

Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients

E McColl^{1*}
A Jacoby¹
L Thomas¹
J Soutter¹
C Bamford¹
N Steen¹
R Thomas²
E Harvey³
A Garratt³
J Bond¹

¹ Centre for Health Services Research, University of Newcastle, UK

² National Centre for Social Research, London, UK

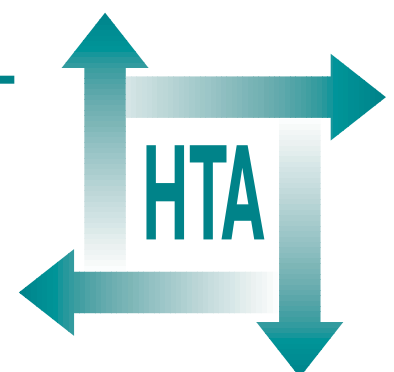
³ Department of Health Sciences, University of York, UK

* Corresponding author

Executive summary

Health Technology Assessment 2001; Vol. 5: No. 31

Health Technology Assessment
NHS R&D HTA Programme





Executive summary

Introduction

Questionnaires are often used to collect primary quantitative data from patients and healthcare professionals. The aim is to gather valid, reliable, unbiased and discriminatory data from a representative sample of respondents. However, the information yielded is subject to error and bias from a range of sources. Close attention to issues of questionnaire design and survey administration can reduce these errors.

Objectives

A selective, narrative literature review was conducted to identify current best practice with respect to the design and conduct of questionnaire surveys, including theories of respondent behaviour, “expert opinion” and high-quality evidence from experimental studies. The principal foci were:

- modes of survey administration (various forms of interviewer administration and self-completion)
- question wording, choice of response formats, and question sequencing
- questionnaire formatting and other aspects of presentation
- techniques for enhancing response rates, with particular emphasis on postal surveys.

Methods of the review and implications for readers

The starting point for this review was “expert opinion”, encapsulated in key textbooks on the design and conduct of surveys. High-grade evidence was then sought from experimental and quasi-experimental studies to support or refute the experts’ recommendations. In addition, information was sought on the theoretical underpinnings of survey response. A deliberate and considered decision was made to include studies from disciplines other than health because it was envisaged that theories of respondent behaviour, as well as methodological messages, are unlikely to be discipline specific. The PsycLIT electronic database was therefore used in addition to MEDLINE, but the

search was confined to articles published in the English language between 1975 and 1996. It is acknowledged that confining the search to two databases only is likely to have led to bias in favour of articles published in the major American and British journals indexed on those databases, and to exclusion of the “grey” literature.

Owing to human error, references identified through MEDLINE for the period 1987–1992, and those identified through PsycLIT for 1979, 1991 and 1993–1996, were excluded from consideration (appendix 5). Although it is acknowledged that these omissions mean that this review cannot be considered to be systematic, the authors believe that their conclusions would not have been materially altered by the incorporation of articles identified through the two key databases for the years in question. In contrast to clinical research, where the accretion of knowledge tends to be incremental, with new studies seeking to replicate or refute the findings of those that have gone before, research into questionnaire construction and survey administration appears to be haphazard, often with little reference to previous studies. Examination of the literature provided little sense of concerted efforts to generalise findings from one study to other settings, populations or modes of administration.

Explicit inclusion, exclusion and quality criteria were applied in a two-stage process of sifting and then synthesising findings from identified studies. However, because of the heterogeneity of studies, no attempts at meta-analysis were made. To facilitate comparisons between studies, findings are presented as relative risks with associated 95% confidence intervals (for differences in percentages), or as differences in means with associated 95% confidence intervals (for continuous data). In setting out the findings, a distinction has also been made between studies on health-related topics and those from other fields. A quality score is included for each identified study.

In defining the scope of this review, an explicit decision was made to exclude certain aspects of the survey process, most notably sampling and pilot testing. These features of survey methodology do not lend themselves readily to experimental

investigation and they are likely to be highly study and topic specific. Indeed, in the context of health technology assessment, definitions of sample inclusion and exclusion criteria and sample size calculations are likely to be predicated on the design of the parent trial. However, in recognition of the importance of these aspects of the survey process, appendix 1 provides brief guidance on key topics omitted from the formal review. This appendix is based primarily on the collective experience of the authors, with limited references to key texts and articles on the chosen topics.

Results

Mode of administration

The two principal modes of administration are self-completion and interviewer administration. Evidence from identified studies provided no consistent picture of the superiority of any one mode in terms of the quantity or quality of the response, or the resources required.

Question wording and sequencing

Evidence from identified studies supported the notion that question wording and framing, including the choice and order of response categories, can have an important impact on the nature and quality of responses.

The conventional wisdom with respect to question ordering is that general questions should precede specific questions; evidence from a number of primary studies supported this assertion.

