

Considerations in Questionnaire Development: A Review

Ipsita Mahapatra¹, Ramesh Nagarajappa², Dharmashree Satyarup³, Sharmistha Mohanty¹

¹Postgraduate Trainee, ²Professor and Head, ³Professor, Department of Public Health Dentistry, Institute of Dental Sciences, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India

Abstract

A fundamental part of good research is concerned with making sure that the design of the questionnaire addresses the needs of research. After a proper administration of the questionnaire, a questionnaire constantly measures what it intends to measure thereby giving us confident results of a study. It has been documented in several studies regarding Validity and Reliability, as an important source of measuring the accuracy and consistency of research instruments (especially questionnaires). But their measure is not observed to be practiced in developing countries among health plus social science researchers. This review article comprehensively explores and describes questionnaire design and measures to validate a questionnaire including its reliability testing.

Keywords: Questionnaire, Reliability, Validity.

Introduction

The questionnaire can be defined as “a tool for collecting and recording information about a particular issue of interest”. It is a proficient means for estimating the “behavior, attitudes, sentiments, opinions and intentions of generally large quantities of subjects” more economically and rapidly against different strategies. Questionnaires are a good research method because it can be accessible to a large number of people relatively easily and economically. They also provide quantifiable answers for a research topic, hence being easy for analysis.

A definite purpose of Questionnaires is to be related to the objectives of the research. They are prepared in various languages, expanding investigation into new measurements by leading examinations in various

cultural groups. Validation of questionnaires in those individual languages lead to exact and strict data about the specific cluster.¹

Design of the Questionnaire: Careful consideration must be given to the design of the questionnaire for gathering useful and relevant information. A lot of thought and effort is required for a well-designed questionnaire along with proper planning and development through various stages (Figure 1).

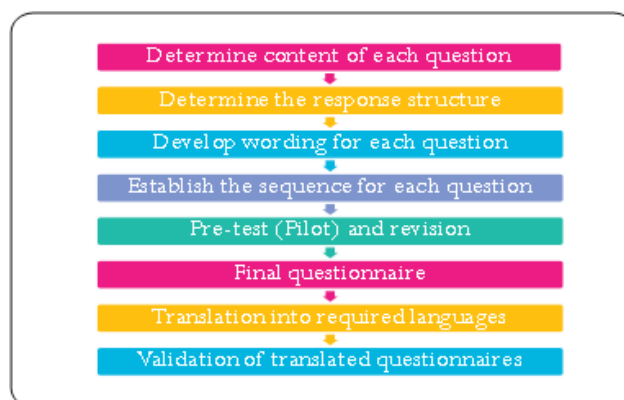


Figure 1: Questionnaire Design

Development of a Questionnaire: The numbering and order of the questions should be in a style that is reasonable in the participant’s view. Usually “Funneling” method is utilized which starts from basic questions

Corresponding Author:

Dr. Ramesh Nagarajappa

Professor and Head, Department of Public Health Dentistry, Institute of Dental Sciences, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India

e-mail: rameshpcd@yahoo.co.in

to comfort participants and afterward centers down to progressively more explicit questions.

Step 1: Determining the contents of the questionnaire: Determining what topics, the questionnaire will cover and how it will be administered is the first task. Based on the study objectives, the number of domains is decided. Posing a large number of questions that are not straight forwardly helpful or significant is the most frequent problem in questionnaire design. It can be organized as follows:

1. The prime objectives of the study should be outlined.
2. Questions relevant to the objectives must be asked and only necessary information should be recorded.
3. The lists should be rearranged and organized into distinguished sub-themes or sections of the questionnaire.
4. The decision for the process of filling up of questionnaire by participants themselves or by an interviewer should be taken.²

Step 2: Formulation of the questions: Once the sections of the questionnaire have been identified, individual questions must be formulated for gathering the particular pieces of information needed. To deliver significance and purpose of the question to the participants, a step called “wording questions” (selection of words or phrases) is practiced. This ensures that they are uniformly understood by the respondents. Below are the general rules on good question formulation:

- i. **Each question must be clear, simple, and specific:** Questions must be indicated in simple, short and clear language. Such as the question can be broken into pieces to assist participants or they may be asked to reveal purposely the different measures taken by them since the beginning of the illness. The following questions may be involved: Did your child suffer from toothache within the past three weeks? (If yes): “Did you visit anyone for advice or treatment? Did you give the child any medicines?” (If participant mentioned medicine(s) were given): “Would you please list any medication(s) that the child has used since the toothache started?”
- ii. **Each question must measure one thing at a time:** Measuring two or more dissimilar things in a question is incorrect. Such as questions like “How do you and your staff normally treat children who present with a toothache?”, requires partition

into numerous questions to permit participants for providing individual answers as there may be a favored method for treating toothache by different prescribers in the office.

