

FEATURES SECTION

How to ...

Write and analyse a questionnaire

A. Williams

University of Bristol Dental School, Bristol, UK

Abstract

Index words:

Questionnaire, response rates, study design.

Questionnaires are used in a wide range of settings to gather information about the opinions and behaviour of individuals. As with any other branch of science, the validity and reliability of the measurement tool, i.e. the questionnaire, needs to be rigorously tested to ensure that the data collected is meaningful. The design and method of administration of a questionnaire will also influence the response rate that is achieved and the quality of data that is collected. This paper discusses the issues that should be considered when designing and undertaking a questionnaire study.

Introduction

Questionnaires can be used in a wide range of settings to gather information about the opinions and behaviour of consumers. The Griffiths Report (Working for Patients)¹ popularized the use of patient satisfaction questionnaires in the National Health Service (NHS) during the 1980s. These questionnaires were written from the perspective of the Providers, however, and the information that was collected tended to be ambiguous and misleading.² More recently, there has been a new emphasis on patient involvement in health care delivery in the NHS.³ Patient-based outcome measures, in the form of questionnaires, are increasingly being used in randomized controlled trials to evaluate new treatments and also as part of audit. Questionnaires are also used to evaluate participant opinion of courses as part of the Continuing Professional Development (CPD) process.

As with other branches of science, a rigorous approach needs to be taken to the design and execution of questionnaire studies. There are many examples in the literature of studies based on questionnaires in which there has been no attempt to test the validity or reliability of the measurement tool. This undermines the results of the study. Achieving an adequate and representative sample is another important issue. Questionnaire surveys are particularly reliant on the willingness of the subjects to take part. Considerable effort is therefore required from the outset to ensure that the questionnaire is acceptable to the target population to maximize response rates. This paper describes the steps that should be followed when

undertaking a questionnaire study and discusses methods of improving response rate.

The stages that you should follow when designing your questionnaire are summarized in Figure 1.

Define your research question and study population

It is important to define your research question, study population, and the objectives of your study at the beginning of your study. You should continually refer back to these during the study design process. This is particularly important in questionnaire studies where there is a temptation to be 'nosey' and delve into a wide range of issues, which although interesting to you, the researcher, are not relevant to your study. You should also consider the associations that you wish to test at the design stage, so that the relevant data, for example, social class indicators, can be collected as part of your questionnaire.

Previous studies⁴ have shown that people are more likely to respond to questionnaires that cover issues that are relevant to them. Qualitative methods, including focus groups and unstructured interviews, are increasingly being used to identify issues of importance to patients as a first stage in questionnaire studies. The selection of participants for qualitative research is quite different from quantitative studies. Rather than selecting subjects at random from a representative group, participants are chosen to take part in qualitative studies on the basis of their particular demographic or treatment char-

1. Define your research question and study population.
2. Decide how the questionnaire will be administered.
3. Formulate your questions.
4. Formulate the responses.
5. Design the layout.
6. Pre-pilot the questions and layout.
7. Pilot study—test validity, reliability, acceptability.
8. Design your coding scheme.
9. Print questionnaire.

Fig. 1 Stages of questionnaire design.

acteristics, or because they are known to hold particular views. This technique is called ‘purposive sampling’ and is useful because a range of possible views can be identified. It is important to appreciate, however, that data collected in this way, although interesting, is not representative of the whole study population.

Focus group meetings are a useful way of identifying issues because the views of a range of subjects can be examined at the same time. Furthermore, interaction between participants can lead to new issues being identified. Running a focus group is a special skill, however, and you will need to use a trained facilitator to ensure that the views of all the participants are included. Unstructured interviews are another method that can be used to identify issues. These are particularly valuable for examining people’s attitudes and beliefs in-depth, and can give insight into some of the reasons behind their behaviour. Again, you will need to employ a trained interviewer for this part of your study to avoid biasing your data.

You should not under-estimate the amount of work and time involved in collecting and analysing qualitative data.⁵ The interviews and focus groups need to be tape-recorded (with the participants’ permission) and you should also take careful written notes to back up the tapes. The tape recordings are then transcribed into scripts with the participants identified by code names only to ensure confidentiality. This can be quite a difficult process for tapes of focus group meetings, since a number of participants may be talking all at once! There are a number of techniques for analysing qualitative data that are beyond the scope of this paper. Put simply, the scripts should be examined by at least two researchers who record new issues as they arise and then produce a list of their relative importance. The main issues identified from the interviews and/or focus groups can then be used to form the basis of your questionnaire. The subjective nature of qualitative data means that it is very easy to bias

your analysis and so it is important to seek advice from an experienced researcher at this stage.

