Analysis of the Items Final Semester Test of Islamic Religious Education

[1]Idil Saptaputra, [2]St. Marwiyah, [3]Muh. Irfan Hasanuddin

[1,2,3]Pendidikan Agama Islam, Fakultas Tarbiyah dan Ilmu Keguruan, IAIN Palopo
Email: idilsaptaputra383@gmail.com

Abstract

This study aims to evaluate the quality of the questions in the exam, focusing on validity, reliability, difficulty, discriminating power, and the effectiveness of misleading choices. This study used a quantitative approach with descriptive explanations. The data obtained were analyzed statistically and described in descriptive form. The population of this study consisted of 197 grade XI students at SMAN 2 Palopo in the 2020/2021 school year. A sample of 40 students was taken using probability sampling technique. Data were collected through observation and documentation, then analyzed. The results showed that 24 (96%) questions were considered valid, while 1 (4%) question was considered invalid. The questions were also proven reliable with a correlation coefficient of 0.731 (r ≥ 0.70). The difficulty level of the questions consisted of 13 (52%) easy, 9 (36%) medium, and 3 (12%) difficult. Questions also had good discriminating power of 10 (40%), fair of 8 (32%), poor of 5 (20%), and 1 (1%) each of very poor and excellent. However, the effectiveness of misleading choices in the questions was lacking, with 22 (88%) non-functional choices and only 3 (12%) functional choices. Based on the analysis of the questions in the final exam of the odd semester of Islamic Education, the quality of the questions was adequate. The implication of this research is that it provides an overview of the quality of the questions.

Keywords: Analysis of Items; Semester Fnal Test; Islamic Education.

Abstrak

Penelitian ini bertujuan untuk mengevaluasi kualitas pertanyaan dalam ujian tersebut, dengan fokus pada validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas pilihan-pilihan yang menyesatkan. Penelitian ini menggunakan pendekatan kuantitatif dengan penjabaran deskriptif. Data yang diperoleh dianalisis secara statistik dan diuraikan dalam bentuk deskripsi. Populasi penelitian ini terdiri dari 197 siswa kelas XI di SMAN 2 Palopo pada tahun ajaran 2020/2021. Sampel sebanyak 40 siswa diambil menggunakan teknik probability sampling. Data dikumpulkan melalui observasi dan dokumentasi, kemudian dianalisis. Hasil penelitian menunjukkan bahwa 24 (96%) pertanyaan dianggap valid, sementara 1 (4%) pertanyaan dianggap tidak valid. Pertanyaan tersebut juga terbukti reliabel dengan koefisien korelasi sebesar 0,731 (r ≥ 0,70). Tingkat kesulitan pertanyaan terdiri dari 13 (52%) mudah, 9 (36%) sedang, dan 3 (12%) sulit. Pertanyaan juga memiliki daya pembeda yang baik sebanyak 10 (40%), cukup sebanyak 8 (32%), jelek sebanyak 5 (20%), dan masing-masing 1 (1%) sangat jelek dan baik sekali. Namun, efektivitas pilihan-pilihan yang menyesatkan dalam pertanyaan masih kurang, dengan 22 (88%) pilihan yang tidak berfungsi dan hanya 3 (12%) pilihan yang berfungsi. Berdasarkan analisis terhadap pertanyaan dalam ujian akhir semester ganjil mata pelajaran Pendidikan Agama Islam, kualitas pertanyaan tersebut sudah memadai. Implikasinya penelitian ini memberikan gambaran kualitas pertanyaan secara kuantitatif.

Kata Kunci: Analisis butir; Ulangan Akhir Semester; Pendidikan Agama Islam.

Introduction

End-of-semester test scores play an important role in determining whether a student advances or stays, providing reports to parents or guardians, and reflecting the success or failure of the learning process.[1] Every year, end-of-semester tests are run as a routine in all educational units.[2] For the 2020/2021 academic year, the final semester test results of Islamic religious education subjects in grade XI of SMAN 2 Palopo during odd semesters can be seen in Table 1 below:

Table 1. Final Semester Exam Results of Islamic Religious Education

No.	Class	Value		Sum
		≥74	≤74	
1.	XI.IPA-1	19	11	30
2.	XI.IPA-2	13	7	20
3.	XI.IPA-3	1	7	8
4.	XI.IPA-4	13	11	24
5.	XI.IPA-5	13	11	24
6.	XI.IPA-6	3	19	22
7.	XI.IPS-1	13	19	32
8.	XI.IPS-2	3	8	11
9.	XI.IPS-3	0	6	6
10.	XI.IPS-4	1	19	20
	Totals	79	118	197

Source: Processed from odd semester final test result data

The minimum threshold score (KKM) for Islamic religious education subjects in grade XI of SMAN 2 Palopo is 74. Based on Table 1, it can be seen that the final semester test results for Islamic religious education subjects are still very low. Of the 197 students, only 79 students managed to achieve the KKM score, while 118 students have not achieved it. The percentage of students who have not succeeded in achieving KKM scores is 60%, while students who succeed are only 40%.