When presumptions are made about the participants, Leading questions may also arise. For instance, the question: “What medication did you give your child when he got toothache?” is biased since the respondent would give the child medication, because of the presumption that the child had a toothache. An improved version may be: “Has your child had toothache within the last week?”

(If yes): “Did you take any action?”

(If yes): “What action did you take?”

- iii. **Questions must be free from ambiguity:** It is advisable to exclude indistinctly defined or words having double meanings. The question: “What kind of patient was he?” Might generate responses like “poor” or “rich,” “simple” or “complex,” “cooperative” or “un-cooperative,” etc.

“Double-barreled” questions like: “Do you think amoxicillin and paracetamol are effective for treating toothache?” can’t be good questions as a participant could like any one of them. In such a situation, having a Yes/No reply would not truly mirror the participant’s opinion. Instead it is worthy to divide it into two questions with individual thought.

Step 3: Sequence of the Questions:

The questions must be organized in a sequence after framing them. Following approach is helpful for participants:

- (a) The topics should follow a logical order. A smooth flow of thoughts running from general questions within a topic to gradually more specific ones.
- (b) Focus should be on interesting and relatively “non-controversial” questions, e.g., “How many years have you worked at this place?”
- (c) Sensitive questions like personal information regarding age, education, occupation, and income may create unwillingness among the participants if asked early. Hence they should be placed at the end, after establishing a relationship of trust.
- (d) Questions influencing answers to a different question should be separated.

(e) Clear indication of instructions for skipping one/ more questions

Step 4: Arrangement of the Questionnaire

The layout of the final draft questionnaire should be made both “consumer” and “user” friendly. It includes:

1. Required headings and spaces for labeling.
2. Identification of all questionnaires by recognizing information for participant, date and place of interview, and name of the interviewer.
3. Start each section of the questionnaire with necessary instructions.
4. Sufficient space between questions.
5. For pre-categorized answers, be constant with codes or boxes.
6. Enough Space to be provided for answers to open-ended questions

Step 5: Conducting Pilot tests:

Pilot testing of the questionnaire is done with the help of applied qualitative method. The developed questionnaire is tried in a group and a similar environment to that of the original study environment. This will provide an opportunity for interviewers to perform all the activities that are estimated to take place in the field.

Step 6: Translation of the Questionnaire: Most interviews are conducted in local languages. A thorough translation of Questionnaires is done before the fieldwork begins for ensuring consistency while using the words and meanings. Back translation of the questionnaire into the original language it was written is done by a different translator. A comparison of both the versions is then done to rule out any differences.

Step 7: Validation of the Questionnaire:

“Validity is the degree to which an assessment measures what it is supposed to measure”. The different types of validity are mentioned below:

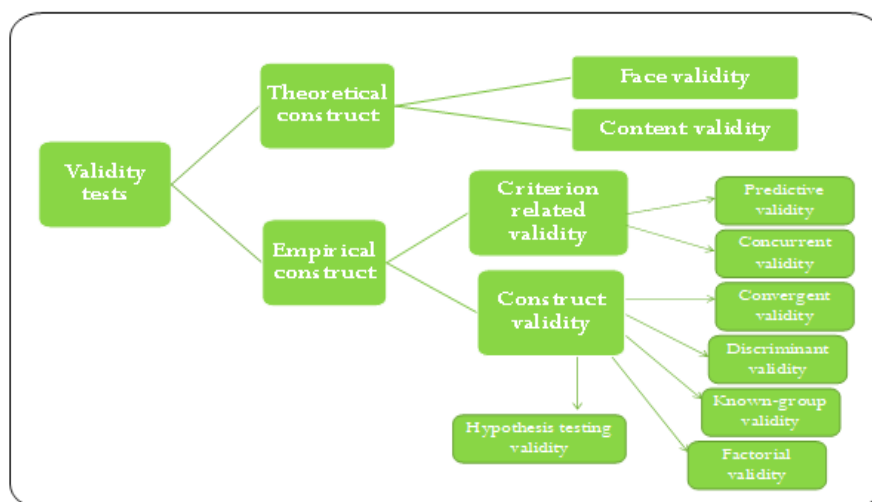


Figure 2. Validity test of the questionnaire

Face Validity: Face validity, “being a form of content validity is assessed by having ‘experts’ (clinicians/clients/researchers) review”. The contents of the questionnaire are being reviewed by an expert on the research subject to see if the items seem appropriate.

