



Self report in clinical and epidemiological studies with non-English speakers: the challenge of language and culture

Sanja M Hunt and Raj Bhopal

J. Epidemiol. Community Health 2004;58:618-622
doi:10.1136/jech.2003.010074

Updated information and services can be found at:
<http://jch.bmjournals.com/cgi/content/full/58/7/618>

These include:

- | | |
|-------------------------------|---|
| References | This article cites 16 articles, 4 of which can be accessed free at:
http://jch.bmjournals.com/cgi/content/full/58/7/618#BIBL

1 online articles that cite this article can be accessed at:
http://jch.bmjournals.com/cgi/content/full/58/7/618#otherarticles |
| Rapid responses | You can respond to this article at:
http://jch.bmjournals.com/cgi/eletter-submit/58/7/618 |
| Email alerting service | Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article |
-
- | | |
|--------------------------|--|
| Topic collections | Articles on similar topics can be found in the following collections

Other Epidemiology (1588 articles) |
|--------------------------|--|

Notes

To order reprints of this article go to:
<http://www.bmjournals.com/cgi/reprintform>

To subscribe to *Journal of Epidemiology and Community Health* go to:
<http://www.bmjournals.com/subscriptions/>

THEORY AND METHODS

Self report in clinical and epidemiological studies with non-English speakers: the challenge of language and culture

Sonja M Hunt, Raj Bhopal

J Epidemiol Community Health 2004;**58**:618–622. doi: 10.1136/jech.2003.010074

Internationally, there is a drive for equality in health care for ethnic groups. To achieve equality, produce sound policies, and provide appropriately targeted services good quality data are essential. Where data are based upon self report, especially from non-English speakers, there are major barriers to the accumulation of reliable and valid information. When data collection instruments designed for English speakers are simply translated into ethnic minority languages, measurement error can result from inadequate translation procedures, inappropriate content, insensitivity of items, and the failure of researchers to make themselves familiar with cultural norms and beliefs. More attention should be paid to conceptual and cultural factors especially in epidemiological and clinical studies where self report is used to gather data. More interdisciplinary collaboration is necessary as well as a modification of customary methods of data collection and the assumptions behind them. The essence of such modifications entails participatory research with members of the linguistic communities concerned.

Ethnic boundaries are imprecise and fluid and the grouping of people has too often ignored significant differences of religion, habits, and language. As this paper is concerned with the gathering of information obtained by self report we are focusing on preferred language rather than ethnic group itself.

A recent editorial in the *British Medical Journal* summarised the general problems of obtaining high quality data from self report by non-English speakers.⁵ However, this is a very complex and important topic that has received scant attention in the epidemiological literature. Accordingly this paper discusses the issues in greater depth, together with the implications for comparative surveys.

THE COLLECTION AND UTILISATION OF SELF REPORT DATA

The problems of meaningful communication with non-English speakers by general practitioners, nurses, and hospital doctors have received some attention.^{6,7} However, there has been insufficient acknowledgement of the issues in relation to epidemiological studies and surveys. In the health domain self report data are gathered for several purposes (box 1) using a variety of methods ranging in precision from casual questioning through unstructured or semi-structured qualitative interviews, to standardised interview schedules and self completed questionnaires. These are produced with varying degrees of refinement. Some undergo extensive testing for validity, responsiveness, and reliability, others little or none.

Measurement error results from four major sources. Firstly, latent variables are not shared across languages. This might be the case with descriptions of mental wellbeing for example.¹¹ Secondly, from the original selection of questions and response constraints, with respect to ambiguous wording, lack of clarity, awkward or inappropriate categories, and the order and context in which the items are presented. Thirdly, from the respondents themselves, in relation to lack of understanding, misinterpretation or confusion, lack of motivation, and/or the perceived social desirability of certain answers. Fourthly, from researchers not being informed about the populations to be sampled, for example not being familiar with cultural norms and conventions and failing to consult translators about the appropriateness of questions.

