



Computer
Assisted
Telephone
Interviewing
Technical
Reference
Group

Population Health Monitoring and Surveillance:
CATI Information Question and Module Development

Principles and Practices

November 2001

CATI Technical Reference Group
National Public Health Partnership

Computer Assisted Telephone Interviewing (CATI) is a methodology widely used for surveillance of health behaviours and health outcomes in populations in Australia. The National CATI Health Survey Technical Reference Group (CATI TRG) is an advisory committee to the National Public Health Information Working Group under the National Public Health Partnership. Members of the CATI TRG include representatives from State/Territory Health Departments, the Commonwealth Department of Health and Ageing (DoHA), the Australian Bureau of the Statistics, the Australian Institute of Health and Welfare and the Public Health Information Development Unit at the University of Adelaide. Since its inception in 1999, the CATI TRG has been a forum for the development and promotion of national standards, valid methods and capacity for CATI health surveys and health surveillance.

To embark in the efforts towards 'harmonisation' of CATI health surveys in Australia, the CATI TRG has identified the need to develop question modules for behavioural risk factor and chronic disease topics based on well-developed conceptual frameworks that underpin the data requirements for health surveillance. The proposed question modules are set to undergo a rigorous process of cognitive and field-testing under the guidance of the CATI TRG and the results will be published in a question module manual as a key reference to those interested in CATI health surveys in Australia.

This paper has been prepared by the CATI TRG as part of a series, with funding predominantly from the DoHA. Its preparation has involved input from all State and Territory jurisdictions, DoHA, the Australian Bureau of Statistics, the Australian Institute of Health and Welfare and the Public Health Information Development Unit at the University of Adelaide as well as recognised content experts.

Any comments or information relevant to the subject matter of this background paper would be welcome. Correspondence should be directed to:

Chair
National CATI TRG
c/- Population Health Data & Information Services Unit
Australian Institute of Health & Welfare
GPO Box 9848 (MDP 16)
CANBERRA ACT 2601

Tel: 02 6244 1000
Fax: 02 6244 1299
Email: info@aihw.gov.au

Population Health Monitoring and Surveillance:
CATI Information Question and Module Development

Principles and Practices

November 2001

CATI Technical Reference Group
National Public Health Partnership

Contents	Page
Executive Summary	iii
Recommendations	iv
Introduction and Overview	1
Data Requirements and Question Selection	6
Pre-Testing	9
Question Design	11
Validity and Reliability	12
Field Testing	13
Usability Testing	14
Reporting	15
Interviewers and Supervisors	15
Additional Issues	17
Bibliography	20
Appendix A: Stages in question development	28
Appendix B: CATI module standard consultation process	29
Appendix C: Additional Topic References	30
Appendix D: Stage One Template Example for the Collation of CATI Questions	32

Executive Summary

In Australia, the proportion of households with fixed (or cordless) telephones has in recent years risen to levels sufficient for the conduct of CATI population surveys, with 97.8% of households having fixed telephones. The National Computer Assisted Telephone Interview Technical Reference Group (CATI TRG) under the National Public Health Partnership has emphasized the need to establish question and module development standards for Computer Assisted Telephone Interviews (CATI) that can be utilised across Australia.

Recently the CATI TRG, a subcommittee of the National Public Health Information Working Group (NPHIWG), developed a detailed appraisal process to evaluate questions that are already in use in State/Territory CATI health surveys. This overview focuses on question development principles and practices and provides a guide to the various stages that are required to ensure that 'best practice' question development is followed.

In recent years there has been a shift to explicitly define any relevant policy and data requirements for health monitoring and surveillance at the beginning of the process. In any question development, the first step should be to incorporate relevant policy and data requirements into any framework that is being used to direct the decisions on what information to collect. Documentation of the policy questions that are relevant to the topic and a statement of data requirements for health surveillance are also recommended. Some recommendations are made that arise from the document and are submitted as essential steps in CATI question development.

