
Editorial

Questionnaire structure – how much do editors need to know?

DOI: 10.20316/ESE.2018.44.18011

Keywords: analysis, editors, questionnaire

Conducting a survey, ie collecting data with a questionnaire, is a very popular method of research, particularly in medicine and the social sciences, with the results frequently published. The data may be collected by self report, ie the participant fills in the data alone, now often online, or by a researcher. The use of a questionnaire as a quantitative method was presented by the Statistical Society of London in the first half of the 19th century. From that time, the methodology of using questionnaires (development, measuring and validation) was mainly established by psychologists. New questionnaires, often called scales, are still being created and described in the literature.

As an editor, author and reviewer, I have noticed that the method is widely (mis)used and the results are frequently misinterpreted. I would like to address some typical issues of which journal editors (with or without a statistical editor or reviewer) should be aware when reading a manuscript describing a (novel) questionnaire.

Questionnaire construction and validation

The purpose of a questionnaire (or any other instrument) is to measure a phenomenon with the least possible bias or error. To fulfil that requirement, the questionnaire must be developed following several methodological steps.¹

Construction begins with the literature search of the research question within a theoretical framework. A questionnaire should be constructed to test a theory - general or specific. For example, to test health-related behaviour, one can use Theory of reasoned actions (general) or Theory of health-related behaviour (specific).¹ After reading and reviewing the existing body of evidence, the research team brainstorms and devises the questions² then sometimes uses focus groups or a Delphi panel method to refine the questions, perhaps produce more, and to gain more insight into the problem.

All questions (items) must be clearly formulated, as specific as possible, and unambiguous.^{2,3} For example in the Croatian version of the Chronic Venous Insufficiency Quality of Life Questionnaire (CIVIQ),⁴ the item: “Feeling nervous/tense” is a double-barrelled question and can be problematic because in the analysis of the answer one does not know whether the respondent was nervous, tense or nervous and tense.⁴ If the questionnaire is measuring knowledge or attitudes, there should be a mix of the same number of positive and negative statements. The order of items in the questionnaire has to be considered carefully, to reduce item order effect.^{2,3} The timeframe is also very important. For example, the question: “Did you use iron supplements?” should be formulated as: “Did you use iron supplements in the last six months”. Five-point rating scales are widely used and called Likert-type, but the true Likert scale is for the

measurement of attitudes, ie how much the participant agrees or disagrees with a particular statement.² Verbal response options are given instead of numerical values.^{2,3} The length of a questionnaire depends on the number of constructs to be measured. In my experience, it is preferable that each scale has at least 10–15 items in the first version.

This phase of research results in a large questionnaire that is usually tested in the second, piloting phase.

Pilot testing of the questionnaire is usually done with a small number of participants, who might not be representative of the target population, for example a convenience or snowball sample. The researcher then analyses the data and refines the questionnaire, ie removes items that do not measure the exact phenomena (validity testing, for example factor analysis or structure equation modelling) or do not measure it reliably (reliability testing, for example internal consistency or split half reliability). The revised questionnaire is then used in the target sample and a control group, preferably together with other already validated instruments that measure the same and different phenomena.⁵

Validation in an independent population is very important, especially in biomedicine. If data are collected, the sample cannot be split in parts, so that one part is used for the development of a scale, and the other for the validation. Another cohort must be chosen and used for the validation.

Sample size and response rate

The number of the respondents should be 5-10 times more than the number of items in the scale.⁶ For example: a manuscript that describes evaluation of a questionnaire consisting of 20 items should include data from at least 100 respondents, preferably more.

The sample of respondents must be representative of the whole population that is being investigated, eg in age, gender, education level, illness characteristics. Sample size calculators can be used to determine how many respondents are needed (for example: <https://www.surveymonkey.com/mp/sample-size-calculator/>), but if the sample is not representative, however large the sample, the data cannot be generalised to the population.

Response rate is also very important. For example *American Journal of Pharmaceutical Education* has set a response rate of 60% or higher in their journal,⁷ but a lot of published studies have a lower response rate. In order to represent the chosen sample and to reduce the nonresponse bias, the response rate has to be as high as possible.⁷ I have read several articles with very low response rate (such as 20%) and do not approve accepting manuscripts with very low response rate even if the sample is relatively big but there could be reasons for non-responding (sensitive topics and delicate questions) and it is up to the editor to decide whether he/she will send such a manuscript to peer review.

Description of the questionnaire

The full evaluation process should be described in detail in the Methods section such that it could be replicated. The results of the pilot and validation studies should be presented in full, with validity (factor analysis) and reliability analysis (Cronbach alpha coefficient). Scree plots, eigen values and reliability coefficients are all required for peer reviewers and readers to decide whether the questionnaire was developed properly.