Decide how the questionnaire will be administered

Questionnaires can be used either as the basis of a structured interview, which is administered by a trained interviewer, or completed by the subject by themselves. When you decide how a questionnaire should be administered you need to achieve a balance between practical considerations, such as the time-frame and funding available for the study, and the issues you wish to examine.

Structured interviews can be undertaken face-to-face, or may be conducted over the telephone or the Internet. Interviewer-administered questionnaires have the advantage that unclear questions can be clarified to the respondent and open-ended questions can be used to collect a range of possible responses. Importantly, the interviewer can also ensure that all the questions are answered by the intended subject. These surveys are, however, expensive because trained interviewers are needed and a large amount of time is needed for each interview.

There are some disadvantages. For example, there is a risk that the interviewer may bias the responses given. Response rates may also be reduced because some people may be unwilling to give up their time to be interviewed. Others may not take part because they prefer the anonymity of a self-complete questionnaire. Finally, response rates to telephone surveys are particularly low because subjects tend to assume that an unfamiliar voice is trying to sell them double-glazing and they therefore hang up before you have had a chance to explain the purpose of your call!

When we consider self-complete questionnaires, these are cheaper to administer and a larger sample can be collected. Respondents can either complete the questionnaire in the research setting or, more commonly, are sent the questionnaire by post or e-mail to complete at home. The alternative option of inviting subjects to complete a questionnaire in the clinic, for example, has the advantage that you can ensure that your target subject completes the questionnaire. You can also clarify any ambiguous questions and make sure that the respondent answers all the questions. These advantages need to be offset against the difficulties of taking up people’s time (your sample will be biased towards people with time to waste!) and disrupting busy clinics. A researcher

also needs to be present throughout the data collection process, which increases the cost of the study and limits the geographical catchment area of the sample that can be studied. There is also a tendency for subjects to feel intimidated by the clinical setting and, as such, they will tend to give responses to please the clinicians and/or researchers, rather than their own opinions.

Most questionnaire studies are based on postal self-complete questionnaires. This study design has the advantage that a large population can be sampled at a relatively low cost over a wide geographical area. The main disadvantages of postal questionnaires are that you have no control over who actually completes the questionnaire. The responses given are often the collective view of a whole household. It is also difficult to ensure that the respondent completes all the questions before returning the questionnaire and there may be problems with literacy or language that are difficult to identify. Good response rates are therefore difficult to achieve with postal surveys and samples tends to be biased towards more educated and non-immigrant populations.⁶

Formulate your questions

Once you've decided how you are going to administer your questionnaire, you can go on to formulate your questions. Questions can be divided into those directly related to the research question; filter questions that explore the characteristics of the different study groups and 'filler' questions that, although not part of the research question, aid the flow of the questionnaire. Wherever possible, you should incorporate questions from existing questionnaires (with the permission of the author). This not only prevents 'reinventing the wheel', but will also allow direct comparison of your data with previous studies. It is important to appreciate, however, that such questions need to be reproduced verbatim to be valid. If you have to change the wording of a question, for example, to suit a British sample, rather than an American one, then, strictly speaking, you will need to retest the reliability and validity of the question.

Studies⁷ have shown that the wording of questions has an important influence on the responses that are given. The principles of question wording are summarized in Figure 2. In general, questions should be short (less than 20 words is recommended), simple and specific. It is important to remember that the average reading age in the UK is around 12 years. More difficult questions will either produce an inaccurate response or, more likely, the respondent will give up and fail to complete the

1. Use simple language.
2. Avoid jargon.
3. Keep questions short and specific.
4. Avoid ambiguities.
5. Avoid double-barrelled questions (those with 'and' or 'or' in the wording).
6. Avoid double negatives.
7. Avoid loaded words.
8. Avoid leading questions.
9. Do not overtax the respondent's memory.
10. Avoid hypothetical questions.

Fig. 2 Principles of question wording.⁸

questionnaire at all. There is a facility in Microsoft Word[®] that you can use to calculate the readability of your questions. Alternatively, you can use the Gunning Fog Index to calculate the reading age for each question.

It is also important that you avoid making assumptions in your questions. For example, 'How often do you visit the dentist?' assumes that the respondent attends the dentist regularly. You should break this question down into two parts, i.e. 'Do you visit the dentist for checkups?' and 'If yes, how often do you go for a check-up?'