It advocates following the basic principles which can be found in any survey methodology text and briefly describes the major features including establishing the validity and reliability of the question usually done by field testing. This document stresses the iterative process of question development pointing out that results of field testing are used to refine questions and inform interview training. Interviewer debriefing is considered an essential part of this process. In addition, focus groups can be used in question development and after the survey has taken place to evaluate the questionnaire and interviewing procedures.

This document focuses on question development and outlines principles and practices for the construction of consistent and uniform questions (or sets of questions) in CATI. It covers the identification and assessment of questions, key processes in the development of questions and the processes to test validity and reliability. Fieldwork issues, interviewer feedback and reporting on questions are briefly discussed. The final section is devoted to briefly discussing some additional issues associated with health surveillance systems that use a CATI mode of administration. Appendices of helpful information are attached for information.

Recommendations

The following recommendations are presented with regard to the development of CATI questions. It is recommended that:

- a structured approach to the selection of variables for health survey data collection and analysis be used;
- policy and data requirements be incorporated into the decision making process on what information to collect;
- questions chosen as part of a core component should depend on the priority given to the topics they cover and any the jurisdictional/national requirement to monitor these topics;
- the process for the development of questions for CATI will be publicly available;
- the appropriateness and performance of questions and question modules be considered on an ongoing basis; and
- special studies are regularly undertaken for groups/areas of interest that are not covered by CATI.

Introduction and Overview

Computer Assisted Survey Information Collection (CATI)

This document presents the principles and practices for the various stages that should be addressed in the development of questions/modules for computer assisted telephone interviews (CATI) in Australian health surveys. CATI is an approach that denotes the use of computers for survey data collection, data capture, data preparation, and associated tasks [24].

In developed countries CATI health surveys are used for monitoring, surveillance, and research to provide information to support policy. Since the mid-1980s the Behavioral Risk Factor Surveillance Survey has been collecting State CATI health survey information in the United States (US) on an ongoing basis. In 1998, the State and Local Area Integrated Telephone Survey has been introduced by the US Centers for Disease Control and Prevention, with an emphasis on children's health survey data. In Europe a number of countries have introduced CATI health surveys and some harmonisation of surveys across countries is underway.

A number of Australian jurisdictions introduced CATI health surveys in the 1990s. In 1998, a working group of the National Public Health Partnership (NPHP) was established with an aim to standardise the collection of CATI health survey data. The material presented in this document has been drawn from a number of sources, including the work of the CATI Technical Reference Group (CATI TRG), a working group of the NPHP.

Over the last twenty-five years internationally there have been major developments in the methods and techniques of data collection for sample surveys and CATI has predominantly become the standard survey technique [39]. The lower cost of data collection, improved sampling methods through random digit dialing and improved technology have been put forward as key reasons for the increased interest in telephone surveys [56]. During the same period there has been recognition that question meaning, wording and ordering have a major impact on responses in survey interviews. In addition, there has been increasing research into the issues of recall biases related to the changing data collection technology [97].

Purpose

This document outlines principles and practices for the development of consistent and uniform questions (or sets of questions) in CATI. It covers the identification and assessment of questions, key processes in the development of questions and the processes to test validity and reliability. In addition, subsequent fieldwork issues, interviewer feedback and reporting on questions are briefly discussed.

The procedures outlined in this document are thought to be essentially the same whether one question or a group of questions is being developed. A more detailed document on the principles and practices of question and module development more broadly covering CATI as well as computer assisted personal interviewing

has been prepared as a Commonwealth Department of Health and Aged Care Occasional Paper.

Traditionally the major contributors to the development of survey questions have been questionnaire designers working closely with subject matter experts, with input from experienced interviewers. In recent years there has been a shift to explicitly define at the beginning of the process the policy and data requirements for health monitoring and surveillance, and later in the process to include an understanding of questions from the perspective of the respondent, with the operational aspects of questionnaire development and testing incorporated into survey planning [29].

A summary of the question development stages is presented in Appendix A. These stages are similar to the processes outlined by Dillman (2000) and those employed by Statistics Netherlands [2]. The document follows the stages emphasising the need for a systematic and rigorous approach to question development. It should be noted that the later stages provide valuable feedback to the earlier development stages thus emphasizing that the monitoring and assessment of questions is a continual process with a need to monitor the ongoing appropriateness of survey questions.

