Al-Tatawur : International journal of Social Science (TIJOSC)

E-ISSN: 3005-5342
P-ISSN: 3005-5334

Research Article

Analysis of the Islamic Religious Education Achievement Test Instrument Using Microsoft Office Excel 2019 : Spearman-Brown Reliability Test

Atabik[1]*, Khafifatul Fian[2]

[1] Pendidikan Bahasa Islam, Universitas Islam Negeri Prof. K. H. Saifuddin Zuhri Purwokerto
[2] Pendidikan Agama Islam, Universitas Islam Negeri Prof. K. H. Saifuddin Zuhri Purwokerto
Address : Jl. A. Yani No.40A, Karanganjing, Purwanegara, Kec. Purwokerto Utara, Kabupaten Banyumas, Jawa Tengah, Indonesia 53126
* Corresponding Author : atabik.zuhdi@gmail.com

Abstract: In quantitative research, there is often a biased study because it does not understand the steps or testing that must be passed. One of the tests that researchers must go through to conduct quantitative research is the instrument reliability test. The purpose of this paper is to analyze the types of reliability tests that the author then tested on the data that the author compiled using the 2019 version of Microsoft Office Excel. The type of research used is library research through data collection techniques by searching for sources either from articles, books, or other sources relevant to the topic of discussion. After the data is collected, it is then selected, and analyzed based on content or content, then draw conclusions. The results showed that the reliability test aims to find out how fixed the respondents' answers to an instrument even though it is done several times. There are several choices of techniques used to conduct reliability tests, including external or internal reliability. Researchers can choose the type of reliability test technique according to their abilities and needs. In addition, the use of Microsoft Office Excel 2019 as a tool in conducting the Spearman-Brown reliability test demonstrates that statistical analysis in educational research can be performed using software that is accessible and widely used. Excel, with its built-in functions and add-ins, allows researchers—especially beginners—to conduct basic reliability tests without requiring more complex statistical software.

Keywords: Instruments; Microsoft Office Excel 2019; Quantitative Research; Reliability Test; Spearman-Brown.

Received: July 28, 2025
Revised: August 11, 2025
Accepted: August 25, 2025
Online Available: August 27, 2025
Curr. Ver.: August 27, 2025

1. Introduction

Quantitative research is a type of research that demonstrates the relationship between variables. A variable refers to anything that is defined by the researcher for the purpose of obtaining information related to the phenomenon under study. When a researcher intends to measure a variable, it is necessary to have a measurement tool that is both accurate and accountable. An instrument is a tool used by researchers to collect data, measure events, and analyze data in accordance with the problems faced by the subjects or sample. Instruments consist of test and non-test forms. An example of a test instrument is learning achievement, while an example of a non-test instrument is a questionnaire. A researcher who has developed an instrument must conduct testing before administering it to respondents. One of the required tests is the reliability test. Through a reliability test, researchers are able to determine the consistency of an instrument when it is used multiple times in a study.

Based on the author's review, studies on reliability testing have been widely discussed. Among them are articles entitled "Analysis of Validity and Reliability of Measurement Instruments for Recognizing Geometric Shapes," "Validity and Reliability of the Achievement Motivation Questionnaire," "Validity and Reliability Test of the Prosocial Behavior Questionnaire," and "Validity and Reliability Test of the VO_2max Test for U-18 Soccer Players." From these studies, it can be seen that reliability testing can be applied to measure various aspects, including cognitive, affective, and psychomotor domains. The difference in this study is that the author specifically focuses on the reliability testing of instruments and on how to conduct reliability tests using Microsoft Office Excel 2019.

2. Proposed Method

This research is a literature study, with data collection carried out through the examination of books, articles, and other sources relevant to the topic under discussion. Once the data were collected, the researcher conducted a selection process to identify materials that were most relevant to the study. The selected data were then analyzed using content analysis techniques, followed by the final step of drawing conclusions.