Content Validity: “It pertains to the degree to which the instrument fully assesses or measures the construct of interest”.³⁻⁷ It is established by experts well-known with the research subject. The experts review

the questionnaire items for “readability, clarity, and comprehensiveness” and get nearer to a certain level of agreement for including the required items in the final questionnaire. Based on the ratings of the items in the questionnaire, a significant level of inclusion of an item requires a CVI (Content Validity Index) of ≥ 0.78 .⁵

Criterion-Related Validity: Criterion-related validity is “assessment of the relationship of scores on a test to a specific criterion”.^{8,9} It is a measurement of

“how well questionnaire findings pile up against another instrument”^{9,10}

Concurrent Validity: Concurrent validity is “the ability of a test to predict an event in the present form”. Here a highly rated existing standard (gold standard) is used for comparison against a newly developed questionnaire.⁴ Concurrent validity comes into consideration when the criterion and the measure coincide at the same time.^{9, 11-14}

Predictive Validity: It is the “ability of a test to measure some event or outcome in the future”.^{11,14}. Predictive validity deals with the capability of the questionnaire to predict future events, behavior, attitudes or outcomes. The correlation coefficient is used here.

Construct Validity: It is the “degree to which an instrument measures the trait or theoretical construct that it is intended to measure”.¹⁵ Instead of using a criterion for comparison, a hypothetical construct is utilized.¹⁵ It measures how meaningful the scale or instrument is when it is in practical use. Depending on the research problem, four types of validity tests serves the function of construct validity:

Convergent Validity: In convergent validity we check the measures that should be related are related. For example, we construct “education” by asking people how much education they have completed. Then, this record can be compared with their institutional records and a test of school-level knowledge. Both the measures should converge or be similar for a good convergent validity.

Discriminant Validity: It proofs that one concept differs from other closely related concepts.¹² For instance, we have 10 items that measure good oral habits. People respond to all 10 in similar ways. But in the same questionnaire, we have also put 5 questions that measure bad oral habits. Our questionnaire holds good discriminant validity of the 10 good oral habits items stuck together as well as show a negative association with the 5 bad oral habit items.⁴

Known-group Validity: Here comparison between two groups is done within a known attribute. Such that a higher score is given to a group with related attributes while a lower score is given to the group with an unrelated attribute.^{16,18-20} For instance, a questionnaire was used to investigate depression between two groups of patients with and without a clinical diagnosis of depression.

As it is predictable in clinically diagnosed patients of depression, the score for the construct of depression in the questionnaire will be scored higher in comparison to those without the diagnosis.

Factorial Validity: Being an empirical extension of content validity, it “validates the contents of the construct employing the statistical model called factor analysis”.¹¹ While analyzing factorial validity within a construct of interest, the several items which measure a particular dimension are meant to be highly associated with each other than those measuring other dimensions.¹⁶ Such as there are eight dimensions in health-related quality of life questionnaire (HRQoL). Hence it is supposed that under the domain social function (one dimension from HRQoL) all items of the questionnaire measuring it must be highly associated than those items measuring another domain like mental health.²⁰

Hypothesis-Testing Validity: It tests “the relationship between a measured concept (variable) or other concepts (variables), derived from a theory, thus providing evidence that a research hypothesis is supported”.¹⁷ Suppose in case of social learning theory revealing how observing physical violence in television leads to violent behavior, a hypothesis is derived from this theory stating a positive correlation between both variables. If the hypothesis is supported by the evidence collected, it is concluded that a higher degree of construct validity is observed in the measurements of physical aggression and viewing of televised physical violence. It is declared after measuring and examining the two theoretical concepts in the hypothesis-testing process.⁴

Reliability of a questionnaire: “Reliability is an extent to which a questionnaire, test, observation or any measurement procedure produces the same results on repeated trials”. In short, it is the stability or consistency of scores over time or across raters.^{3,4} A sign of lack of reliability exists due to divergences between observers or instruments. Usually carried out by a pilot test, it could be assessed in three major forms discussed below:

1. Test-retest reliability: “The ability of the questionnaire to provide similar results when delivered to the same participant on two separate occasions to see how stable the responses are”. Generally, a correlation coefficient (r) value ≥ 0.70 is considered as good.²⁰ There is a scope for the memory of the last answer to affect test-retest practice as the participants become known with the terms from the previous test.