There has been a great deal of research, in the social sciences, on the way in which respondents interpret and respond to oral or written questions. When a person is confronted with a

The assessment of the health and healthcare needs of ethnic minority populations are acknowledged priorities for both health and social services in the USA, UK,¹ and increasingly in Europe. The UK Race Relations (Amendment) Act 2000 places responsibility for equality between “races” on all public sector bodies. Internationally, there are increasing numbers of policies and legal requirements based on similar principles. A sound scientific database is necessary to ensure that services and initiatives are targeted appropriately for the groups in question. In the medical domain, reliable information is required on symptoms, health related behaviours, disease patterns, healthcare needs, use of services, and outcomes. Non-English speakers have been excluded from clinical trials and epidemiological studies, one of the several reasons for which may be the lack of valid and reliable instruments.^{2,3}

There are major barriers to the accumulation of reliable information on ethnic minorities, particularly newer and older immigrants and refugees who may have little or no competency in English. About 23% of immigrants to Britain born in China, Bangladesh, India, and Pakistan have no functional skill in English and 70% cannot function fully in an English speaking social environment.⁴

See end of article for authors' affiliations

Correspondence to: Professor R Bhopal, Public Health Sciences, Division of Community Health Sciences, University of Edinburgh, Medical School, Teviot Place, Edinburgh EH8 9AG, Scotland; Raj.Bhopal@ed.ac.uk

Accepted for publication 15 November 2003

Box 1 Uses of self report data

- As an integral part of the clinical interview
- As an adjunct to clinical measures, for example using standardised questionnaires such as the SF36⁸ for the assessment of general health status, to assess degree of disability, to check for the presence and severity of symptoms, or to gather patient assessed outcomes, increasingly included in clinical trials and other studies of treatment efficacy.
- As a part of epidemiological studies, for example gathering data on health related behaviour to monitor changes in personal habits, the use of diagnostic tools such as the Rose Angina Questionnaire,⁹ and for periodic international data collection, such as that on child and adolescent health.¹⁰
- By the NHS in relation to health needs assessment for planning and targeting of services.
- In studies of satisfaction with health care, for example women's views of maternity services.¹¹

questionnaire item a series of cognitive processes are set in train (box 2). These processes affect the quality of the data obtained.¹²

Self report data may assess some phenomenon within a particular group, say, satisfaction with care and, sometimes, to monitor changes over time. This requires only that the measure used is salient to and appropriate for the group concerned and is valid and reliable. If the data are to be used to make comparisons between groups as in clinical trials and most epidemiological studies, then the questions must be conceptually and functionally equivalent and salient for all the groups compared.

TRANSLATION AND ADAPTATION

It has been customary to translate questions originally developed for native English speakers into the requisite language(s). An assumption is made that the modes of inquiry, types of assessment, and research methods appropriate for native English speakers can be applied to other linguistic groups. Experience in translating questionnaires and interview schedules began primarily in the USA, in the field of cross cultural psychology.¹³ Initially one or more professional translators would take material in English and translate it into the target language. Where more than one translator was involved translations would be compared and agreement negotiated. The focus was on achieving linguistic equivalence. Consequent field testing led to the realisation that bi-lingual people are not at all representative of the population from which they come being biased by age,

Box 2 Cognitive processes triggered by questioning

- Comprehension and interpretation of the purpose of the question
- Information retrieval and reconstruction
- Judgment about what is required
- Evaluation of the situation and a decision about what to reveal
- Selection of a response.

Key points

- Collecting self reported data by ethnic group in multi-ethnic settings is necessary and difficult
- There are major ethnic variations in many measures of self reported health and risk factor status
- Translation into appropriate languages and back translation are necessary but insufficient steps
- The social sciences literature offers guidelines on how to conduct surveys in cross cultural settings
- There is an urgent need to improve the cross cultural validity of survey methods, particularly in epidemiological and public health research in multilingual, multi-ethnic societies

education and, often, gender and they produce translations that are too formal and literary for most people.¹⁴

Currently, the most sophisticated translation techniques are applied in the field of patient assessed outcomes where methods have evolved to a prolonged process of item selection, testing, and retesting and consultations with people monolingual in the target language(s). This development has been fuelled by the availability of funding from pharmaceutical companies carrying out clinical trials where there will be insufficient numbers of patients in any one country and the increasing requirement of regulatory bodies to include information from patients themselves.

"State of the art" translation/adaptation uses an iterative process with several stages^{15 16} as summarised in box 3.