If possible, the questionnaire should be published in the article or as a supplement. Clear instructions about the scale scoring should be given. Manuscripts presenting non-standardised questionnaires should be rejected.

Data analysis

The exact number of responses to each item should be presented.⁸ The responses should also be analysed for any variables that might affect the responses, such as gender,⁹ income level, urban vs rural living.¹⁰

To interpret results that have been measured using different scales, results are usually transformed to z-values or linearly combined. For example, CIVIQ questionnaire consists of 14 items and has 3 scales: pain (3 items), physical (5 items) and psychological (6 items) dimension with different number of items per scale. In order to compare results on each scale a linear combination is used: $([\text{Final score} - \text{minimal possible score}] / [\text{Theoretical maximal} - \text{minimal score}]) \times 100$. The result can be converted to a 100 point scale (<https://www.civiq-20.com/scoring-missing-data/calculation-global-index-score/>) and the comparison of results on different scales was possible.⁴

Questionnaires measure construct(s) indirectly and these scales are continuous variables. The results usually present the relation between several variables, using different models of regression analyses. For multiple regression, variables have to be normally distributed and correlated.⁸ Data that do not follow a normal distribution (Kolmogorov-Smirnoff test) cannot be used in multiple regression analysis.¹¹

Many people think that it is very easy to conduct and publish a survey: they just list some questions and send them to some respondents. In fact, creating a valid, standardised questionnaire is a complex endeavour that usually requires a team of experts and considerable investment of time and resources. Every manuscript describing a new questionnaire should present a clearly defined objective and rationale for the study, ie why there is a need for a (new) questionnaire construction or attitudes measurement, within a theoretical framework.^{1,3} A statistical peer reviewer can be consulted in order to help with the review of the questionnaire construction and validation.

Among 485 guidelines on the EQUATOR network (<https://www.equator-network.org/>), four address questionnaires and surveys, the latest being published in 2009, and one includes a checklist for web surveys (CHERRIES). Editors could give more advice on how to report questionnaires in their journal instructions for authors and maybe develop guidelines and checklists for reporting standardised questionnaires.¹²

Ksenija Baždarić

References

- 1 Sutton S. Determinants of Health-Related Behaviours: Theoretical and Methodological Issues. 2004;94-126. doi:10.4135/9781848608153.n4. Available at: <https://www.phpc.cam.ac.uk/pcu/wp-content/files/2012/01/Ch-04.pdf>
- 2 Price Paul C, Jhangiani Rajiv CI. Constructing Survey Questionnaires. In: *Research Methods in Psychology*. 2nd ed. 2015;Ch. 9. <https://opentextbc.ca/researchmethods/chapter/constructing-survey-questionnaires/>.
- 3 Lee SH. Constructing Effective Questionnaires. In: *Handbook of Human Performance Technology*; 2001:760-779. <http://www.davidlewisphd.com/courses/EDD8006/fall11/2006-Lee.pdf>.
- 4 Sinožić T, Baždarić K, Šverko D, Ružić A, Katić M. Validation of the Croatian version of CIVIQ quality of life questionnaire in patients with chronic venous disorders. *Croatian Medical Journal*. 2017;58(4). doi:10.3325/cmj.2017.58.292
- 5 Price Paul C, Jhangiani Rajiv CI. Reliability and Validity of Measurement Title. In: *Research Methods in Psychology*. 2nd ed; 2015. <https://opentextbc.ca/researchmethods/chapter/reliability-and-validity-of-measurement/>.
- 6 Costello AB, Osbourne JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*. 2005;10(7):1-9. doi:10.1.1.110.9154
- 7 Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *American Journal of Pharmaceutical Education*. 2008;72(2):43. doi:10.5688/aj720243
- 8 Lang T. Twenty statistical errors even you can find in biomedical research articles. *Croatian Medical Journal*. 2004;45(4):361-370.
- 9 Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016. doi:DOI 10.1186/s41073-016-0007-6
- 10 Wark S. A missing factor in the reporting of medical research outcomes: Geographic classification of participants. *European Science Editing*. 2017;43(4):76-80. doi:10.20316/ESE.2017.43.016
- 11 Lang, TA; Secic M. *How to Report Statistics in Medicine*. 2nd ed. American College of Physicians; 2006.
- 12 Marsh J. Should editors be more involved in the development of reporting guidelines? *European Science Editing*. 2018;44(1):2-3. doi:10.20316/ESE.2018.44.17024