

Review Article

Breaking Down Difficult Theory of Cronbach's Alpha in Simple Concept: A Guide for Questionnaire Reliability Assessment for Medical Researchers

Noureen Durrani^{*1}, Muhammad Imran², Muhammad Kashif³

¹Department of Publication, Liaquat National Hospital, Karachi, Pakistan.

²Center of Excellence in Women and Child Health, Aga Khan University Hospital, Karachi, Pakistan.

³Research Department, Tabba Heart Institute, Karachi, Pakistan.

Abstract: Medical research frequently uses questionnaires to measure abstract notions like behavior, knowledge, and attitude. Determining these instruments' dependability is essential to arriving at accurate and significant interpretations and conclusions. Reliability is often measured by evaluating internal consistency. A statistic that measures internal consistency is Cronbach's alpha. Despite the tool's usefulness, medical undergraduates struggle to understand the concept due to their lack of statistical knowledge. The main focus of this guide is clarify the widespread misconception that higher alpha values are preferable. The articles explain through examples that good internal consistency is achieved by the items' shared meaning, which contributes to the creation of a consistent measure of the same construct. The range of Cronbach's alpha is from 0-1. The range of alpha >0.7 to ≤ 0.9 is regarded as adequate to good reliability. In contrary, redundancy may also be represented by extremely high alpha values (>0.9). Medical researchers who comprehend Cronbach's alpha will be able to create better research tools and, in the end, produce better data for clinical and educational settings.

Keywords: Medical research, Clinical research, Surveys, Reliability, Questionnaire.

BACKGROUND

In medical research, questionnaires are usually used for gathering data on attributes such as knowledge, attitudes, and practice. It is not enough to simply ask questions when designing a questionnaire; there is also a need to design the appropriate questions, in the proper way, with accuracy and consistency. The reliability of the questionnaire is a crucial factor to take into account when trying to provide reliable and rigorous data, whether in surveys, clinical investigations, or academic assessments [1-3]. Making incorrect decisions and drawing false conclusions are likely to result from poorly designed questionnaires or inconsistent data. This makes the concept of reliability, which is defined as the degree to which a measure will yield particular outcomes that are steady and consistent [4-6].

Reliability is crucial for evaluating the abstract variables such as stress, anxiety, satisfaction, knowledge, etc., concepts that are impossible to measure directly like blood pressure or weight. To illustrate the abstract variables, we use collections of connected questions, or items. However, the results could be unclear if the items in a questionnaire don't work or function as intended [5-8]. For instance, if you are designing a questionnaire to measure stress burden and half of the question in study tools are related to stress and rest are related to sleep patterns or physical activity, the overall score for stress would be misleading and questionable.

Cronbach's alpha is one of the most common statistical tool for measuring internal consistency of a questionnaire [9-12]. Despite its usefulness and importance as a measure of reliability, medical researchers often find the concept of statistical knowledge confusing. Investigators may find the term "internal consistency," "reliability coefficient," and "item correlation" too difficult to comprehend and are unable to assess how their understanding of these concepts relate to actual research practice.

By connecting Cronbach's alpha to real-world situations and common questionnaire designing problems, breaking down its formula, explaining its components in simple words, and offering guidance on how to interpret the results it yields, this research aims to simplify its use and comprehension for medical researchers. Medical researchers and those who intend to use academic research or evidence-based practice in their careers should be aware of this basic knowledge.

BASIC CONCEPT OF RELIABILITY AND INTERNAL CONSISTENCY OF A QUESTIONNAIRE

In simple words, questionnaire reliability is the degree to which a questionnaire consistently measures the constructs it is intended to measure [13-15]. Assume, you are using a sphygmomanometer to record a patient's blood pressure. A calibrated sphygmomanometer will provide you approximately identical readings each time if three consecutive are taken while the patient is relaxed, seated correctly, and there is no other cause for the variation. However, sphygmomanometer is not reliable if it reads 140/90, 100/60, and 180/110 while the patient remains stable under the same condition. Now imagine a questionnaire was developed to

* Address correspondence to this author at the Department of Publication, Liaquat National Hospital, Karachi, Pakistan.
Email: noureen.durrani@lnh.edu.pk

assess medical students' knowledge about diabetes. A reliable set of questions acts as an accurate sphygmomanometer; if the fact or conditions are not changed, the questionnaire would give the similar results every time it will be used on the same participant. However, if some of students' scores 90% in the first time and then it drops to 50% on re-evaluation after few hours assuming that in real their knowledge has not been changed, that's an indication of unreliable questionnaire most probably due to ambiguous questions. A reliable set of questions always gives consistent responses, not random ones.