However, even this prolonged degree of testing has been criticised for failure to achieve items in other languages that are comparable to the original English in terms of appropriateness and meaning.¹⁷ Where multiple languages are involved single languages are translated separately so although each may bear some resemblance to the English, the different languages are not necessarily close to one another. Often there remain considerable differences of concept between languages simply because it is impossible to find equivalent translations. For example the term "feeling blue", which is used in the original American version of the SF36 has different connotations when translated into different languages.¹⁸ In relation to languages such as Arabic, Cantonese, Punjabi, or Swahili, which have different roots to those of English, these issues are starkly highlighted.

Box 3 State of the art translation/adaptation procedures

- Translation of items by a team of bi-linguals
- Comparison of translations
- Negotiation of "best" items
- Consultations with people who are monolingual in the target language(s)
- Item refinement
- Field testing with monolinguals
- Refinements as needed
- Testing for face, content, construct, and criterion validity in each language.
- Testing for reliability and responsiveness
- Statistical analysis of ratings of quality of translation across different countries

For example, the terms “check up” and “Pap smear” have no conceptual equivalent in any Chinese language.¹⁹

LANGUAGE AND CULTURE

It is important to consider conceptual matters, cultural relevance, and the subtle connotations of words and phrases within a particular group.²⁰ A distinction can be made between language and culture. For example, both England and USA are primarily English speaking but their citizens do not necessarily share the same values with respect to health. In addition, because of the differing healthcare systems, socialised and private medicine, the readiness of patients to admit to health problems may be affected, as may their satisfaction and compliance with medical regimens. It is important to bear in mind that the content of a questionnaire reflects not only the language of the originating country but also the standards, expectations, values, and preoccupations of both the researchers and the lay people involved in the developmental procedures. Even when people belonging to another culture speak fluent English they do not necessarily share the beliefs and values of native English speakers. Moreover, their interpretation of the meaning of a question may be somewhat different because there is evidence that bi-linguals process information differently than do monolinguals.²¹

Research and health needs assessment and have assumed that data from different ethnic groups can be compared. Western concepts of “health”, “risk”, and “need” for example are so dominant that it is easy to forget that there are alternative views. According to the evidence cognitive processes are universal across cultures but the content of those processes is clearly not, although many elements may be shared.²² One example would be in the distinctions drawn between family and friends in English. This is not nearly so clear cut in Asian and African cultures. Thus, a question such as “Has anyone in your family had heart disease?” may elicit a response drawn from a wider frame of reference in these populations than in northern Europeans.

Comparability of data across languages is not always achievable. For example, an analysis of datasets from six different languages based upon the World Health Organisation’s Quality of Life Scale were subjected to the Rasch model of measurement.²³ Although there were some health related concepts that were similar across cultures, there were others which were very dissimilar, particularly those related to mental health. This study strongly suggested that equivalence of a questionnaire across several cultures is unrealistic at all but a very basic level.

EPIDEMIOLOGY AND SURVEY RESEARCH

Questionnaires

In epidemiology and health survey research the methods for obtaining equivalent information from different language groups trail behind the fields of sociolinguistics and anthropology. Indeed, the design and development of questionnaires used in epidemiology in general, has been criticised on the grounds they are too often ad hoc and of poor quality.²⁴ Two research projects carried out in this department have demonstrated these issues.

Fifteen local and national surveys of alcohol and tobacco use that had involved ethnic minorities were examined. Prevalence data on the consumption of alcohol and cigarette smoking differed by ethnic group. However, these differences were not consistent across studies. As the data were obtained, in many cases, from translated interview schedules and questionnaires doubts were raised about the quality of the resulting information. Accordingly the way in which items had been translated from the English was examined and compared with state of the art criteria. Only two of the

studies had used more than one translator. One had used consultations with members of the language group concerned to investigate cultural sensitivities and one had consulted monolingual people on the adequacy of the translations. None of the studies had tested the questions for validity, reliability, or responsiveness and none had compared the translated questionnaires with languages other than English. Consultations held with Bengali speakers indicated that there were serious problems both with the translations and with cultural appropriateness, for example, asking Muslims about alcohol use, using expressions such as “weekend”, “Christmas”, and “hangover”, which had low relevance for the respondents.²⁵

Another study concerned translations into Punjabi and Cantonese of the Rose Angina Questionnaire (RAQ). These translations were made by professional translators without any input from monolingual people representative of those to whom they would eventually be administered. A review of translations of the RAQ showed that, most often, no details of translation methods were given at all and face and content validity were not assessed in any of the studies.²⁶ In depth interviews with Punjabi and Cantonese speaking peoples have highlighted issues of incorrect translations, inappropriateness, and lack of cultural relevance (Hanna L, PhD thesis forthcoming; personal communication).