The low end-of-semester test results for Islamic religious education subjects in grade XI of SMAN 2 Palopo in the 2020/2021 school year caused curiosity for researchers to find the root of the problem that caused it. After making initial observations, researchers found that the quality of the question items on the final semester test was not yet known. Therefore, researchers

---

[1] Farida Gimo & Nugrahani, "Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016," *Stilistika* 5, no. 1 (2019): 36, https://doi.org/https://doi.org/10.32585/.v5i1.609.

[2] Nur Triyatno, Sriyono Sriyono, and Nur Ngazizah, "Bias Gender Ujian Akhir Semester Genap Fisika Kelas X SMA Negeri Kabupaten Purworejo Tahun Pelajaran 2012/2013," *Radiasi : Jurnal Berkala Pendidikan Fisika* 4, no. 1 (April 30, 2014): 34, https://jurnal.umpwr.ac.id/index.php/radiasi/article/view/416.

suspect that the low results of the final semester test of Islamic religious education subjects in grade XI SMAN 2 Palopo for the 2020/2021 academic year are caused by inadequate quality of question items.

This research has practical and solutive value. This type of research can be the basis for improving the quality of question items on end-of-semester tests, improving the competence of human resources in the field of education, and improving the overall quality and quality of education, especially in Islamic religious education subjects. Based on existing data, facts, and theories, researchers are interested in carrying out research entitled "Analysis Item of Final Semester Test of Islamic Religious Education Subjects".

Method

The research method used is quantitative research with a descriptive approach. The data obtained were analyzed using statistical methods. The research data is in the form of students' answers to the final semester test questions for Islamic religious education subjects. The data is converted into numerical form and then analyzed to evaluate the quality of the question items based on validity, reliability, difficulty, discriminating power, and effectiveness of deceivers. After the analysis is carried out, the results will be described in the discussion of this study.

The population in this study refers to all objects, subjects, and sources of research.[3] The population in this study is grade XI students of SMAN 2 Palopo who are Muslims and have done the final semester test questions for Islamic religious education subjects in the odd semester of the 2020/2021 academic year, with a total of 197 students. The sample represents a portion of the population that represents the specified criteria.[4] The sample selection in this study used probability sampling techniques, where each member of the population has an equal chance.[5] The sample of this study consisted of students in class XI of SMAN 2 Palopo who were considered representative of each class. The number of samples taken was 20.5% of the total population, which amounted to 40 students (20.5% x 197 = 40.4).

The research instruments that will be used are laptops or Microsoft Excel to store and analyze data in the form of list documents, stationery to perform manual analysis or calculations, and mobile phones or cameras to collect data in the form of screenshots of students' answers to the final semester test questions for Islamic religious education subjects in grade XI SMAN 2 Palopo

---

[3] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif Dan R&D* (Bandung: Alfabet, 2016), 215.

[4] Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik* (Jakarta: Rineka Cipta, 2011), 112.

[5] Sugiyono, *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, Dan R&D* (Bandung: Alfabet, 2015), 82.

in the 2020/2021 school year which were carried out online. The data analysis technique used is a quantitative analysis technique, which involves reviewing empirical data in the form of students' answers to the final semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo in the 2020/2021 academic year. Quantitative data analysis will provide information about the quality of question items based on validity, reliability, difficulty, discriminating power, and effectiveness of deceivers. The stages of data processing and analysis include: (1). Validity, (2). Reliability, (3). Level of Difficulty, (4). Discriminating Power, and (5). Effectiveness of Deceivers.[6]

Analisys Of The Items Final Semester Test Of Islamic Religious Education

This study was conducted to determine the quality of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. The quality of the question items in question is validity, reliability, level of difficulty, discriminating power and effectiveness of deceivers.