RELIABILITY – STATISTICAL CONCEPT

In statistics, the questionnaire reliability is defined as the degree to which a questionnaire reliability evaluates a construct (concept/variable) throughout time, across items, and across raters [13, 16]. Reliability can be expressed statistically as the extent to which a measurement is free from random error and produces consistent findings under specific conditions [17-19].

Reliability is conceptually defined as the percentage of total variance that cannot be attributed to measurement error. As a result, it provides information on the signal-to-noise ratio in a group of data, the overall consistency of a measure, and the distinguishability of individual measurements. Accordingly, a reliability of 0 indicates that measurement error accounts for all variability, whereas a reliability of 1 indicates that all variability is due to actual differences and that there is no measurement error. True and error-related variance are equivalent when the reliability is 0.5 [16].

INTERNAL CONSISTENCY OF QUESTIONNAIRE

In cross-sectional studies that measure knowledge, attitudes, or constructs such as stress, self-esteem, burnout, etc., "reliability" primarily refers to "internal consistency". In a questionnaire

designing context, internal consistency is the extent to which each item measures the same underlying construct and is consistent with the others items [20-22]. Stated differently, internal consistency assesses how well a questionnaire's items "work together". For example, if a questionnaire is used to gauge medical students' stress levels, then each item should evaluate stress in the same way. If the survey participants responds strongly agree to one item indicating higher stress, they are more likely to respond similar to other stress related items, reflecting a good internal consistency.

Two hypothetical versions of a questionnaire measuring stress among medical professionals are compared in Table 1; one has a good internal consistency (good questionnaire) and the other has low internal consistency (bad questionnaire). Every question on the surveys explains its relevance and how it affects the internal coherence of the ones being constructed.

Every question on the Good Questionnaire has something to do with stress. One item, headaches during hectic workdays, is indirectly associated with stress experiences since it may be a physical indication of stress, while three items including feeling overburdened, difficulty relaxing, and tension with patient care, are directly related to stress experiences. Good internal consistency is achieved by the items' shared meaning, which contributes to the creation of a consistent measure of the same construct.

The Poor Questionnaire, on the other hand, has both relevant and unrelated items. In particular, "I like to try new cuisines at restaurants" and "I prefer listening to music when driving," these two items in a questionnaire, have nothing to do with the stress construct. Because inconsistent responses add a layer of noise, the presence of some irrelevant items causes response variability, which runs counter to the concept of internal reliability. This illustration emphasizes how important it is for achieving desirable instrument's reliability to only include elements that conceptually belong together.

Table 1. Comparison of Good and Poor Questionnaire items and their Effect on Internal Consistency.

Item	Good Questionnaire (High Internal Consistency)	Relevance	Effect on Internal Consistency	Poor Questionnaire (Low Internal Consistency)	Relevance	Effect on Internal Consistency
Item 1	I feel overwhelmed at work.	Relevant	Enhances internal consistency	I feel overwhelmed at work.	Relevant	Enhances internal consistency
Item 2	I find it hard to relax after my shift.	Relevant	Enhances internal consistency	I enjoy trying new cuisines at restaurants.	Irrelevant	Weakens internal consistency
Item 3	I feel tense or anxious during patient care.	Relevant	Enhances internal consistency	I feel tense or anxious during patient care.	Relevant	Enhances internal consistency
Item 4	I experience headaches during busy workdays.	Relevant (indirect)	Mildly enhances internal consistency	I prefer listening to music while driving.	Irrelevant	Weakens internal consistency

CRONBACH'S ALPHA

For instance, a researcher is designing a survey questionnaire to gauge medical students' stress levels before exams. Ten questions are drafted, some of which deal with sleep, some with

attention, and some of which just ask about anxiety. Now, the goal is to ensure that all these items are truly measuring stress. This is the point where Cronbach's alpha is used for determining whether all items are working together to reflect the same underlying phenomenon [23].

