktor berfungsi sebanyak 25 soal atau 71,43% dan distraktor tidak berfungsi sebanyak 10 soal atau 28,57% serta analisis kualitas butir soal menurut taksonomi Bloom pada ranah kognitif. Oleh karena itu, dari 35 butir soal pilihan ganda ditemukan 51,43% termasuk dalam kategori HOTS dan 48,57% termasuk dalam kategori LOTS.

Introduction

The results of the Program for International Student Assessment (PISA) and Trends in International Match and Science Survey (TIMSS) show that Indonesian students have not been able to occupy top positions since they participated in 1999 (Kemendikbud 2023). This survey shows that the majority of Indonesian students are still at the LOTS (Lower-Order Thinking Skill) level. This indicates that students' scientific literacy is still low. There is still a lot of memorized content in the short-term memory. Thinking skill still tends to center around remembering (recalling), restating, or referring without processing (reciting). Various studies show the urgency of using HOTS in learning and assessment, including Asfiyah (2021), Baharun and Sa'diyah (2018), Wicaksono (2021), Nurhayati, Jamilah and Astuti (2022), Putri (2022), Syarifuddin, Mutmainah and Fauziah (2022), and Nurjanah, Erni., Ramadhan., Gilang Mas and Diana, Lesti (2022). Considering the importance of item quality analysis and HOTS-based assessment in learning to improve students' critical thinking skills, the researchers are interested in conducting this study.

Along with the implementation of the independent curriculum, it is expected that there will be a paradigm shift in the implementation of learning. Learning which was initially teacher-centered has changed to become student-centered. In this way, teachers are expected to be more creative and innovative in delivering lessons. The application of several learning models such as project-based learning, problem-based learning, learning with a problem-solving approach, and inquiry learning is an opportunity for teachers to implement learning activities at the HOTS (Higher Order Thinking Skill) level.

The government has made several efforts to improve students' thinking skills. One of them is something new in the 2013 curriculum technical guidance concept, namely the application of Higher-Order Thinking Skills (HOTS) in Islamic Education assessments. HOTS is the ability to think that does not just remember (recall), restate, or refer without processing (recite). In principle, tests that are designed based on HOTS aim to measure students' abilities in (1) transferring one concept to another; (2) processing and applying information; (3) looking for connections between different pieces of information; (4) use information to solve problems; and (5) examine ideas and information critically.

Islamic education is one of the subjects that must be taught at schools and is very important. To have a better delivery of Islamic education subjects, it is very necessary to implement HOTS. In Indonesia, Islamic education is placed in a strategic position. This can be seen in Law no. 20 of 2003 concerning the National Education system, Article 3, namely:

National Education functions to develop abilities and shape the character and civilization of a dignified nation to educate the life of the nation to develop the potential of students so that they become human beings who believe in and are devoted to God Almighty, with healthy morals, knowledgeable, capable, creative, independent, a democratic and responsible citizen.

Assessment designed by referring to HOTS in Islamic Education has not been fully carried out by teachers, the questions developed only test students' low-level thinking skills. This can also be seen based on the results of research conducted by Ahmad and Sukirman (2019) who found that the problems that arise from the questions developed for the final exam do not fully have HOTS characteristics, and there are still many questions that only measure low level of thinking skills.

Likewise, based on initial observations made with Islamic Education teachers at Tarbiyah Islamiyah Hamparan Perak, Deli Serdang Regency, there are several problems in implementing Islamic Education, including (1) The test prepared by teachers is obtained from Islamic education textbooks and worksheets, (2) The test has never been analyzed, (3) The test has not been prepared completely based on HOTS (High Order Thinking Skill), and (4) based on mid-semester test obtained from Islamic education teachers, of the Islamic education questions given to students are still in the LOTS category which places more emphasis on aspects of knowledge, understanding and application.

Research Methods

The research method employed in this study is a descriptive qualitative method. This study seeks to reveal various facts in the research field related to the teacher's ability to compose HOTS questions. Primary data sources from primary sources in this research are Islamic Education teachers and questions from mid-test. The secondary data sources are library materials, literature, and previous studies. Data analysis techniques include data reduction, data presentation, and conclusion drawing, then data validity checking techniques include credibility, transferability, trustworthiness, and confirmability.

Results

The findings of this study are described in the following table:

Table 1. Qualitative Analysis of the Test Items

No	Aspect	Category

		Qualified		Unqualified	
		F	%	F	%
1	Learning Content	33	94,29%	2	5,71%
2	Construction	35	100,00%	-	-
3	Language	21	60,00%	14	40,00%

Based on Table 1 above, it can be seen that of the 35 items, there are 33 items, or 94.29% qualified in the material aspect, then in the construction aspect all the items are qualified and 60.00% are qualified in terms of the language.