Validity

The validity of the end-of-semester test question items is calculated using the biserial correlation coefficient formula and the results of this calculation are compared with the biserial correlation value in the table (r-table) with a certain level of significance (α). If the biserial correlation coefficient is greater than the r-table value, then the question item is considered empirically valid. The r-table value for a sample of N = 40 with a significance level of 5% ($\alpha = 0.05$) is 0.312.[7] The results of validity analysis can be seen in Table 2 and Figure 1.

Table 2. Results of the validity analysis of UAS PAI in grade XI SMAN 2 Palopo for the school year 2020/2021

Validity Interpretation	Item Totals	Nomor Soal
Valid	24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13,14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25.
Non-Valid	1	12

Source: Processed from the validity data of UAS question items

---

[6] M. Chabib Thoha, *Teknik Evaluasi Pendidikan* (Jakarta: Raja Grafindo, 2001), 143.

[7] Wahyu Bagja Sulfemi, "Hubungan Persepsi Peserta Didik Tentang Kompetensi Guru Mata Pelajaran Sejarah Dengan Hasil Belajar Peserta Didik Mata Pelajaran Sejarah Di Kelas X SMA Negeri 1 Pamijahan Kabupaten Bogor," *Jurnal Fascho* 5, no. 2 (2016): 65.

Table 2 displays the results of the analysis of the validity of question items in the final semester test of Islamic religious education subjects in grade XI of SMAN 2 Palopo in the 2020/2021 academic year. Of the 25 questions, 24 questions were declared valid, while 1 question item was declared invalid because it had a negative value. Question items number 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, and 25 are question items that are declared valid, while question items number 12 are not declared valid.
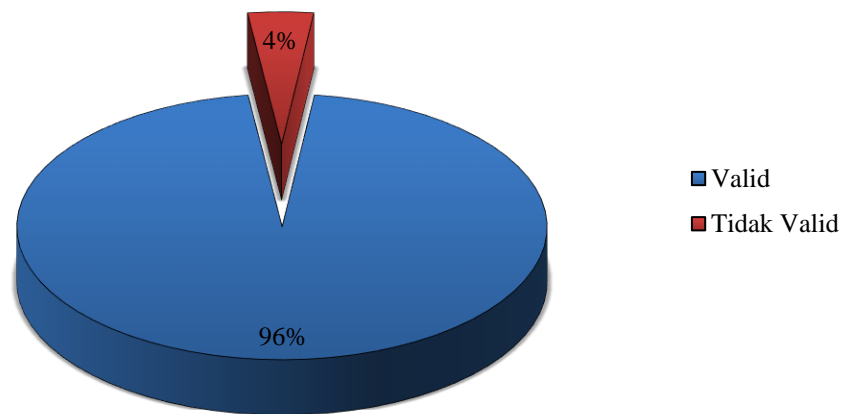

Figure 1. The percentage of validity of odd UAS question items for PAI subjects in grade XI SMAN 2 Palopo for the 2020/2021 school year

Figure 1 shows the percentage of validity of question items in the final semester test of Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. It can be seen that the percentage of question items that are declared valid is higher than the percentage of question items that are declared invalid. Question items that are declared valid have a percentage of 96%, while question items that are declared invalid only have a percentage of 4%. Thus, overall, the final test questions for the odd semester of Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year are dominated by the question items stated by valid.

Reliability

Rumus Kuder Richardson or KR21 is used incalculating the reliability of the question item. The results of the reliability calculation (ri) are interpreted with reference to the threshold value. If the ri value ≥ 0.70, then the question item can be considered reliable. Conversely, if the ri value ≤ 0.70, then the question item is considered unreliable.[8] The reliability coefficient obtained is 0.731. This value shows that the reliability (ri) of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year is classified as reliable, because the score is greater than the threshold of 0.70.

Difficulty Level

The calculation of the difficulty level of the question items is carried out using the difficulty index formula or proportion (P = B / JS). The difficulty level of the question items can be interpreted based on the results of the proportion calculation. The difficulty level of the question items can be categorized as difficult (0.00-0.30), medium (0.30-0.70), and easy (0.70-1.00).[9] The results of the analysis of the difficulty level of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year can be found in Table 3 and Figure 2.