STATISTICAL DESCRIPTION IN SIMPLE WORDS

In 1951, Lee Cronbach developed alpha statistic [24]. In terms of statistics, Cronbach's alpha (α) evaluates the proportion of total variance in questionnaire scores that is real versus variation because of random error or inconsistencies in the item themselves [25]. Both true variation and mistake variation are reflected in a person's overall score each time they respond to a series of questions. Differences that accurately show how people differ in a certain notion (stress, knowledge, contentment, etc.) are indicative of true variance. Differences brought about by unclear questions, ambiguity, guesses, and other random variation are indicative of error variation.

CRONBACH'S ALPHA FORMULA

$$\alpha = \frac{N}{N-1} \left(1 - \frac{\sum \sigma_{item}^2}{\sigma_{total}^2} \right)$$

α = The reliability coefficient (ranging from 0-1)
 N = Number of items in a questionnaire
 $\sum \sigma_{item}^2$ = The sum of variances of each individual item
 σ_{total}^2 = Variance of total score

UNDERSTANDING STATISTICAL FORMULA OF CRONBACH'S ALPHA

Despite its complexity, the Cronbach's alpha formula is actually quite straightforward when viewed in light of simple words. The number of items is N. For instance, if a questionnaire has six questions, N=6 and the sum of the variances of each question is $\sum \sigma_i^2$. The variance of each individual item denotes that how participants' responses vary on that item. This number would be high if each person's response to the question varied greatly. The variance of the total score is denoted by σ^2 . Following the summation of participant's responses on each item, this represents the overall variation in total scores among survey participants. It's simply explains how each participant's overall score varies.

Since the numerator part in formula represents a variation in each individual question and the denominator part shows variation in total score. Thus the variance ratio in formula indicate how much of the total variance is because of item-level variance which is considered error. However, we are interested in knowing true differences in scores rather than knowing the random error. So to extract this, we subtract it from 1 which gives the true variance that is not due to random error-and simply shows a reliable part. In last, the estimates are adjusted for the sample by multiplying it with correction factor. This correction factor helps to compensate for the tendency to slightly underestimate reliability when the number of items is small, thus making the final alpha value a more accurate estimate of internal consistency.

DESIRABLE VALUE OF CRONBACH'S ALPHA AND ITS INTERPRETATION

Cronbach's alpha values ranges from 0-1 [26, 27]. Based on rule of thumb these values are classified as:

≥ 0.9 indicates excellent internal consistency

-0.8-0.9 indicates good internal consistency

-0.7-0.8 indicates acceptable internal consistency

< 0.7 indicates improvement is required in developed questionnaire

The formula of Cronbach's alpha poses a question that "How much of the differences in total scores are due to random errors in individual questions, and how much reflects real and reliable differences between the individual's answers?"

An alpha value of 0.4, indicates that 40% of variation is from real reason of actual differences between responses whereas 60% of variance is because of random error or inconsistencies between questions. Conversely, an alpha value of 0.85 indicates that only 15% of the variations in total scores are due to chance errors or inconsistencies in the questions, while roughly 85% of the variations represent actual, consistent variances in scores between the respondents [25]. This high estimate suggests the questionnaire's excellent reliability showing the all the questions working together to measure the underlying concept. This means that alpha will be high if the variations of the individual questions are minor (everyone answers the same questions) and the variance of the total score is large (different people score differently overall). It's a positive indication! It shows that the same underlying notion is being measured consistently by the items.