Table 2. The Quantitative Analysis of the Test Items

No	Aspect	Kategori			
		Qualified		Unqualified	
		F	%	F	%
1	Validity	34	97,14%	1	2,86%
2	Reliability	0,959 (Very reliable)			
3	Level of Difficulty	Easy	Medium	Difficult	
		3	32	-	
		(8,57%)	(91,43%)		
4	Level of Differences	Not good	Fair	Good	Very Good
		-	1	31	3
			2,86%	88,67%	8,57%

Based on Table 2 above, it can be seen from the results of the validity testing that there was 1 test item that failed, and the results of the reliability testing questions showed a reliability coefficient of 0.959. The results of the difficulty level of the questions described 3 questions in the easy category, 32 questions in the medium category, and 0 questions in the difficult category. The results of the differentiating level of the questions are 1 test item in the fair category, 31 test items in the good category, and 3 test items in the very good category.

Discussion

The research findings show that of 35 questions, it appears that questions number 3 and 28 are not suitable to be tested based on the content aspect because the questions do not match

the competency and do not have answers. The research findings show that from 35 questions based on construction analysis, it shows that all the questions are worthy of being tested, however, several aspects are not fulfilled, such as: (1) the main aspect of the questions does not provide clues, the answer key found in the main question number 1 is the same as the answer key, (2) the length of the answer choices being relatively the same is found in questions number 2, 4, 19 and 24, which have answer choices that are not the same length, and the answer key choices are different from each other, (3) almost all questions do not meet the construction feasibility aspect regarding points 7, 9 and 10 because this question is not equipped with pictures, diagrams except question number 7, does not use the statement "all the answers above are correct", and the question items depend on the answer to the previous question.

In the language feasibility analysis, several questions did not meet the criteria as in; (1) use language that follows Indonesian rules, not fulfilled in questions 2, 3, 4, 9, 11, 12, 14, 15, 16, 18, 19, 22, 26, and 28, and (2) Answer choices do not repeat the same words/groups of words unless they are a single unit of understanding, not fulfilled in questions, 1, 4, 6, 14, 14, 23, 24, 30, and 33.

Based on Bloom's taxonomy analysis of 35 multiple choice questions, there are 17 questions (48.6%) including the LOTS question category, namely number questions; 1, 2, 4, 5, 6, 8, 14, 17, 20, 23, 24, 25, 30, 31, 32, 33 and 35. In the LOTS question category, the question criteria are spread across each level with details: C1 (remember) = 8 questions (22.8%) questions number 1, 4, 23, 24, 25, 30, 33 and 35, C2 (understand) = 3 questions (8.7%) questions number 8, 31 and 32, C3 (applying) = 6 questions (17.1%) questions number 2, 5, 6, 14, 17 and 20. Furthermore, 18 (51.4%) questions fall into the HOTS question category. This shows that half of the total questions are HOTS questions, namely number questions; 3, 7, 9, 10, 11, 12, 13, 15, 16, 18, 19, 21, 22, 26, 27, 28, 29 and 34, but all are in criteria C4 (Analysis) = 18 questions (51.4%) so there is no variation in HOTS-based preparation.

Referring to the data above, it can be seen that the questions developed by the teacher at the difficulty level show that 91.43% (32 questions) are in the medium category and 8.57% (3 questions) are in the easy category. This is referred to Arikunto's (2020:207) explanation that good questions are questions that are not too easy or too difficult. Questions that are too easy do not stimulate students to increase their efforts to solve them. On the other hand, questions that are too difficult will cause students to become discouraged and not have the enthusiasm to try again because it is beyond their reach. Referring to Arikunto's statement, the composition

of the questions in this study is satisfactory because the majority of questions designed by teachers are in the medium category.

Furthermore, regarding the percentage proportion of questions in the easy category, the percentage proportion of questions in the medium category, and the percentage proportion of questions in the difficult category, Arikunto (2020:210) did not state it emphatically, he only said that questions that were too easy or too difficult did not mean they were not allowed. used, this depends on the use.

The same thing is stated in the guidelines of the Ministry of Education and Culture of the Republic of Indonesia in 2019 in the Module for Preparing High-Order Thinking Skills (HOTS) Questions for Islamic Education and Character Education, relating to what is the percentage of questions in the easy category, the percentage proportion of questions in the medium category, the percentage proportion of questions in the easy category, and the difficult category is also not stated explicitly.