3. Results and Discussion

3.1. Definition of Quantitative Research

In relation to quantitative research, the definition emphasizes that it is a type of research that involves calculation processes or, in other words, makes use of numerical data. The research must begin by clearly determining the population or sample to be studied. Quantitative research is required to be value-free; therefore, the principle of objectivity is strictly applied. To achieve such objectivity, researchers must employ instruments that have been tested for both validity and reliability.

According to Untung Nugroho, the quantitative method is referred to as a traditional method characterized by systematic, structured, and well-planned procedures, in which numerical data are involved from the stages of data collection, interpretation, to the presentation of research findings. Meanwhile, Indriyaningrum et al. explain that quantitative research employs a hypothesis-verification logic, which begins with deductive reasoning to formulate hypotheses, followed by testing in the field, and concludes with drawing conclusions.

From the various definitions described above, it can be further concluded that quantitative research is a systematic form of inquiry that involves numerical data from the process of data collection to the presentation of findings. Through quantitative research, the outcomes become measurable and structured. The analysis in quantitative research is carried out deductively, beginning with general information and proceeding to more specific insights.

3.2. Characteristics of Quantitative Research

With regard to the characteristics of quantitative research, Jaya explains that it possesses several distinct features. First, quantitative research is conducted with the purpose of measuring variables, identifying relationships between variables, determining the influence of one variable on another, and distinguishing among variables. Second, the problems addressed in quantitative research typically concern the effects, relationships, or differences between variables. Third, quantitative research is carried out to test specific theories that have already been established and selected by the researcher. Fourth, this type of research places greater emphasis on the function of theory as a starting point for exploring the concepts contained within the theory itself. Fifth, quantitative research employs hypotheses once the researcher has determined which theory will be applied [11].

The sixth characteristic is the prioritization of data collection techniques through questionnaires. Seventh, the data presentation is usually in the form of distribution tables based on respondents' answers. Eighth, quantitative research applies an etic perspective, meaning that the data collected is restricted or determined by the researcher based on certain indicators. Ninth, it makes use of operational definitions. Tenth, it employs specific criteria for determining the number of respondents or samples. Eleventh, the process of drawing conclusions follows a deductive approach, moving from general to specific. Twelfth, the main research instrument is the questionnaire used for data collection. Thirteenth, data analysis is conducted statistically, or in other words, it involves numerical calculations. Fourteenth, the conclusion in quantitative research generally reflects the degree of relationship among the variables studied [11].

3.3. Learning Achievement in Islamic Religious Education

Rosyid et al. explain that learning achievement refers to the results of measuring students' cognitive, affective, and psychomotor domains after participating in the learning process, assessed using test instruments or other relevant tools [12]. Abduloh et al. state that learning achievement is the outcome obtained by students through learning activities conducted both at school and at home [13]. According to Harefa, learning achievement is the attainment demonstrated in a particular skill or body of knowledge [14]. Chaerunisa and Latief further explain that learning achievement refers to students' outcomes in a specific topic or learning process within a certain period of time [15].

From the various definitions above, it can be concluded that learning achievement is the result of the learning process, which involves the cognitive, affective, and psychomotor domains, and is measured using relevant instruments to determine students' level of understanding of the material taught [16]. When connected with Islamic Religious Education anderstood as a body of studies on Islam presented through teaching materials that function as both guidance and training within the scope of Islamic studies it can be concluded that learning achievement in Islamic Religious Education refers to the results of students' activities in the learning process through structured training within the framework of Islamic studies.

3.4. Definition of Reliability Testing

Marisa explains that reliability is an index indicating the extent to which a measuring instrument can be relied upon or, in other words, trusted [17]. Makkira et al. further describe reliability testing as a procedure conducted to measure the consistency of a series of measurements. Such testing demonstrates the degree of accuracy and the extent to which an instrument can be trusted and depended upon in measurement activities [18]. Fadhila et al. define reliability testing as a tool used to measure questionnaires, citing Ghozali in their explanation that a questionnaire is considered reliable if the responses to its items remain consistent over time [19]. According to Sugiyono, a reliable instrument is one that, when used repeatedly to measure the same object, produces identical data [20]. Similarly, Riyanto and Hatmawan state that reliability testing refers to the consistency of a measuring instrument in assessing what it is intended to measure [21].