Questionnaire appearance

Through careful attention to the design and layout of questionnaires, the risk of errors in posing and interpreting questions and in recording and coding responses can be reduced, and potential inter-rater variability can be minimised.

Evidence from experimental and quasi-experimental studies on aspects of questionnaire appearance was scanty. However, a number of articles were identified that outlined a theoretical basis to aspects of design, which suggested that questionnaire appearance can influence respondents' decisions at several stages, from arousal of interest in questionnaire completion, through task evaluation, to initiation and monitoring of the process of completion. There is a need for consistency in the presentation of visual information and an understanding and application of "graphic non-verbal language" (i.e. the spatial arrangement of information and other visual phenomena such as colour and brightness).

Enhancing response rates

High survey response rates are desirable because they increase the precision of parameter estimates and reduce the risk of non-response bias.

Many factors may combine to influence the decision of a recipient of a questionnaire to respond. Potential respondents must have both the means to complete the questionnaire and the will to do so; the perceived costs of responding must not exceed the benefits.

"Saliency" – the apparent relevance, importance and interest of the survey to the respondent – is a very important influence on response rates. Fortunately, health-related surveys are likely to be perceived as salient. Perhaps surprisingly, questionnaire length appears to be less important.

The number of contacts made with sampled individuals is another powerful factor. Some researchers advocate prenotification, so that recipients are primed for the arrival of the questionnaire. Almost all experts in survey design advocate the use of reminders, a recommendation supported by evidence from primary studies.

Other factors that have been shown to influence response rates include making a self-interest/utility appeal to the respondent and the use of incentives (particularly enclosed monetary incentives). Perhaps surprisingly, anonymity has not been demonstrated to have any consistent effects on the rate or quality of response.

Conclusions

Recommendations for practice

The heterogeneity of findings indicates that there can be no universal recommendations on best practice in respect of questionnaire design and survey conduct. Rather, individual survey researchers need to take into account the aims of the particular study, the population under investigation and the resources available; trade-offs between the ideal and the possible are likely to be needed. However, some general principles can be offered.

The principal objective should always be to collect reliable, valid and unbiased data from a representative sample, in a timely manner and within given resource constraints.

In choosing a mode of questionnaire administration, consideration needs to be given to the availability ►

of an appropriate sampling frame, anticipated response rates, the potential for bias from sources other than non-response, acceptability to the target population, the time available, the financial budget, and the availability of other resources (e.g. skills or equipment).

In formulating questions and response categories, and in determining question order, researchers should bear in mind that survey respondents employ a wide range of cognitive processes in formulating their responses. To minimise bias and to reduce spurious inter-respondent variation, careful attention must be given to these issues.

The “task analysis” model, the theory of social exchange and theories of perception and cognition should inform decisions regarding the physical design of questionnaires, as well as strategies for delivering and returning them. The aim should be to enhance the perceived and actual benefits of responding and to minimise the perceived and real costs. The effort required to interpret questions and provide responses should be made as easy as possible. Strategies for reducing the monetary cost to respondents include the use of prepaid return envelopes and the provision of financial incentives (unless ethical imperatives preclude the latter).

Recommendations for research

Both quantitative research (in the form of experimental manipulations of various aspects

of questionnaire design and administration) and qualitative research (in the form of cognitive interviews addressing the processes by which respondents react to questionnaire stimuli) are required.

Assessing the reproducibility of previous findings should be afforded higher priority than embarking on totally new lines of enquiry. In particular, it will be important to investigate whether findings from social, educational or market research also apply to health-related surveys. It will also be important to test whether observed effects of manipulating different aspects of questionnaire design are equally applicable to interviewer-administered and self-completed questionnaires.

Multiple measures of outcome or “success” should be examined, including those of quantity (e.g. questionnaire and item response rates) and quality (e.g. non-response bias; and validity, reliability and distribution of responses), as well as resource implications.

Publication

McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, *et al.* Design and use of questionnaires: a review of best practice applicable to surveys of health staff and patients. *Health Technol Assess* 2001;5(31)

How to obtain copies of this and other HTA reports

Copies of this report can be obtained by writing to:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.

Or by faxing us at: +44 (0) 23 8059 5639

Or by emailing us at: hta@soton.ac.uk

Or by ordering from our website: <http://www.ncchta.org>

NHSnet: <http://nww.hta.nhsweb.nhs.uk>

The website also provides information about the HTA Programme and lists the membership of the various committees.

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 93/49/02.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Methodology Programme Director: Professor Richard Lilford
HTA Programme Director: Professor Kent Woods
Series Editors: Professor Andrew Stevens, Dr Ken Stein, Professor John Gabbay
and Dr Ruairidh Milne
Monograph Editorial Manager: Melanie Corris

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2001

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to HMSO, The Copyright Unit, St Clements House, 2-16 Colegate, Norwich, NR3 1BQ.

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.