The requirements for different topics will vary depending on a number of factors, particularly the availability of existing questions. Other factors include the time available for testing, the level of funds available, human resources, and the aims and objectives of the researchers (such as consistency over time versus maximizing item quality).

For example, in the topic of smoking there are a wide range of questions that have been developed in the Australian setting and it could be expected that most of the stages outlined in this document will have already been conducted. On the other hand questions related to community capacity and gambling have not been widely developed and would require full development.

CATI Advantages and Disadvantages

CATI methods can reduce costs, improve timeliness and improve the quality of collected data compared to traditional paper-and-pen based data collections. CATI methods have also shown improvements in questionnaire administration, as well as interviewer and respondent acceptance. CATI procedures can have a dramatic effect with:

1. more complicated routing patterns of respondents through the questionnaire;
2. on-line range, consistency, and data completeness checks;
3. automated prompting and standardised prompts for error resolution;
4. on-line 'help' utility;
5. customised wording of complex questions based on preloaded data or answers to previous questions;
6. ability to access data from external sources and use it in the interview; and
7. randomisation of question order and response options [55, 103].

Lavrakas (1987) identifies that the major disadvantages are both the limitations in the forms of questions that can be asked and the length of questioning, particularly in CATI interviews. In addition, he believes that “the quality of data that (are) gathered (via CATI) is directly related to the quality of interviewing that is performed. This in turn is a function of the skill of individual interviewers and the rigor of (the) systematic routine that interviewers are expected to follow”. [59]. One might argue that this is true of all interviewing modes of administration. It would be recommended that for all questions that are to be administered by interview, question development and interviewer instructions are thoroughly tested and evaluated. Further if the interview is to be conducted by CATI, then the questions are tested in the CATI environment as the quality of CATI surveys is affected by several factors, including the quality of responses obtained from respondents [93].

CATI health surveys conducted by properly structured interviews have an advantage over other in-person and mail surveys with the data collection closely monitored and supervised.

Is CATI appropriate for a surveillance system?

In Australia, the proportion of households with fixed (or cordless) telephones has in recent years risen to levels sufficient for the conduct of CATI population surveys. Currently 97.8% of households have fixed telephones [8]. However, there are important groups such as the remote indigenous population, the homeless and those in institutions who would not be covered by CATI health surveys despite this high percentage of Australian households with fixed telephones.

Another issue to consider is the recent dramatic changes in the telecommunication industry with the high take-up of mobile phones in Australia. This could be expected to have a major impact on CATI. In 2000, 58.5% of households had at least one mobile phone [8]. This means that CATI may have to make some changes to sampling techniques and other issues related to the ‘individual’ versus the ‘household’.

CATI Questionnaire Development Issues

DeLeeuw (2000) explains that the main advantages of computer assisted interviewing are improved data quality, efficiency and lower costs. Particularly with the automatic routing of questions and the use of range and consistency checks, data quality has improved [28]. However, there has been a re-examination of other non-sampling errors, including errors in survey measurement. She observes that

“computer assisted data collection is no panacea for good data quality. It requires one to do almost everything that is needed with a good paper-and-pen interview or questionnaire, and to add extra effort in computer implementation, in testing the questionnaire, ... in extra interviewer training, and in designing a respondent friendly and trustworthy questionnaire.”

Her views are not new but they highlight a number of key areas that require specific attention with the development of questions for CATI. Questionnaire construction and questionnaire administration are two of the major sources of bias in data collection [38]. To quote Platek (1985) “errors resulting from unrealistic demands on the respondent’s knowledge or memory, the use of overly difficult and technical language, or the excessive demands on respondent’s patience are all sources of non-response which have their roots in the questionnaire” [79].

Deming (1944) highlighted the imperfections in the design of questions and questionnaires [32].