From these explanations, it can be concluded that reliability testing in quantitative research is a process that determines the consistency of measurements. Consistency in this context means that an instrument yields stable results when used repeatedly at different times on the same object. Through reliability testing, researchers can identify whether an instrument is accurate and dependable for use in further studies.

3.5. Types of Reliability Testing

Reliability testing can be conducted externally or internally. External reliability testing includes the test–retest method, the equivalent method, and the combined method. Meanwhile, internal reliability testing is carried out through item consistency analysis using specific techniques [20]. The test–retest method is an external reliability test conducted by administering the same instrument several times to the same respondents [22]. The equivalent method involves administering different but equivalent (comparable) instruments once to the same respondents [23]. The combined method is performed by administering two equivalent instruments several times to the same respondents:

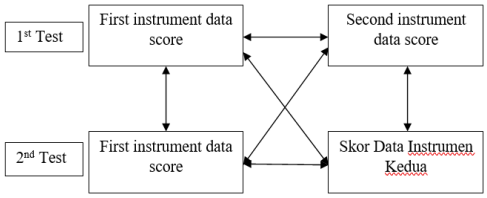

Figure 1. Combined External Reliability Testing.

Whether using the external reliability test of the test–retest type, the equivalent type, or the combined type, all three share the same criterion: an instrument is considered reliable if the correlation coefficient is positive or significant [22]. The correlation between the results of the first test and the subsequent test is calculated using the Product Moment Correlation formula [24]:

$$r_{xy} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n(\sum x_i^2) - (\sum x_i)^2][n(\sum y_i^2) - (\sum y_i)^2]}}$$

Description:
r_{xy} : Product Moment correlation coefficient
n : number of respondents
x_i : score of each item on the first test
y_i : score of each item on the second test

The significance of the correlation can be tested in two ways: (1) by comparing the correlation coefficient with the r table of the Product Moment, where the correlation is significant if the calculated r value is greater than the table value; or (2) by using a t-test [22]. The t-test formula is as follows [25]:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Description:
t : calculated t-value
r : correlation coefficient
n : number of respondents

After the calculated t value is obtained, it is then compared with the t table. The t table values correspond to the levels of significance available in the r table, namely 0.50, 0.25, 0.20, 0.05, 0.02, 0.01, and 0.0005. However, the levels most commonly used are 0.01 and 0.05. The degree of freedom (df) is determined by subtracting two from the number of respondents (df = n – 2). A result is considered positive or significant if t calculated > t table [22].

Furthermore, internal reliability testing consists of several techniques: the Spearman-Brown formula, Kuder Richardson (KR-20), Kuder Richardson (KR-21), and the Hoyt Analysis of Variance (ANOVA Hoyt) [20].

First, the Spearman-Brown test is a reliability test applied only to items that have already been proven valid. The requirements are that the instrument data must be dichotomous (scored 1 and 0), and the number of items must be even. This test is conducted by dividing the scores into two groups, either odd–even or first–last items. The Spearman-Brown formula is as follows [26]:

$$r_{11} = \frac{2 \times r_{1/2,1/2}}{1 + r_{1/2,1/2}}$$

Description:

r_{11}: instrument reliability coefficient

$r_{1/2,1/2}$: correlation index between the two halves of the instrument

Next, the Kuder Richardson Reliability Test (KR-20) is used when the instrument data are dichotomous (scored 1 and 0) and the valid items are odd in number. The KR-20 formula is as follows [20]:

$$r_i = \frac{k}{k-1}\left(\frac{s_t^2 - \sum p_i q_i}{s_t^2}\right)$$

Description:

k: number of items in the instrument

p_i: proportion of respondents who answered item i correctly

q_i: $1 - p_i$

s_t^2: total variance

Next, the Kuder Richardson Reliability Test (KR-21) is applied when the instrument data are dichotomous (scored 1 and 0). The KR-21 formula is as follows [20]:

$$r_i = \frac{k}{k-1}\left(1 - \frac{M(k-M)}{k\,s_t^2}\right)$$

Description:

k: total number of items in the instrument

M: mean of the total scores

s_t^2: total variance

The next reliability test is the Hoyt Analysis of Variance (ANOVA Hoyt), which is applied when the instrument data are dichotomous (scored 1 and 0). The formula for the Hoyt analysis is as follows [20]:

$$r_i = 1 - \frac{MS_e}{MS_s}$$

Description:

MS_s: mean square between subjects

MS_e: mean square error

r_i: reliability coefficient of the instrument

6. Internal Reliability Testing of Test Instruments Using Microsoft Office Excel 2019

In relation to internal reliability testing, the author used a test instrument for measuring learning achievement in Islamic Religious Education, with dichotomous scoring (1 and 0) and a total of 10 respondents. Before conducting the instrument trial, the criteria for reliability index interpretation are presented as follows [27]:

Table 1. Interpretation of Reliability Coefficient Values.

Value Range	Category
$0.80 \le r_{11} < 1.00$	Very High Reliability
$0.60 \le r_{11} < 0.80$	High Reliability
$0.40 \le r_{11} < 0.60$	Moderate Reliability
$0.20 \le r_{11} < 0.40$	Low Reliability
$0.00 \le r_{11} < 0.20$	Very Low Reliability

Reliability Testing Using the Spearman-Brown Formula

Table 2. Test instrument data.

No	Respondent	P1	P2	P3	P4	P5	P6
1	X1	1	1	1	1	1	1
2	X2	1	0	1	1	1	1
3	X3	1	1	1	1	1	1
4	X4	1	1	1	1	1	1
5	X5	1	1	1	0	1	1
6	X6	1	0	1	0	1	0
7	X7	1	1	0	1	1	0
8	X8	1	1	1	1	1	1
9	X9	1	1	1	1	1	1
10	X10	1	1	1	1	1	1

Description:

Value 1 = true

Value 0 = false

Fist step grouping data based on odd-even numbers.

Table 3. Grouping odd-even data.

No	Respondent	P1	P3	P5	P2	P4	P6
1	X1	1	1	1	1	1	1
2	X2	1	1	1	0	1	1
3	X3	1	1	1	1	1	1
4	X4	1	1	1	1	1	1
5	X5	1	1	1	1	0	1
6	X6	1	1	1	0	0	0
7	X7	1	0	1	1	1	0
8	X8	1	1	1	1	1	1
9	X9	1	1	1	1	1	1
10	X10	1	1	1	1	1	1

Step Two: Calculate the number of questions answered correctly on odd items (X) and even items (Y), then determine X^2, Y^2, and XY.

Table 4. Calculation of Item Scores Based on Odd-Even Data.

No	Responden	Ganjil (X)	Genap (Y)	X^2	Y^2	XY
1	X1	3	3	9	9	9
2	X2	3	2	9	4	6
3	X3	3	3	9	9	9
4	X4	3	3	9	9	9
5	X5	3	2	9	4	6
6	X6	3	0	9	0	0
7	X7	2	2	4	4	4
8	X8	3	3	9	9	9
9	X9	3	3	9	9	9

10	X10	3	3	9	9	9
Score		29	24	85	66	70

Step three: determine the product moment

Table 5. Product Moment Calculation.

No	Responden	Ganjil (X)	Genap (Y)	X^2	Y^2	XY
1	X1	3	3	9	9	9
2	X2	3	2	9	4	6
3	X3	3	3	9	9	9
4	X4	3	3	9	9	9
5	X5	3	2	9	4	6
6	X6	3	0	9	0	0
7	X7	2	2	4	4	4
8	X8	3	3	9	9	9
9	X9	3	3	9	9	9
10	X10	3	3	9	9	9
Score		29	24	85	66	70
Rxy	0,25					

Step Four: determine the spearman-brown reliability test

Table 6. Uji Reliabilitas Spearmen Brown Calculation.