DATA CONSIDERATIONS FOR CRONBACH'S ALPHA CALCULATIONS

Cronbach's alpha can be used depending on the type of data obtained from samples at question although it works best with questionnaires generating continuous data where the variability is large and therefore there is a little less variance in the reliability of results calculated from that type of input. Generally, though, the coefficient is most frequently used with data generated using ordinal scales such as Likert for data generated via questionnaires typically classified into 5 or 6 categories (e.g., Strongly Disagree - Strongly Agree). While responses are technically ordinal scales, they can still be treated as including relatively constant-scale results so that samples producing continuous data produce coefficient results very similar to those derived from interval data, therefore, the coefficients would indicate generally a higher reliability than if treated as true ordinal scales. If the sample is dichotomous, such as "yes" or "no" or "correct" or "incorrect", the calculation of coefficient will work; however, the input must be converted to binary numbers in order to maintain compatibility with application form of interval data. For example, if you give the respondent a list of 10 different options in an MCQ, the respondent must choose only one correct answer per question, which allows for the proper use of Cronbach's alpha to determine whether each item measures the same underlying attribute (e.g., knowledge of a subject area). On the other hand, if you give the respondent a list of 10 different response options from which they may choose multiple options, and each option is independent from the others, the use of Cronbach's alpha for

these types of questions will not be appropriate because they do not represent a single underlying (latent) construct. Therefore, either descriptive statistics or comparative statistical methods such as the Cochran's Q TEST would typically be used for the analysis of this type of data.

A STEP-BY-STEP MANUAL FOR CRONBACH'S ALPHA CALCULATION

Assume a 5-item scale was developed to measure burn-out in final year medical students. Five students are surveyed, and data is gathered. Assume that each student's response fell between Strongly Disagree and Strongly Agree on a scale of 1 to 5 (Table 2). An alpha value of 0.94 is obtained by inserting the values displayed in the Cronbach's alpha formula. This means that 94% of the variations in the overall scores is genuine, consistent differences between students rather than sporadic, random errors or mistakes in the questions. Random noise or variances in responses account for just around 6% of the differences.

Table 2. Hypothetical Students' Feedback on Burn Out Questionnaire (5-point Likert scale) with Item Variances.

Questionnaire items	Number of students					Variance
	1	2	3	4	5	
I feel emotionally drained after clinical rounds.	4	3	5	2	4	1.3
I find it hard to relax after shifts.	3	2	4	2	3	0.7
I dread going to the hospital.	5	4	5	3	4	0.7
I feel overworked.	4	3	5	2	4	1.3
I often think about quitting medicine.	3	2	4	3	3	0.5
Total	19	14	23	12	18	18.2

CONFUSION IN CRONBACH'S ALPHA CONCEPT

One of the common confusions is that the variation within individual questionnaire items must be greater than the variation in the overall scores. This confusion arises because different students will respond differently to the same item. However, it is true that responses on individual item varies due to different participants answering to it but the total score of each individual captures something more than this. The total score variation of a participant captures how consistently participant response to all question. Students who perform well on one question also typically perform well on the others and generally the same could be said about lower scores. Because of this fact, their total scores are more variable compared to the variability of any single item score. A higher variance in total score shows that questionnaire items are working together to measure the same underlying concept because it differs those performing well and scoring low. Therefore, rather than evaluating random differences within items individually, Cronbach's Alpha measures

degree of internal consistency based on overall score which is a good indicator of differentiating between those having high scores and low scores.

Interpreting an extremely high Cronbach's alpha (beyond 0.9) as continuously positive is another common misconception. Since alpha indicates reliability, many people assumes that the value closer to 1 is better. An excessively high alpha can imply redundancy, which means that responders are perceiving responses that are too similar or repeating, even if a high alpha indicates consistency [25, 28]. This could make the questionnaire boring and excessively lengthy without offering more in-depth information. Thus, be aware that a very high alpha can necessitate evaluating the scale to see if it can be shortened as much as feasible without losing enough reliability.