Another potential problem is time-memory bias. People find it particularly difficult to recall events that have happened some time ago and so it is unwise to ask about things that have happened more than 6 months in the past. If this is unavoidable then you should either use wide time scales, for example 6–12 months, or put the event in question into some sort of context. For example, the question 'When did you first meet the surgeon?', in a survey of orthognathic patients, could be rewritten 'Did you meet the surgeon before your braces were fitted?'

The order of your questions is another important issue in questionnaire design. Studies have shown that a better response rate is achieved if general questions precede specific questions.⁸ It is better to start with easy, factual non-personal questions to relax the respondent. Questions that involve some research from the respondent, for example, looking up clinic dates, should come towards the end of the questionnaire. This has the advantage that your respondent will have hopefully developed an interest in the subject and feel some ownership of the questionnaire by this stage. If not, then at least they will have completed the rest of the questions and you will be able to collect some data. It is also better to ask personal questions towards the end of the questionnaire when again the respondent is more relaxed.

pre-pilot is not a formal procedure, more an information-gathering exercise. You should sit down with a few suitable subjects, who may be friends or colleagues, and go through the questions together to identify potential problems. After each session, you should amend the questionnaire before re-piloting with another group of testers. This process needs to continue until you are confident that your questions are unambiguous, appropriate and acceptable to respondents. You can also test the layout at this stage to ensure that people can navigate their way easily through the questionnaire.

Pilot study to test the questionnaire

As with other forms of science, you need to be able to show that the data collected from your questionnaire are valid and reliable. Ideally, every questionnaire should undergo a formal pilot during which the acceptability, validity, and reliability of the measure is tested. You should also pilot the data collection process and covering letters to participants. The pilot should be based on subjects from a similar population to that being examined in your survey. Since patients will be involved, you will need to obtain ethical approval for this part of your study. Testing a questionnaire can be a very time-consuming process and this stage of questionnaire design has often been overlooked in the past in researchers' eagerness to start collecting the data. Unfortunately, this then casts doubt over the robustness of the data that are collected.

Testing validity

A questionnaire can be said to be 'valid' if it examines the full scope of the research question in a balanced way, i.e. it measures what it aims to measure.⁸ There are several aspects of validity that need to be tested. Criterion validity is assessed by comparing a new measure with an existing 'gold standard' scale. If such a scale exists, however, one would question the need to develop a new questionnaire. The factual validity of a questionnaire can be assessed by comparing responses about clinical events with information recorded from the clinical notes. The face validity of a questionnaire can be examined by interviewing people, either face-to-face or over the telephone, after they have completed the questionnaire to find out whether the responses they have given in the questionnaire agree with their real opinions. When testing face validity, it's important to word your questions in the interview differently from those in the questionnaire

otherwise all you will be testing is the reliability of your questions.

Testing reliability

Reliability is defined as an assessment of the reproducibility and consistency of an instrument. For self-complete questionnaires, two aspects of reliability should be examined. You can assess test-retest reliability by asking people to complete the questionnaire on two separate occasions approximately 2 to 3 weeks apart, assuming that their circumstances will not have changed in the interim. The two sets of responses can then be compared statistically using weighted Kappa¹² for categorical data and Spearman's Rank Correlation Coefficient (or its non-parametric equivalent, the Wilcoxon Rank Sum test) for continuous data. You can determine the internal consistency of your questionnaire by asking a question or questions in more than one way during the questionnaire. The responses given can then be compared as before.

Testing acceptability

Qualitative methods can be used to assess the acceptability of a questionnaire. You can either ask the subjects included in your pilot study to write their comments about the questionnaire on a separate sheet or you can ask them over the telephone how they found answering the questionnaire during the validity testing. It's also a good idea to ask people in the pilot study how long it took them to complete the questionnaire. You can then include this information in the cover letter that you use to accompany the questionnaire in your main survey.

Design your coding scheme

Coding is the process of converting questionnaire data into meaningful categories to facilitate analysis. You need to think about your coding scheme at the beginning of your study and, wherever possible, build it into your questionnaire, for example by numbering the response tick boxes for each question. This will allow you to enter data directly from the questionnaire into your database for analysis. The numbers within the boxes should correspond to the variables in the database where the responses will be stored. The alternative is to code the questionnaire responses onto a separate coding sheet and then to enter the data from the coding sheet into the database. This process is not only laborious, but also

doubles the margin for error and observer bias. It's a good idea to test your coding scheme and data entry process during the pilot study, so that problems can be rectified before the main study starts. It's also prudent to discuss the coding scheme for your questionnaire with a statistician at this stage so that any mistakes can be rectified before you go to the printers.