No	Respondent	Ganjil (X)	Genap (Y)	X^2	Y^2	XY
1	X1	3	3	9	9	9
2	X2	3	2	9	4	6
3	X3	3	3	9	9	9
4	X4	3	3	9	9	9
5	X5	3	2	9	4	6
6	X6	3	0	9	0	0
7	X7	2	2	4	4	4
8	X8	3	3	9	9	9
9	X9	3	3	9	9	9
10	X10	3	3	9	9	9
Score		29	24	85	66	70
Rxy	0,25					
r_{11}	0,4					

Step Five: Interpretation of the r11 Result with the Reliability Index Criteria. Based on Table 6, it can be concluded that the reliability test result (r11) using the Spearman-Brown formula obtained a score of 0.4. This means that, when compared to the reliability index criteria presented in Table 1, the score falls into the low category.

4. Conclusion

Reliability testing is an essential stage in quantitative research before a researcher conducts the study. This is because, through reliability testing, it can be determined how consistent respondents' answers are to the research instrument when administered at different

times. In addition, reliability testing allows the instrument to be analyzed according to the obtained reliability index. Several techniques can be applied to test reliability, and researchers may choose the most suitable method depending on their abilities and research needs, based on the techniques that have been mentioned and explained in the previous sections.

References

Abduloh, et al. (2022). Peningkatan dan pengembangan prestasi belajar peserta didik. Ponorogo: Uwais Inspirasi Indonesia.

Amiruddin, Z. (2010). Statistika pendidikan. Yogyakarta: Teras.

Amruddin, et al. (2022). Metode penelitian kuantitatif kualitatif. Jawa Barat: Media Sains Indonesia.

Asfiati. (2020). Visualisasi dan virtualisasi pembelajaran pendidikan agama Islam. Jakarta: Kencana.

Chaerunisa, Z., & Latief, J. (2021). Pengaruh disiplin belajar terhadap prestasi belajar IPS di sekolah dasar. Edukatif: Jurnal Ilmu Pendidikan, 3(5), 2952-2960. https://doi.org/10.31004/edukatif.v3i5.1043

Hanafi, H. (2018). Ilmu pendidikan Islam. Yogyakarta: Deepublish.

Harefa, D. (2020). Peningkatan prestasi belajar IPA siswa pada model pembelajaran learning cycle dengan materi energi dan perubahannya. Trapsila: Jurnal Pendidikan Dasar, 2(1), 25-36. https://doi.org/10.30742/tpd.v2i01.882

Hidayat, A. A. (2021). Menyusun instrumen penelitian & uji validitas-reliabilitas. Surabaya: Health Books Publishing.

Indriyaningrum, K., et al. (2022). Pelatihan metodologi penelitian kuantitatif bagi guru-guru SMK Bhina Tunas Bhakti, Juwana-Pati. Jurnal Penamas, 6(2), 143-148.

Jaya, I. M. L. M. (2020). Metode penelitian kuantitatif dan kualitatif: Teori, penerapan, dan riset nyata. Yogyakarta: Anak Hebat Indonesia.

Kurniawan, H. (2021). Pengantar praktis penyusunan instrumen penelitian. Yogyakarta: Deepublish.

Marisa, S. (2022). Reliabilitas konstruk multidimensional pada structural equation modeling (SEM). Akrab Juara: Jurnal Ilmu-ilmu Sosial, 7(1), 361-368. https://doi.org/10.58487/akrabjuara.v7i1.1783

Misri, A. (2019). Perjuangan muslim di negara konflik. Banda Aceh: Ar-Raniry Press.