Interviews

Where face to face interviews use an interview schedule complications can arise because there exist somewhat different forms of the same language, for example, Bengali and the Sylheti variant of Bengali, which has no written form. For some languages, for example, Arabic and Cantonese the written and spoken forms are not the same. This means that if an interviewer asks questions in one of those languages the actual questions asked will not be the same as the questions as written with unknown effects on data quality.

The customary interview situation may violate expectations of other cultures about the normal way to interact, especially in face to face situations. Such interviews constitute highly unnatural social situations regardless of culture. For example, the questions have been decided in advance by some unknown party and not discussed with the respondent, the topic may switch abruptly, the response options may not fit the situation, the interviewer uses standard wording regardless of the respondent’s response, the interviewer is not supposed to elaborate or discuss to preserve standardisation. Thus requests for clarification meet with a standard response and the answers may not be grounded in the everyday life of the interviewee.²⁷ These issues may be exacerbated in the case of people from cultures where surveys are uncommon or unknown.

SOME SOLUTIONS

In the health research and healthcare fields translation/adaptation of interview schedules and questionnaires from English into other languages have suffered from faults in relation to cultural hegemony, failure to ensure that the phenomenon of interest is present in all target groups, lack of salience of content, the non-equivalence of concepts, assumptions about willingness to disclose certain types of information, and the use of levels of language not easily comprehensible to the less well educated. More interdisciplinary cooperation could help ensure that advances in one field pertinent to another may be shared, for example cognitive aspects of survey methodology, health measurement, epidemiology, survey research, linguistics, and ethnomedicine.

Rather than pursuing cross cultural equivalence an alternative is to search for both emic and etic items referring to the topic of interest. Emic refers to those issues that are salient and meaningful only within particular cultures while etic refers to universal concerns such as the welfare of children or the inability to carry out daily activities. In relation to health behaviours asking if a person smokes tobacco products may be universally appropriate. However, asking if a person chews paan is unlikely to be. This strategy requires a much less ethnocentric and more participatory approach whereby monolingual and bi-lingual representatives of the target group(s) are involved at all stages of the research. Initial inquiries would ascertain which matters are of importance to the group concerned and generating items for inclusion in a mode of inquiry congenial to that group. The process consists of a collaborative spiral of inquiry, reflection, action, planning, and discussion. The end result would be a set of questions that would share common items supplemented by culture specific information. Measures developed in this way would be different in different languages with respect to some of the content but fully appropriate and salient to each. Such a procedure would allow for comparisons within groups over time and between groups for the shared items. The application of Rasch analysis, as previously mentioned, can assist in testing for the similarity of latent variables across languages.

Another approach is to focus upon the similarity of concepts rather than upon equivalence of items. For example, we might assume that the notion of physical wellbeing exists in all cultures but the implications could be different. For groups where prayer is important the ability to kneel may be essential to wellbeing, for others this may be less salient than the ability to play or even stand and watch football. It may not be necessary to have exact comparisons as long as the underlying purpose of the question is the same.

Translators should be required to advise not solely on the target language but also upon the cultural acceptability of the questions to be asked. In some cases bi-lingual people may have become so far removed from traditions and the community under study as to be culturally if not linguistically alienated. Unless requested to do so translators may not regard it as part of their task to comment on the salience or sensitive nature of the questions asked.

Researchers doing research with ethnic minorities should be cognisant of the customs, values, and beliefs of the target group(s) before designing any project. Issues of cross language data collection should be seen as a challenge and not as an obstacle, a stimulus to innovative thought and the development of new techniques of investigation. Cultural and linguistic differences have yet to be incorporated as fundamental to sound public health, primary and secondary care, and health promotion. Categories and concepts in health research based upon a Western epistemological order may impede the accumulation of high quality data from people born outside Western countries.²⁸ Individuals' reactions to illness and discomfort, their concepts of health, their behaviour, their help seeking is intimately bound up with cultural beliefs, values, and experience.