ACTIONS NEEDED ON THE BASIS OF CRONBACH'S ALPHA

Cronbach's alpha aids in understanding how reliable a questionnaire is, and depending on its value, different actions are recommended:

- An alpha of less than 0.5 is unacceptable [23]. This suggests that a large number of the questions are either not measuring the same underlying construct or are ambiguous or inconsistent. Actions to be taken include carefully go over the items and make any necessary revisions, such as rewording the questions, eliminating those that might be problematic, or redesigning the scale entirely.
- A value of alpha between 0.5 and 0.7 indicates moderate or doubtful consistency [29]. This suggests that the questionnaire has consistency but still it needs improvement. To deal with this, examine the item to determine which questions are the weakest, then make focused, minor adjustments to ensure consistency and clarity.
- The range of alpha >0.7 to ≤ 0.9 is regarded as adequate to good reliability [11]. The measure or construct being assessing is adequately captured by the questionnaire, which is consistent. Although the tool can be utilized with confidence but there is still need for some small improvements based on the content review by other subject expert.
- An Alpha above 0.90 indicates great reliability, which is usually a desirable; however, it may indicate redundancy, meaning that the questions are excessively repetitive or similar. Determine whether any items are unnecessary; consider condensing the questionnaire while preserving a valid measurement [30].
- Many students believe that an instrument measures only one construct if the alpha is high. However, Cronbach's alpha does not inform users of the instrument's uni-dimensionality; rather, it merely speaks to the test's internal consistency. For instance, if you develop the

questionnaire with items regarding stress and time management by random (which is possible), the items may be related to each other and provide a high alpha value, which indicates that the questionnaire evaluates two things. Consequently, alpha should be used in conjunction with exploratory factor analysis.

CONCLUSION

In medical education and research, a valuable and consistent data can be obtained only when a reliable questionnaire is used. Cronbach's alpha is a key tool for measuring internal consistency which shows whether the items in a questionnaire are working well together. The consistency pattern is evident with high Cronbach's alpha value preferably 0.7 to 0.9. The variance in overall score is an indicator of variation among individual responses to overall questionnaire. A low alpha value indicates that there may be poor item alignment and that random error or even poorly constructed questions may account for a large portion of the variance in the final score. In contrary, redundancy may also be represented by extremely high alpha values (>0.9). Medical researchers who comprehend Cronbach's alpha will be able to create better research tools and, in the end, produce better data for clinical and educational settings.

AUTHORS' CONTRIBUTION

Nooreen Durrani: Conceptualization, Literature review, Work Design, interpretation, Writing Draft, Final approval, final proof to be published.

Muhammad Imran: Writing Draft, Final approval, final proof to be published.

Muhammad Kashif: Critical review and revision the manuscript, Final approval, final proof to be published.

ACKNOWLEDGEMENTS

Declared none.

ETHICAL CONSIDERATIONS

Data Availability Statement

The data used for illustrative purposes are entirely hypothetical and were generated using artificial intelligence. No real patient data were used or shared.

Consent to Participate

No human participants were involved; therefore, ethical approval was not required.

Consent for Publication

All author take the responsibility of the work and provide consent for publication.

Conflict of Interest

Declared none.

Competing Interest/Funding

Declared none.

Use of AI-Assisted Technologies

Claude AI technology was used to generate data of hypothetical examples and language editing.

REFERENCES

- [1] Ahmed I, Ishtiaq S. Reliability and validity: importance in medical research. *Methods* 2021; 12: 2401-6.
- [2] Hossan D, Wolfs B, Petkovic M. Questionnaire validity and reliability: A review with practical guidelines. *J Entrep Bus Econ* 2025; 13(1): 135-86.
- [3] Koo M, Yang SW. Questionnaire use and development in health research. *Encyclopedia* 2025; 5(2): 65.
- [4] Karnia,R. Importance of reliability and validity in research. *Psychol Behav Sci* 2024; 13(6): 137-41.
- [5] Govindasamy P, Cumming TM, Abdullah N. Validity and reliability of a needs analysis questionnaire for the development of a creativity module. *J Res Spec Educ Needs* 2024; 24(3): 637-52.
- [6] Tourangeau R. Survey reliability: Models, methods, and findings. *J Surv Stat Method* 2020; 9(5): 961-91. doi: 10.1093/jssam/smaa021. PMC8665769.
- [7] Taherdoost H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *Int J Acad Res Manag* 2016; 5: 28-36.
- [8] Sürücü L, Maslakci A. Validity and reliability in quantitative research. *Bus Manag Stud: An Int J* 2020; 8(3): 2694-726.
- [9] Streiner DL. Starting at the beginning: An introduction to coefficient alpha and internal consistency. *J Pers Assess* 2003; 80: 99-103.
- [10] Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ* 2018; 48(6): 1273-96.
- [11] Barbera J, Naibert N, Komperda R, Pentecost TC. Clarity on Cronbach's alpha use. *J Chem Educ* 2020; 98(2): 257-8.
- [12] Edelsbrunner PA, Simonsmeier BA, Schneider M. The Cronbach's alpha of domain-specific knowledge tests before and after learning: A meta-analysis of published studies. *Educ Psychol Rev* 2025; 37(1): 4.