2. Alternate-form reliability (or equivalence):

“It is referred to as the measure of agreement between two or more research instruments like two different questionnaires on a research construct that are delivered at nearly the same point in time”.³ Followed by a parallel form procedure, in which alternative forms of the same measure is delivered to either the same group or different group of participants. Questions or responses are reworded and rearranged to create two items that are related and un-identical. Delivery of the various forms can be at the same time or following some intervals. More similarity between the two forms indicates a higher degree of correlation between them.⁴

3. Internal consistency: “It is the degree to which the subjects answer similar questions similarly”. Two equivalent forms of a questionnaire are administered at the same time to a subject and assessed for internal consistency. Other method include the “split-half reliability index” and “coefficient alpha index”.²¹ Sometimes, “Kuder–Richardson formula 20 (KR-20) index” can also be used.³

Intraclass Correlation (ICC): Intraclass correlation measures “the reliability of ratings or measurements for clusters- data that has been collected as groups or sorted into groups”. It describes how strongly the units in the group resemble each other. The ICC ranges from 0 to 1. High similarity exists between values from the same group if the ICC value is close to 1. If the ICC value is close to zero, it indicates that values from the same group are not similar.²¹

Conclusion

Good questionnaire design is a key to obtaining good survey results. Based on validity and reliability testing, questions are re-written, removed or included, until the researcher feels the questionnaire met the standards set by him. In this article we have discussed the various ways of testing the validity and reliability of a questionnaire which will help the researcher to improve his knowledge and skills.

Conflict of Interests: None

Ethical Permission: Approved

Funding: None

References

1. Patel HR, Joseph JM. Questionnaire Designing Process: A Review. *J Clin Trials* 2016; 6: 255.
2. Structured Questionnaires. Available at http://archives.who.int/PRDUC2004/RDUCD/INRUD_2000_CDROM/Manuals/Qualitative_Method_Manual/qm_ch6.doc [Last accessed on Jan 27, 2020].
3. Miller MJ. Graduate Research Method. Available from: www.michaeljmillerphd.com/res500.../reliability_and_validity.pdf. [Last accessed on Jan 27, 2020].
4. Bolarinwa OA. Principles and method of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J* 2015;22:195-201.
5. Sangoseni O, Hellman M, Hill C. Development and validation of a questionnaire to assess the effect of online learning on behaviors, attitude and clinical practices of physical therapists in United States regarding of evidence-based practice. *Internet J Allied Health Sci Pract* 2013;11:1-12.
6. DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh* 2007;39:155-164.
7. Polit DF, Beck CT. The content validity index: Are you sure you know what’s being reported? Critique and recommendations. *Res Nurs Health* 2006;29:489-497.
8. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: Theory and application. *Am J Med* 2006;119:166.e7-16.
9. Liang Y, Laua PW, Huang YW, et al. Validity and reliability of questionnaires measuring physical activity self-efficacy, enjoyment, social support among Hong Kong Chinese children. *Prev Med Rep* 2014;1:48-52.
10. Wong KL, Ong SF, Kuek TY. Constructing a survey questionnaire to collect data on service quality of business academics. *Eur J Soc Sci* 2012;29:209-221.
11. Drost EA. Validity and reliability in social science research. *Educ Res Perspect* 2011;38:105-123.
12. Booth ML, Okely AD, Chey TN, Bauman A. The reliability and validity of the adolescent physical

- activity recall questionnaire. *Med Sci Sports Exerc* 2002;34:1986-1995.
13. Pedisic Z, Bennie JA, Timperio AF, et al. Workplace sitting breaks questionnaire (SITBRQ): An assessment of concurrent validity and test-retest reliability. *BMC Public Health* 2014;14:1249.
 14. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. *J Clin Hypertens (Greenwich)* 2008;10:348-354.
 15. Wong KL, Ong SF, Kuek TY. Constructing a survey questionnaire to collect data on service quality of business academics. *Eur J Soc Sci* 2012;29:209-221.
 16. Engel RJ, Schutt RK. Measurement. *The Practice of Research in Social Work*. 3rd ed., Ch. 4. Sage Publication Inc. (Online); 2013. p. 97-104. Available from: https://www.us.sagepub.com/sites/default/files/upm-binaries/45955_chapter_4.pdf. [Last accessed on Jan 28, 2020].
 17. Wells CS. Reliability and Validity; 2003. Available from: <http://www.journalism.wisc.edu/~dshah/./Reliability%20and%20Validity.pdf>. [Last accessed on Jan 28, 2020].
 18. Deniz MS, Alsaffar AA. Assessing the validity and reliability of a questionnaire on dietary fibre-related knowledge in a Turkish student population. *J Health Popul Nutr* 2013;31:497-503.
 19. De Vellis RF. *Scale Development: Theory and Applications*. 3rd ed. Thousand Oaks, California: SAGE; 2012.
 20. Singh AS, Vik FN, Chinapaw MJ, et al. Test-retest reliability and construct validity of the ENERGY-child questionnaire on energy balance-related behaviours and their potential determinants: The ENERGY-project. *Int J Behav Nutr Phys Act* 2011;8:136.
 21. Anderson AS, Bell A, Adamson A, Moynihan P. A questionnaire assessment of nutrition knowledge: Validity and reliability issues. *Public Health Nutr* 2002;5:497-503.