Table 3. Results of the analysis of the difficulty level of UAS PAI question items in grade XI SMAN 2 Palopo for the 2020 school year/2021

Difficulty Level Interpretation	Item Totals	Item
Difficult	3	1, 10, 12
Medium	9	2, 3, 8, 9, 11, 16, 17, 18, 20
Easy	13	4, 5, 6, 7, 13, 14, 15, 19, 21, 22, 23, 24, 25

Source: Processed from the difficulty level data of UAS question items

[8] Febrianawati Yusup, "Uji Validitas Dan Reliabilitas Instrumen Penelitian Kuantitatif," *Tarbiyah : Jurnal Ilmiah Kependidikan* 7, no. 1 (July 24, 2018): 21, https://doi.org/10.18592/TARBIYAH.V7I1.2100.

[9] Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: Bumi Aksara, 2016), 207.

Table 3 displays the results of the analysis of the difficulty level of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. There are 3 questions categorized as difficult, 9 questions categorized as medium, and 13 questions categorized as easy. The question items included in the difficulty level are question items number 1, 10, and 12. The question items included in the medium difficulty level are question items number 2, 3, 8, 9, 11, 16, 17, 18, and 20. Meanwhile, the question items included in the easy difficulty level are question items number 4, 5, 6, 7, 13, 14, 15, 19, 21, 22, 23, 24, and 25.
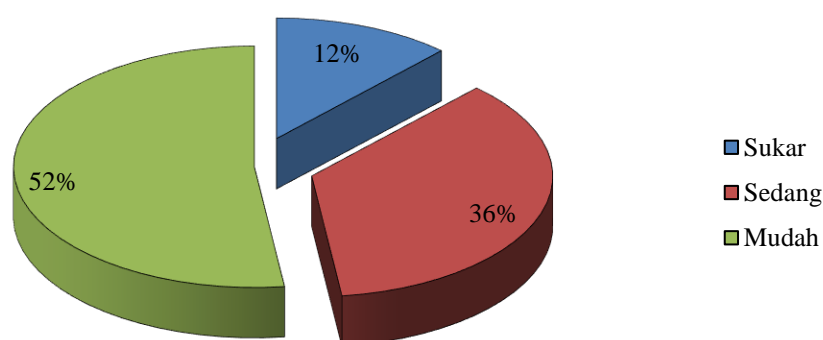

Figure 2. The percentage of difficulty level of odd UAS question items in PAI subjects in grade XI of SMAN 2 Palopo for the academic year of 2020/2021.

Figure 2 shows the percentage of difficulty of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. It can be seen that the percentage of question items at the level of easy difficulty is greater than the percentage of question items at the level of medium difficulty and the level of difficult difficulty. Question items at the easy difficulty level have a percentage of 52%, question items at the medium difficulty level have a percentage of 36%, while question items at the difficulty level are difficult only 12%.

Differentiating Power

The calculation of differentiating power is carried out using a formula to find the differentiating power index or discrimination (D = PA - PB). The distinguishing power of the question items can be interpreted as bad (0.00-0.20), sufficient (0.20-0.40), good (0.40-0.70), very good (0.70-1.00), or negative very ugly.[10] The results of the analysis of the distinguishing power of the odd semester final test questions for Islamic religious education subjects

---

[10] Arikunto, 226.

in grade XI of SMAN 2 Palopo for the 2020/2021 academic year are contained in Table 4 and Figure 3.

Table 4. Results of the differentiating power analysis of PAI UAS question items in grade XI SMAN 2 Palopo for the school year 2020/2021

Interpretasi Daya Pembeda	Item Totals	Item
Excellent	1	11
Good	10	2, 5, 8, 9, 15, 17, 18, 20, 21, 22
Enough	8	1, 3, 6, 7, 16, 19, 24, 25
Bad	5	4, 10, 13, 14, 23
Very Bad (Negative)	1	12