Increasing your response rate

Achieving a good response rate is a particular problem for postal questionnaire surveys. Studies¹³ suggest that the response rate may be reduced by up to 20% compared to an interviewer-based survey. A response rate of 75% for a postal questionnaire survey is considered to be extremely good. Poor response rates can lead to bias because people from higher social groups and non-ethnic minorities are more likely to complete questionnaires than other groups.⁶

It's unethical to offer people financial incentives to take part in healthcare surveys and so researchers need to develop other ways of encouraging subjects to take part. Figure 3 lists the methods that are commonly used to increase response rates in postal questionnaire studies. Everyone is bombarded by junk mail these days, much of which ends up in the dustbin unopened. Therefore, it's a good idea to use white stamped envelopes, which look more personal, to send out initial contact letters, rather than brown franked envelopes. Most ethics committees insist that you send subjects information in advance about a study and that you ask them to indicate whether they wish to take part in a survey, rather than simply sending them a questionnaire through the post with a cover letter asking them to complete it. This stage of a questionnaire survey not only increases the cost of your study, but also tends to reduce response rates.

All the letters that you send to potential participants need to be approved by the ethics committee before you begin your study. Each letter should state the aim of the study, the sponsor and emphasize that the information given will be confidential. It's also important that you state in your initial contact letter that if a subject chooses not to participate in the study then their future healthcare will not be affected. Postal questionnaires should be sent out with a pre-paid envelope for their return. All contact letters should be written on official headed paper and signed in blue ink by the researcher.

It's expensive to employ interpreters solely for questionnaire surveys. One way around this, if you expect a significant proportion of your sample to come from a

1. Use white, stamped envelopes to contact subjects.
2. Always include a pre-paid addressed envelope for respondent's replies.
3. Include a personal cover letter that has been signed in blue ink.
4. Use official-looking headed paper for all correspondence.
5. Stress the anonymity of the survey.
6. Focus the questionnaire on issues of importance to the subjects.
7. Design a questionnaire that is easy to navigate.
8. Use brightly coloured paper and add a logo to the cover of the questionnaire.
9. Send at least two reminders to non-respondents.
10. Include another copy of the questionnaire and a pre-paid envelope with each reminder.

Fig. 3 Recommendations for increasing response rates.

non-English speaking group, is to include a cover letter that has been translated into common languages for your target population. A contact number for an interpreter who people can telephone to help them complete the questionnaire should be included in the letter. It's difficult to control for bias if a subject asks an interpreter to help them complete the questionnaire, but the advantages of increasing the response rate by providing an interpreter usually outweigh the disadvantages.

Sending out reminders has also been shown to improve response rates in questionnaire surveys. Each questionnaire should be labelled with a unique ID number so that reminders are only sent to non-respondents. It's quite possible that a subject may have lost their original questionnaire and so every reminder should be accompanied by another questionnaire and another pre-paid envelope. It's a good idea to label the reminder questionnaires with the ID number and prefix, for example, 'a', 'b' or 'i', 'ii', etc. This will allow you to identify those instances where the first questionnaire that was completed has been delayed or lost in the post, and the subject has then been kind enough to complete a repeat questionnaire for you.

It's important to remember that people have given up their time to complete your questionnaire and so it's only considerate to send each participant a letter thanking them for taking part in your study. Recent Government Standards for Research¹⁴ state that the results of studies involving NHS patients should be fed back to participants. Some of the information you've collected could be distressing to some patients, however, and you should therefore seek advice about which information should be disseminated.

Data coding

It is good practice to check through each completed questionnaire when you receive it to check for missing data. If there are any factual gaps in people's responses, for example, the number of visits made to a clinic, then it may then be possible to enter these data from other sources. Subjects who have failed to follow the filters and answered every single question can also be identified at this stage, and if necessary, contacted to verify information.

Two approaches can be used when analysing the responses to open-ended questions. You may choose to examine all the responses to a question, devise categories for the answers and then code the data in the same way as a closed response question. Alternatively, interesting responses can be quoted verbatim in the final report. The choice depends on the nature of the question and the range of answers given.

Once you have coded as many of the open-ended questions as possible you can enter your data directly from the completed questionnaires into a computer programme for analysis. The Statistical Package for Social Sciences (SPSS)¹⁵ is a useful software package for questionnaire surveys because it is flexible and easy to use. For large surveys there are commercial data entry firms who will undertake this process for you at a price of around £2 per questionnaire. Optical scanners are often used to enter data from large surveys. These are only suitable for single coded responses however, and there is a high error rate.