POLICY IMPLICATIONS

Policies to improve the health and health care of ethnic and linguistic majorities in the health and social services will not achieve their goals unless the cultural dimensions of self report are given fuller attention than hitherto. This poses a formidable challenge to policy makers, as this is no small task. In London, as in many metropolitan cities, numerous languages are represented.²⁹ The implications for research, if the concepts in this paper are to be implemented, are

substantial. Research funding bodies need to devise policies that consider the implications.

ACKNOWLEDGEMENTS

We are indebted to Lisa Hanna, Colin Fischbacher, and Leslie Alexander for their helpful comments on the several drafts of this paper.

CONTRIBUTORS

SMH had the original idea and wrote the paper; RB contributed ideas, helped redraft and commented at all stages.

Authors' affiliations

S M Hunt, R Bhopal, Public Health Sciences, Division of Community Health Sciences, University of Edinburgh, Medical School, Edinburgh, Scotland

Funding: none

Competing interests: none declared.

REFERENCES

- Raftery J**, ed. Health care needs assessment: black and minority ethnic groups. In: *Health care needs assessment. The epidemiologically based needs assessment reviews*. Third series. Abingdon: Radcliffe, Medical Press (in press). <http://hcna.radcliffe-online.com/bmeqframe.htm>.
- Bartlett C**, Davey P, Dieppe P, *et al*. Women, older persons and ethnic minorities: factors associated with their inclusion in randomised trials of statins 1990 to 2001. *Heart* 2003;**89**:327–8.
- Free C**, McKee M. Meeting the needs of black and minority ethnic groups. *BMJ* 1998;**316**:380.
- Carr-Hill R**, Passingham S, Wolf A, *et al*. *Lost opportunities: the language skills of linguistic minorities in East and West London*. London: Basic Skills Agency, 1996.
- Hunt S**, Bhopal R. Self report in research with non-English speakers. *BMJ* 2003;**327**:352–3.
- Cohen S**, Moran-Ellis, Smaje C. Children as informal interpreters in GP consultations: pragmatics and ideology. *Sociology of Health and Illness* 1999;**21**:163–86.
- Elderkin-Thompson V**, Silver R, Waitzkin H. When nurses double as interpreters: a study of Spanish-speaking patients in a US primary care setting. *Soc Sci Med* 2001;**52**:1343–58.
- Ware J**, Sherborne C. The MOS short form health survey (SF36): international conceptual framework and item selection. *Med Care* 1992;**30**:473–81.
- Rose G**, Blackburn H, Gillum R, *et al*. *Cardiovascular research methods*. Geneva: World Health Organisation, 1982.
- Michaud P-A**, Blum R, Slap G. Cross-cultural surveys of adolescent health and behaviour: progress and problems. *Soc Sci Med* 2001;**53**:1237–46.
- Fenton S**, Sadiq-Sangster A. Cultural relativism and the expression of mental distress. South Asian women in Britain. *Sociology of Health and Illness* 1996;**18**:66–85.
- Jabine T**, Straf M. *Cognitive aspects of survey methodology, building a bridge between disciplines*. Washington DC: National Academic Press, 1984.
- Triandis H**. Major theoretical and methodological issues in cross-cultural psychology. In: Davidson J, Louner W, eds. *Readings in cross-cultural psychology*. Hong Kong University Press, 1974:26–38.
- Hendricson WD**, Russell J, Prihoda T, *et al*. An approach to developing a valid Spanish language translation of a health status questionnaire. *Med Care* 1989;**27**:959–64.
- Hunt SM. Cross-cultural comparability of quality of life measures. In: Guggenmoos-Holzmann I, Bloomfield K, Brenner H, *et al*, eds. *Quality of life and health: concepts, methods and applications*. Berlin: Blackwell Wissenschafts-Verlag, 1995:15–26.
- Guillemin F**, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;**46**:1417–32.
- Stewart AL**, Napoles-Springer A. Health-related quality of life assessments in diverse population groups in the United States. *Med Care* 2000;**38**(suppl 9): I1102–24.
- Bullinger M**, Alonso J, Apolone G, *et al*. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Association. *J Clin Epidemiol* 1998;**51**:913–23.
- Pasick R**, Stewart S, Bird J, *et al*. Quality of data in multiethnic health surveys. *Public Health Rep* 2001;suppl 116:223–43.
- Hunt SM**. Cross-cultural issues in the use of quality of life measures in randomized controlled trials. In: Staquet M, Hays R, Fayers P, eds. *Quality of life assessment in clinical trials*. Oxford: Oxford University Press, 1998:51–68.
- Holme J**. *An introduction to sociolinguistics*. London: Longmans, 1992.
- Warnecke R**, Johnson T, Chavez N, *et al*. Improving question wording in surveys of culturally diverse populations. *Ann Epidemiol* 1997;**7**:334.
- Leplège A**, Ecosse E and the WHOQOL Rasch Project Scientific Committee. The analysis of four WHOQOL-100 data sets (Argentina, France, Hong-Kong, UK). *Journal of Applied Measurement* 2001;**4**:389–418.