- [13] Mo Z, Li X, Zhai Y, Men Y, Tang Y, Qiao J, *et al.* Reliability and validity of a questionnaire measuring knowledge, attitude and practice regarding “oil, salt and sugar” among canteen staff. *Scientific Rep* 2023; 13: 20442.
- [14] Ferrando PJ, Navarro-Gonzalez D. Reliability and external validity of personality test scores: The role of person and item error. *Psicothema* 2021; 33(2): 259-67.
- [15] Kumar S, Dinesh N, Periasamy P. Testing validity and reliability of the questionnaire in soft skills research: A perspective from b-school alumni. *Int J Entrep* 2021; 25: 1-10.
- [16] Matheson GJ. We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *Peer J* 2019; 7: e6918. doi: 10.7717/peerj.6918.
- [17] Turgut MF, Baykul Y. Measurement and evaluation in education. 4th ed. Ankara: Pegem Academy 2012.
- [18] Mokkink LB, Eekhout I, Boers M, van der Vleuten CPM, de Vet HCW. Studies on reliability and measurement error of measurements in medicine - from design to statistics explained for medical researchers. *Patient Relat Outcome Meas* 2023; 14: 193-212. doi: 10.2147/PROM.S398886.
- [19] Elgestad Stjernfeldt P, Sjögren P, Wårdh I, Boström AM. Systematic review of measurement properties of methods for objectively assessing masticatory performance. *Clin Exp Dent Res* 2019; 5(1): 76-104. doi: 10.1002/cre2.154.
- [20] Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess* 2003; 80: 99-103.
- [21] Ranganathan P, Caduff C, Frampton CMA. Designing and validating a research questionnaire - Part 2. *Perspect Clin Res* 2024; 15(1): 42-5. doi: 10.4103/picr.picr_318_23.
- [22] Internal Consistency. [cited February, 1, 2026]. Available from: <https://www.sciencedirect.com/topics/nursing-and-health-professions/internal-consistency>
- [23] Izah SC, Sylva L, Hait M. Cronbach's alpha: A cornerstone in ensuring reliability and validity in environmental health assessment. *ES Energy Environ* 2023; 23: 1057.
- [24] Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16: 297-334.
- [25] Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53-5.
- [26] Sheposh R. Cronbach's alpha. 2024; Available from <https://www.ebsco.com/research-starters/science/cronbachs-alpha>
- [27] Bujang MA, Omar ED, Baharum NA. A Review on sample size determination for cronbach's alpha test: A simple guide for researchers. *Malays J Med Sci* 2018; 25(6): 85-99. doi: 10.21315/mjms2018.25.6.9.
- [28] Alkhadim GS. Cronbach's alpha and semantic overlap between items: A proposed correction and tests of significance. *Front Psychol* 2022; 13: 815490. doi: 10.3389/fpsyg.2022.815490.
- [29] Schrepp M. On the usage of cronbach's alpha to measure reliability of UX scales. *JUX* 2020; 15(4): 247-8.
- [30] Zakariya YF. Cronbach's alpha in mathematics education research: Its appropriateness, overuse, and alternatives in estimating scale reliability. *Front Psychol* 2022; 13: 1074430.

Received: February 09, 2026

Revised: March 03, 2026

Accepted: March 11, 2026

© 2026. The Authors, National Journal of Health Sciences.

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.