Data cleaning

Once you've entered all your questionnaire responses into the database it is necessary to spend some time cleaning the data to identify inconsistencies and outliers. This is a laborious process and there is a strong temptation to leave this stage out, particularly if there is a scientific conference looming at which you've promised to present the results of your study! This is a false economy, however, since you may then have the embarrassment of having to amend your abstract or even withdraw your paper when you discover that your 'exciting' finding is merely the result of a data entry error!

The quickest way to clean your data is to produce frequency figures for each question and examine the outliers. Simple cross-tabulations can then be used to identify nonsensical responses, for example, patients who have undergone a mandibular setback for a Skeletal II discrepancy.

Data analysis

It's particularly important to refer back to the original aims of your study and the hypotheses that you wish to test at this stage to keep your analysis focused.

The usual sequence is:

1. Descriptive analysis: describe the distribution and range of responses to each variable and examine the data for skewness.
2. Recode data into categories where appropriate, for example, ages into age ranges, to enable statistically meaningful comparison of sub-groups.
3. Bivariate analyses: use simple cross-tabulations to identify trends and examine possible associations between one variable and another.
4. Multivariate analyses/regression analysis techniques can then be used to test the effect of effect of one variable on an outcome, whilst controlling for another.

Examination of the data for selection bias

It is difficult to avoid the tendency for postal questionnaire surveys, in particular, to be biased towards educated, English-speaking populations. The extent to which this applies to your survey should be stated in your report to enable the reader to judge for themselves how valid your results are for a given population. It's also important to record as many demographic and treatment details as possible about non-respondents. Simple cross-tabulations can then be used to compare the characteristics of respondents and non-respondents. If you have undertaken an interviewer-based survey involving more than one interviewer, then it will be necessary to compare the responses given to each interviewer to identify observer bias.

Conclusion

In this paper I have described some of the important principles of undertaking a questionnaire survey. The steps needed for designing and testing a questionnaire have been discussed, together with issues surrounding the coding and analysis of the data. It should be clear to the reader by now that developing a questionnaire from scratch is a very time-consuming process. Wherever possible existing measures should be used with the permission of the author.

Recommended reading

Bowling A. *Research Methods in Health*. Buckingham: Open University Press, 1997.

Black N, Brazier J, Fitzpatrick R, Reeves B. Patient and staff questionnaires. In: *Health Services Research Methods—a guide to best practice*. London: BMJ Books, 1998.

References

1. Griffiths R. *NHS Management Inquiry*, the Griffiths Report. London: Department of Health and Social Security, 1983.
2. Hall JA, Dornan MC. What patients like about their medical care and how often they are asked: a meta-analysis of the satisfaction literature. *Soc Sci Med* 1988; **27**: 935–939.
3. Department of Health. *The NHS Plan. a Plan for Investment. a Plan for Reform*. Norwich: Stationery Office, 2000.
4. Heberlein TA, Baumgartner R. Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature. *Am Soc Rev* 1978; **43**: 447–462.
5. Bowling A. Unstructured interviewing and focus groups. In: *Research Methods in Health*. Buckingham: Open University Press, 1997.
6. Cartwright A. Some experiments with factors that might affect the response of mothers to a postal questionnaire. *Stat Med* 1986; **5**: 607–617.
7. Larsen JD, Mascharka C, Toronski C. Does the wording of the question change the number of headaches people report on a health questionnaire? *Psychological Record* 1987; **3**: 423–427.
8. Black N, Brazier J, Fitzpatrick R, Reeves B. Designing and using patient and staff questionnaires. In *Health Services Research Methods—a guide to best practice*. London: BMJ Books, 1998.
9. Bowling A. Questionnaire design. In: *Research Methods in Health*. Buckingham: Open University Press, 1997.
10. Hawkins DI, Coney KA. Uninformed response error in survey research. *J Market Res* 1981; **18**: 370–374.
11. Poe GS, Seeman L, McClaughlin J, Mehl E. ‘Don’t know’ boxes in factual questions in a mail questionnaire: effects on level and quality of response. *Public Opinion Quarterly* 1988; **52**: 212–222.
12. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991
13. Cartwright A. Interviews or postal questionnaires? Comparisons of data about women’s experiences with maternity services. *Milbank Q* 1988; **66**: 172–189.
14. Department of Health. *Research Governance Framework for Health and Social Care*. London: DoH, 2000 Available at: www.doh.gov.uk/research/rd1/researchgovernance/researchgovindex.htm (accessed July 2003)
15. SPSS (1999). *SPSS for Windows*. Release 10.0. : SPSS.