- 24 **Olsen J.** Epidemiology deserves better questionnaires. *Int J Epidemiol* 1998;**27**:935.
- 25 **Vetini A,** Bhopal R, Hunt S, *et al.* *Measurement of risk factors for cancer in ethnicity and health research: a case study of tobacco and alcohol. Report to the Scottish Cancer Group of the Scottish Executive.* Edinburgh: Section of Public Health Sciences, Edinburgh University, 2002.
- 26 **Fischbacher C,** Hanna L, Bhopal R. International and cross-cultural comparisons of angina prevalence using the Rose Angina Questionnaire: have Rose's warnings been heeded? Joint conference of the international epidemiological association and society for social medicine, Oxford, 12–15 Sep.
- 27 **Cicourel A.** Interview surveys and the problem of ecological validity. *American Sociologist* 1982;**17**:11–20.
- 28 **Airhihenbuwa C.** *Health and culture: beyond the Western paradigm.* Thousand Oaks, CA: Sage, 1995.
- 29 **Baker P,** Eversley J. *Multi-lingual capital. The languages of London schoolchildren and their relevance to economic, social and educational policies.* London: Battlebridge Publications, 2000.

ECHO

Responses to a fourth dose of *Haemophilus influenzae* type B conjugate vaccine in early life

M H Slack, D Schapira, R J Thwaites, M Burrage, J Southern, D Goldblatt, E Miller



Please visit the *Journal of Epidemiology and Community Health* website [www.jech.com] for a link to the full text of this article.

Objective: To describe the immune response of preterm infants, with a reduced response to primary *Haemophilus influenzae* type B (Hib) immunisation, to a fourth dose of Hib conjugate vaccine given in early life.

Design: Prospective observational study.

Setting: Five Wessex Neonatal Units.

Patients: Infants born at < 32 weeks and immunised with three doses of combined acellular pertussis-Hib vaccine, with a Hib IgG geometric mean concentration (GMC) < 1.0 µg/ml after these primary immunisations.

Interventions: An additional fourth dose of Hib conjugate vaccine given before 1 year of age. Blood taken to assess Hib IgG concentration and avidity after immunisation.

Main outcome measures: Hib IgG GMC and avidity index.

Results: Ninety six infants (mean gestational age at birth 29.1 weeks) received a fourth dose of Hib at a mean age of 7.8 months. Hib IgG GMC after the primary immunisations was 0.17 µg/ml (95% confidence interval (CI) 0.14 to 0.20) rising to 4.68 µg/ml (95% CI 3.36 to 6.57) after the fourth dose ($p < 0.0001$). The IgG response to the fourth dose correlated positively with the response after the primary immunisations ($p < 0.001$). Hib IgG geometric mean avidity index (GMAI) after the primary immunisations was 30.87 (95% CI 20.40 to 46.73). This increased to 124.73 (95% CI 109.93 to 141.51) after the fourth dose ($p < 0.0001$).

Conclusion: Preterm infants with very low IgG responses to Hib after primary immunisations with a combined acellular pertussis-Hib vaccine mount a good response to a fourth dose of Hib. This study suggests that all infants will benefit from a fourth dose of Hib, regardless of the age at which it is given.

▲ *Archives of Disease in Childhood Fetal and Neonatal Edition* 2004;**89**:F269–F271.