Furthermore, the research findings show that analysis of the quality of the items according to Bloom's taxonomy as revised by Anderson and Krarthwohl in the cognitive domain can be concluded from the 35 multiple choice questions that 51.43% (18 questions) are in the HOTS question category and 48.57% (17 questions) are in the HOTS question category. In particular, the HOTS category questions are all in the C4 (Analysis) criteria so there are no variations in HOTS-based preparation in C5 and C6.

Regarding the percentage of HOTS questions of 51.43% and LOTS questions of 48.57%, basically, it is not too far from the theoretical criteria put forward by Suparman (2021:2) that 40% of lower order thinking skills (LOTS) and 60% higher order thinking skills (HOTS).

The state of the teacher's ability to compose HOTS questions is unique because, in the case of this study, the percentage of questions in the LOTS category and the HOTS category is not much different in percentage. This is a particular concern in the future to further develop teachers' abilities in compiling HOTS category questions even though teachers have already received training related to preparing HOTS questions.

Islamic Education test items developed by teachers in C1, C2, and C3 are LOTS category questions while C4, C5, and C6 are HOTS category questions as explained in the 2019 Ministry of Education and Culture of the Republic of Indonesia in the Module for Preparing High-Order Thinking Skills (HOTS) Questions for Islamic Education and Budi Pekerti who explained that C1 (remembering), C2 (understanding), and C3 (applying) were included in the

Low Order Thinking Skill (LOTS) category, while C4 (analyzing), C5 (evaluating) and C6 (creating/making) is included in the High Order Thinking Skill (HOTS) category).

The same thing was emphasized by Purnomo (2019:41) and Suparman (2021:31) that Bloom's taxonomy, namely levels C1 (remembering), C2 (understanding), and C3 (applying) are included in the LOTS category, while C4 (analyzing), C5 (evaluate) and C6 (create/create) belong to HOTS categories.

In general, the ability of PAI teachers at Tarbiyah Islamiyah Hamparan Perak Middle School to compile HOTS questions has been said to be good from the material aspect because 91.42% (32 questions) are qualified, while in the construction aspect it can be said to be good because it has reached 88.57% (31 questions) are qualified, while the teacher's ability to design questions in the linguistic aspect still needs to be improved because only 42.85% (20 questions) are qualified.

The questions designed by Islamic Education teachers at Tarbiyah Islamiyah Hamparan Perak Middle School were made in multiple choice form. In this case, multiple choice questions are easier to examine and can be handed over to others (Nursalam, 2021), but working on multiple choice questions allows test takers to only guess the answers and does not represent the test taker's level of knowledge. Apart from that, multiple choice questions are full of cheating in that test takers try to find answers using certain codes, for example using their hands.

Conclusion

The conclusion of this study is described as follows (1) qualitative analysis of the quality of the questions in the material feasibility aspect shows that 32 questions (91.43%) are qualified, and several other questions that do not meet the material eligibility criteria are 3 questions (8.57 %). Analysis of the quality of items on the construction feasibility aspect of 31 questions (88.57%) which were qualified for use, and several other questions which do not meet the construction feasibility criteria, namely 4 questions (11.43%). Analysis of items on the language suitability aspect of 20 questions amounting to (57.14%) qualified for use, and several other questions that do not meet the language eligibility criteria, namely 15 questions (42.86%), (2) quantitative analysis of the quality of the items in the validity aspect showed 34 questions (97.14%) in the valid category and 1 question (2.86%) in the invalid category. Analysis of the quality of the questions in the reliability aspect shows a reliability coefficient of 0.959 in the reliable category. Analysis of the quality of the questions in the difficulty level aspect shows that 32 questions (91.43%) are in the medium category 3 questions are (8.57%) in the easy category, and 0

questions (0%) are in the difficult category. Analysis of the quality of the questions in the aspect of different power of the questions showed that 31 questions were (88.57%) in the good category, 3 questions (8.57%) were in the very good category, and 1 question was (2.86%) in the fair category.

This study suggests that the preparation of questions must be based on question indicators obtained from competency achievement indicators and basic competencies that have been determined by the government, not just based on students' textbooks or workbooks, and the questions prepared must also have answers, and meet the eligibility criteria based on the content, construction, and language.

References

Ahmad, Iqbal Faza dan Sukirman. Analisis Higher Order Tingking Skills (HOTS) Pada Soal Ujian Akhir Siswa Kelas 6 KMI Dalam Kelompok Mata Pelajaran Dirasah Islamiyah di Pondok Modern Tazakka Batang. Jurnal Pendidikan Agama Islam, Vol. XVI, No. 2, Desember 2019.