“Faulty design ... can be the cause of considerable bias. Faulty design often arises from lack of knowledge of the subject matter ... An understanding of the subject is accordingly demanded as one of the qualifications for planning a survey.” (p364)

‘Best Practice’ has been defined as a “holistic strategy”, a jigsaw with all elements interlocking and connected [10]. In this situation, question development can be seen as the integration of the selection, evaluation and testing of questions within the CATI environment, combined with interviewer procedures and supervision. Willis et al. (1999) identified that the use of different techniques in question development (specifically the use of cognitive testing, field testing and the use of ‘experts’) were complementary rather than conflicting with one another [105].

The pretesting policy of the US Census Bureau presents mandatory standards for question and questionnaire development. This policy emphasizes that a survey question and questionnaire must ‘work’ and if there is insufficient evidence about how well a question performs then the question must be subjected to some form of questionnaire pretest along with field interviews [101]. This pretest policy also requires quantitative measures of validity and reliability to ensure that question and questionnaire results can be generalised to the population of interest.

Similarly, Statistics Netherlands has an extensive program of questionnaire development. The Questionnaire Design Resource Center provides design, development and testing that incorporate ordinary interviews, in-depth interviews, cognitive interviews and focus groups [3].

This overview of question development principles and practices does not provide the specific prescription on issues such as question wording and format. It concentrates on the development of uniform and consistent questions, although Platek (1985) observes that a questionnaire should not be too rigid and must be flexible to adapt to respondents of different age/sex groups, language and social backgrounds. Different words or examples may be needed in order to convey the desired meaning to all respondents, and the questionnaire must be able to anticipate all possible answers. It is therefore necessary to develop clear guidelines for interviewers and supervisors, providing the range and extent of prompts and expansion of questions that will assist respondents to provide the required information without leading or introducing bias into the responses [79].

It is expected that with the development of questions for CATI the preparation and dissemination of reports that document the development process will be publicly available. The acceptance of the questions into the National Health Data Dictionary would also ensure the broad utilisation of standard (comparable) questions across data collections [9].

Data Requirements and Question Selection

There are a number of steps that can be followed in identifying the data requirements for health monitoring and surveillance, operationalisation of those data requirements, and selection of possible questions. They include the identification of relevant policy questions, a review of the literature, the input from subject matter experts, other stakeholders and key informants, and a systematic appraisal of existing questions.

Policy Questions

The specification of data requirements for health surveillance and selection of subject topics within a health survey can be a less than trivial task. Certainly, the topic (eg smoking) may well have been determined at an early stage, but the more precise data requirements/variables/questions both within the topic and associated with the topic require close attention. Accordingly, there is a need to examine (or re-examine) the concepts to be measured and the subsequent explicit specification of data requirements. From this it could be expected that there may be a need to develop new questions or modify existing questions.

A structured approach to the selection of variables for health survey data collection and analysis is required. Documentation of the policy questions that are relevant to the topic and a statement of data requirements for health surveillance by each jurisdiction is recommended. This should then be followed by an assessment of the relevant variables or indicators for each policy question and operationalisation of the data requirements [77].

To assist in this approach, the NPHP has been developing a performance indicator framework for population health that has been based on recent Canadian and US initiatives and more recently has developed a Chronic Disease Framework [72]. Individual jurisdictions will also have their own set of indicators and/or frameworks. The important thing is that questions are anchored to a documented framework and this framework provides a mechanism to identify variables/questions requiring development. Embedding question development within a framework reinforces the associations between the health indicators and outcome measures.

Literature Review

A thorough review of the literature to identify existing data collections and the method of collection used previously, is a key step in the identification of possible questions and variables related to the particular topic and associated variables/questions. It is desirable to identify as broad a range of existing questions as possible.

The preparation of the General Nordic Questionnaire provides an example of the use of an extensive review of the literature on sets of questions to derive a comprehensive survey program [63]. The review of the literature for each set of questions was lead by a researcher familiar with the subject. A summary of the

review, along with a draft questionnaire, was published by the Nordic Council of Ministers for wide consultation prior to testing.

One important consideration when identifying existing questions from other sources is that the same question in a different collection mode is a different question [32]. Accordingly, having gained questions, it is important that the question is tested and evaluated in the environment of the CATI (particularly CATI) health survey and also in the Australian setting.