Mufidah, A. (2021). Analisis validitas dan reliabilitas instrumen pengukuran mengenal bentuk-bentuk geometri. Jurnal Kualita Pendidikan, 2(3), 192-296. https://doi.org/10.51651/jkp.v2i3.144

Mulyasana, D., et al. (2020). Khazanah pemikiran pendidikan Islam dari wacana lokal hingga tatanan global. Bandung: CV Cendekia Press.

Munte, S., & Polewangi, Y. D. (2022). Pengaruh harga, variasi produk dan kualitas produk terhadap keputusan pembelian keripik singkong saat pandemi COVID-19 di UKM Cap Rumah Adat Minang Medan. Jurnal Ilmiah Teknik Mesin dan Industri, 1(1), 53-68. https://doi.org/10.31289/jitmi.v1i1.1222

Nugroho, U. (2018). Metodologi penelitian kuantitatif pendidikan jasmani. Jawa Tengah: CV. Sarnu Untung.

Pasaribu, E. Z., et al. (2020). Belajar statistika siapa takut dengan SPSS. Indonesia: Yayasan Kita Menulis.

Pramdhan, K., et al. (2021). Pengaruh latihan menggunakan media karet ban dan skipping terhadap hasil long pass sepakbola. Journal of Physical and Outdoor Education, 3(1), 94-100. https://doi.org/10.37742/jpoe.v3i1.122

Riyanto, S., & Hatmawan, A. A. (2020). Metode riset penelitian kuantitatif penelitian di bidang manajemen, teknik, pendidikan, dan eksperimen. Yogyakarta: Deepublish.

Rosita, E., et al. (2021). Uji validitas dan reliabilitas kuesioner perilaku prososial. FOKUS (Kajian Bimbingan & Konseling dalam Pendidikan), 4(4). https://doi.org/10.22460/fokus.v4i4.7413

Rosyid, M. Z., et al. (2019). Prestasi belajar. Malang: Literasi Nusantara Abadi.

Salekha, R. D., et al. (2021). Studi komparasi pembelajaran kelas V SD N Sojomerto 01 Kabupaten Batang-Indonesia dan Feu Roosevelt College Inc Marikina Metro Manila-Filipina. Praniti, I(2), 87-91.

Setiarini, N. I. (2021). Muslim minoritas dan budaya Muslim Melayu masyarakat Pattani Thailand Selatan. Jurnal Penelitian Agama, 22(1), 127-137.

Sugiyono. (2016). Metode penelitian kuantitatif, kualitatif, dan R&D. Bandung: Alfabeta.

Sukesi, K., et al. (2017). Migrasi perempuan, remitansi dan perubahan sosial ekonomi pedesaan. Malang: UB Press.

Susanto, P. H. N., et al. (2022). Metodologi penelitian kebidanan. Sumatera Barat: PT. Global Eksekutif Teknologi.

Syakir, M., et al. (2021). Pengaruh stres kerja, komunikasi kerja dan kepuasan kerja terhadap prestasi kerja karyawan pada PT. Prima Karya Manunggal Kabupaten Pangkep. Amkop Management Accounting Review, 2(1), 20-27. https://doi.org/10.37531/amar.v2i1.141

Tugiman, et al. (2022). Uji validitas dan reliabilitas kuesioner model UTAUT untuk evaluasi sistem pendaftaran online rumah sakit. JATISI (Jurnal Teknik Informatika dan Sistem Informasi), 9(2), 1621-1630. https://doi.org/10.35957/jatisi.v9i2.2227

Ulfah, A. K. (2022). Ragam analisis data penelitian (sastra, riset, dan pengembangan). Madura: IAIN Madura Press.

Veronica, A., et al. (2022). Metodologi penelitian kuantitatif. Sumatera Barat: PT. Global Eksekutif Teknologi.

Yuliani, & Purwanto, E. (2022). Pendidikan Islam di Thailand. Educational Journal: General and Specific Research, 2(1), 51-61.

Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan, 7(1), 17-23. https://doi.org/10.18592/tarbiyah.v7i1.2100