Source: Processed from data on the discriminating power of UAS question items

Table 4 displays the results of the differentiating power analysis of odd end-semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. Of the 25 questions, 1 question item has very good discriminating power, 10 question items have good discriminating power, 8 question items have sufficient discriminating power, 5 question items have bad discriminating power, and 1 question item has very bad discriminating power. Question point number 11 has a very good distinguishing power. Question items number 2, 5, 8, 9, 15, 17, 18, 20, 21, and 22 have good distinguishing power. Question items number 1, 3, 6, 7, 16, 19, 24, and 25 have sufficient distinguishing power. Question items number 4, 10, 13, 14, and 24 have poor distinguishing power. While question point number 12 has a very bad distinguishing power.
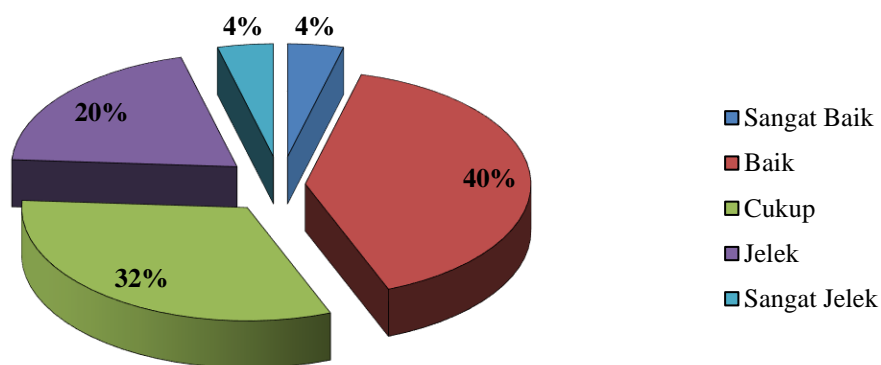

Figure 3. The percentage of distinguishing power of UAS PAI subjects in grade XI SMAN 2 Palopo for the 2020/2021 school year.

Figure 3 illustrates the percentage of distinguishing power of odd end-semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. The percentage of question items with good discriminating power is greater than the percentage of

question items with excellent discriminating power of once, enough, bad, and very ugly. Question items with good distinguishing power have a percentage of 40%, enough by 32%, bad by 20%, and very good and very bad each have a percentage of 4%.

Based on these results, it can be concluded that the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year are dominated by question items with good discriminating power. This shows that questions can be effective in distinguishing between students who have high abilities and students who have low abilities.

Effectiveness of Deceivers

Table 5 displays the results of the analysis of the effectiveness of deception of odd end-semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. The calculation of the effectiveness of the deceiver uses the deception index (IP) formula to evaluate whether the alternative answer choices that serve as deceptors or not. The odd end-of-semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year consist of 25 questions with 5 alternative answer choices, so there are 4 alternative answer choices that function as deceivers. A question item is said to have a good deceiver if at least 5% of all test participants choose the deceiver.[11] In this study, using a sample of 40 participants, so 5% of 40 is 2. Therefore, the question items are categorized as having good deceivers if each deceiver is chosen by at least 2 students. The results of the analysis of the effectiveness of deception of odd end-semester test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year can be found in Table 5 and Figure 4.

Table 5. Results of the analysis of the effectiveness of deception of UAS PAI question items in grade XI SMAN 2 Palopo for the school year of 2020/2021

Interpretation Effectiveness of Deceivers	Item Totals	Item
Function	3	6, 8, 11
Non-Function	22	1, 2, 3, 4, 5, 7, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25

Source: Processed from data on the effectiveness of UAS question item trickers

---

[11] Anita Anita, Sulis Tyowati, and Zuldafrial Zuldafrial, "Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas," *Edukasi: Jurnal Pendidikan* 16, no. 1 (June 19, 2018): 42, https://doi.org/10.31571/edukasi.v16i1.780.

Table 5 presents the results of the analysis of the effectiveness of deceivers on the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year. Of the total 25 questions, there are 3 questions with the effectiveness of the deceptor that works and 22 questions with the effectiveness of the deceiver that does not work. Question items 6, 8, and 11 have the effectiveness of a working trickster, while question items number 1, 2, 3, 4, 5, 7, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, and 25 have the effectiveness of a deceptor that does not work.
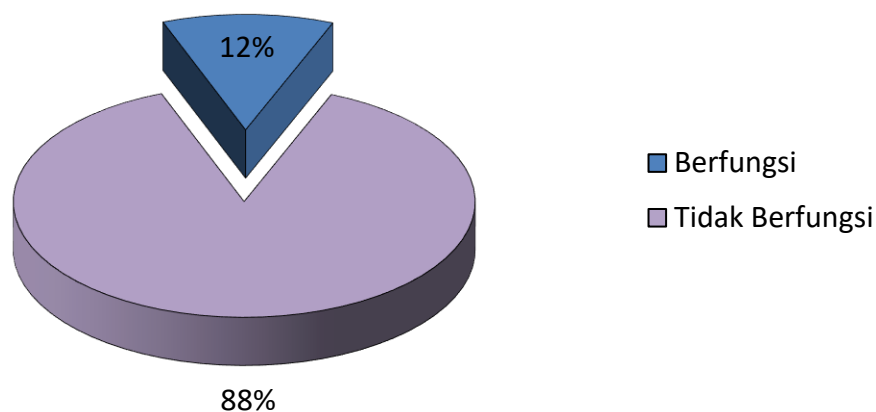

Figure 4. The percentage of effectiveness of PAI UAS question item deceivers in grade XI SMAN 2 Palopo for the academic year of 2020/2021.