Arikunto, Suharsimi. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara, 2020.

Asfiyah, Siti. Implementasi Penilaian Berbasis High Order Thinking Skills Pada Mapel PAI dalam Meningkatkan Berpikir Kritis dan Kreatif Siswa di Tingkat SMP. Jurnal: Quality, Vol. 9 No. 1 Tahun 2021.

Baharun H, Sa'diyah K. Penilaian Berbasis HOTS Berdasarkan Taksonomi Bloom Pada Pembelajaran PAI. Riau: Hikmah, Jurnal Pendidikan Islam, Vol. 7 No. 2 Juli-Desember 2018.

Hidayat, Imam. *Kompetensi Guru Dalam Pembelajaran PAI Berbasis Higher Order Thinking Skills Di Sekolah Menengah Pertama*. Jurnal: Khazanah Pendidikan Islam, Vol. 2 No. 2, 2020.

Kementerian Pendidikan dan Kebudayaan Republik Indonesia. *Modul Penyusunan Soal HOTS Pendidikan Agama Islam dan Budi Pekerti*, 2019.

Nurhayati, Jamilah, dan Astuti, Reni. Analisis Kemampuan Pemecahan Masalah Dalam Menyelesaikan Soal HOTS. Jurnal: Prodi Pendidikan Matematika (JPMM), Volume 4, Nomor 2, Tahun 2022.

Nurjanah, Erni., Ramadhan, Gilang Mas dan Diana, Lesti. 2022. *Analisis Soal Tipe High Order Thinking Skill (HOTS) Pada Soal Penilaian Akhir Semester Tematik Ganjil Kelas V*

SDN Selakaso Tahun Ajaran 2020/202. Jurnal: Journal of Elementary Education Volume 05 Number 02, March 2022.

Nursalam. Pengukuran dalam Pendidikan. Ujungpandang: Alauddin University Press, 2021.

Risdiana, Aulia., Erna, Maria dan Holiwarni, Betty. 2022. *Pengembangan Soal HOTS* (*Higher Order Thinking Skills*) *Pada Materi Asam-Basa untuk Kelas XI SMA/MA Sederajat.* Jurnal: Jurnal Inovasi Pendidikan Kimia 16 (2) 2022.

Purnomo, Pajar. 2019. Penilaian Pembelajaran HOTS (Higher Order Thinking Skill). Cilacap: Candradimuka Pers.

Setiawati, S. 2019. *Analisis Higher Order Thinking Skills (HOTS) Siswa Sekolah Dasar dalam Menyelesaikan Soal Bahasa Indonesia.* In Prosiding Seminar Nasional Pendidikan KALUNI (Vol. 2).

Sumiati, Ati, dkk. 2018. Teknik Analisis Butir Soal dalam Meningkatkan Kompetensi Guru di SMK Cileungsi Bogor. Jurnal Pemberdayaan Masyarakat Madani (JPMM) Vol 2 No 1 (Juli) 2018. E-ISSN: 2580-4332.

Suhady, Wandy., Roza, Yenita dan Maimunah. 2020. *Pengembangan Soal untuk Mengukur Higher Order Thinking Skill (HOTS) Siswa.* Jurnal: Gantang. September 2020, Vol. 2. No. 1.

Suparman, Ujang. 2021. *Bagaimana Meningkatkan Kemampuan Berfikir Tingkat Tinggi (HOTS) Peserta Didik.* Bandar Lampung: Pusaka Media.

Syarifuddin, Mutmainah dan Fauziah, Ade Fitri.2022. Analisis Higher Order Thinking Skill (HOTS) Siswa dalam Menyelesaikan Soal Matematika Materi Fungsi Kuadrat. Jurnal: Pendidikan MIPA Volume 12. Nomor 4, Desember 2022.

Wicaksono, Ari Reza. Pengembangan Soal Berbasis Hots Matapelajaran PAI Di SMK 17 Seyegan. Jurnal: Bintang : Jurnal Pendidikan dan Sains Volume 3, Nomor 1, April 2021.

Widihastuti. 2015. Model Penilaian untuk Pembelajaran Abad 21 (Sebuah Kajian untuk Mempersiapkan SDM Kritis dan Kreatif. Dalam Siti Hamidah, Sri Wening, dan Yuswati (Penyunting), Prosiding Seminar Nasional Pengembangan SDM Kreatif dan Inovatif untuk Mewujudkan Generasi Emas Indonesia Berdaya Saing Global (hlm.77-86). Yogyakarta: Universitas Negeri Yogyakarta.