Figure 4 shows the percentage of effectiveness of deceivers on the odd semester final test questions for Islamic religious education subjects in grade XI SMAN 2 Palopo for the 2020/2021 academic year. It can be seen that the percentage of question items with the effectiveness of the deceptor that does not work is higher than the percentage of question items with the effectiveness of the deceiver that works. The percentage of question items with the effectiveness of non-functioning deceivers reached 88%, while those that functioned were only 12%. Thus, it can be concluded that the dominance of the odd semester final test questions for Islamic religious education subjects in grade XI of SMAN 2 Palopo for the 2020/2021 academic year lies in the question items with the effectiveness of deceptors that do not work.

Conclusion

This study aims to evaluate the quality of question items in the final test of odd semesters of Islamic religious education subjects in grade XI SMAN 2 Palopo for the 2020/2021 academic year. The quality of the question items studied includes validity, reliability, level of difficulty, discriminating power,

and effectiveness of deceivers. The results showed that as many as 24 questions (96%) were declared valid, while 1 question item (4%) was invalid. In addition, the question items were also proven reliable based on the calculation of the correlation coefficient of 0.731 (ri ≥ 0.70). In terms of difficulty, there are 13 questions (52%) that are categorized as easy, 9 questions (36%) are categorized as medium, and 3 questions (12%) are categorized as difficult. Furthermore, the question items are dominated by good discriminating power, with 10 questions (40%) having good discriminating power, 8 questions (32%) having sufficient discriminating power, 5 questions (20%) having bad discriminating power, and 1 question item (1%) having very bad or very good discriminating power. However, the effectiveness of the problem item deceiver showed the dominance of the non-working deceiver, with 22 question items (88%) having a malfunctioning trickster and only 3 question items (12%) having a working trickster. Based on this analysis, it can be concluded that the quality of the odd semester final test questions for Islamic religious education subjects in grade XI SMAN 2 Palopo for the 2020/2021 academic year has reached an adequate level in measuring student abilities, taking into account aspects of validity, reliability, level of difficulty, and discriminating power.

References

Anita, Anita, Sulis Tyowati, and Zuldafrial Zuldafrial. "Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas." *Edukasi: Jurnal Pendidikan* 16, no. 1 (June 19, 2018): 35–47. https://doi.org/10.31571/edukasi.v16i1.780.

Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara, 2016.

———. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta, 2011.

Gimo & Nugrahani, Farida. "Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016." *Stilistika* 5, no. 1 (2019): 35–46. https://doi.org/https://doi.org/10.32585/.v5i1.609.

Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif Dan R&D*. Bandung: Alfabet, 2016.

———. *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, Dan R&D*. Bandung: Alfabet, 2015.

Sulfemi, Wahyu Bagja. "Hubungan Persepsi Peserta Didik Tentang Kompetensi Guru Mata Pelajaran Sejarah Dengan Hasil Belajar Peserta Didik Mata Pelajaran Sejarah Di Kelas X SMA Negeri 1 Pamijahan Kabupaten Bogor." *Jurnal Fascho* 5, no. 2 (2016): 52–70.

Thoha, M. Chabib. *Teknik Evaluasi Pendidikan*. Jakarta: Raja Grafindo, 2001.

Triyatno, Nur, Sriyono Sriyono, and Nur Ngazizah. "Bias Gender Ujian Akhir Semester Genap Fisika Kelas X SMA Negeri Kabupaten Purworejo Tahun

Pelajaran 2012/2013." *Radiasi : Jurnal Berkala Pendidikan Fisika* 4, no. 1 (April 30, 2014): 34–37. https://jurnal.umpwr.ac.id/index.php/radiasi/article/view/416.

Yusup, Febrianawati. "Uji Validitas Dan Reliabilitas Instrumen Penelitian Kuantitatif." *Tarbiyah : Jurnal Ilmiah Kependidikan* 7, no. 1 (July 24, 2018): 17–23. https://doi.org/10.18592/TARBIYAH.V7I1.2100.