From a review of the literature there can be an identification of key questions/variables associated with the topic being investigated. The literature review may result in the elimination of unnecessary questions and the incorporation of new questions. For example, if there has been demonstrated to be no difference in health status between those who are self-employed and those who are salaried workers, then this question might be excluded. However, if new evidence is being reported on the relationship between employment security, control at work and health status then questions on work organisation and health should be investigated and developed.

Key Informants

It is fundamental to include subject matter experts, policy and program developers, other stakeholders and key informants in the process of definition of data requirements for health surveillance and subsequent question selection. Key informants are usually those individuals who have demonstrated extensive knowledge in the topic, such as researchers, academics, policy makers and representatives of special interest groups. They provide an important link to the existing literature and can provide expansion of possible collections that have not been reported in the (predominantly) academic literature.

Question Appraisal

Recently the CATI TRG, a subcommittee of the National Public Health Information Working Group (NPHIWG), developed a detailed appraisal process to evaluate questions that are already in use in State/Territory CATI health surveys. These initiatives are summarised in Appendix B which presents the processes for the question module development. This model follows best practice and is recommended for use with any sort of health related data element development.

For some additional sources to appraise alternative questions, Appendix C provides some key references that might be helpful for further information. In addition, Appendix C also includes a number of Australian and international web-sites that provide questions currently in use.

Stage 1 of the CATI TRG appraisal process involves the collation of all available questions: an example of the template developed by the CATI Technical Reference Group is presented in Appendix D. In particular, it can be seen that this collation process incorporates a structured approach that documents specific policy questions relevant to the question topic [21].

Summary

Policy Questions

- what are the policy questions relevant to the topic?
- how does the topic fit into the performance indicator framework?

Literature Review

- what questions have already been used for the topic?
- do questions exist in major collections such as the BRFSS or CCHS?

Key Informants

- are there researchers, academics, policy makers or representatives of special interest groups who can advise on the topic?

Question Appraisal

- can existing questions be used?
- do CATI questions meet the data requirements of the CATI TRG and NPHIWG?

Data Requirement Consultation

- what are the data requirements for health surveillance?
- are they anchored to a documented framework either jurisdictional or national?
- do they include advice from national expert and national strategy groups?
- have data requirements achieved endorsement of the CATI TRG and NPHIWG?

Pre-Testing

Once the data elements have been identified and questions developed or found that are pertinent to these, a pre testing of the questions is recommended, if that has not been done before. In recent years there have been significant improvements in the art and science of survey question development, with the application of the principles of cognitive psychology, the value of qualitative and ethnographic research methods, and understanding of language intended for respondents. Focus groups and ethnographic interviewing have been used to solicit opinions on research topics, with the particular procedures chosen dependent on the nature of the topic. For example, in-depth interviews may be preferred when the topic relates to more personal issues. The use of design oriented focus groups and cognitive interviewing in a laboratory setting are now seen as complementary processes that can evaluate many aspects of question design [1, 33, 89].

With the use of standardised and sometimes technical language in survey questions, it has been recognised that many respondents may have difficulty understanding the meaning and intent of survey questions [1]. Valuable insight into the issues involved in a question topic can be gained through the use of group interviews that focus discussion on the particular topic that the survey questions are endeavouring to measure [54]. Focus groups can be used in question development as well as after the survey has taken place to evaluate the questionnaire and interviewing procedures. [41].

Cognitive Testing

Survey design in developed countries has incorporated the use of cognitive testing since the 1980s. With cognitive testing, survey researchers attempt to understand how respondents have interpreted the questions in relation to the identification of potential problems with perception, sensitivity, comprehension, memory, context, format, vocabulary, reference periods, judgment and response categories [51, 107]. Questions may be difficult to answer if they are difficult to understand, if response requires detailed memory recall, or if response categories fail to cover the range of respondent experience [35].

Currently there is no set of standards for cognitive interviewing that have been developed, evaluated and systematically implemented. Cognitive interviewing approaches have evolved independently across organisations and reflect the preferences of the researchers in them [14].

There have been concerns that the results of cognitive testing need to be accepted by survey sponsors. Beatty et al. (1996) reported that the survey sponsors were concerned that given the interviewing methodologies used in the cognitive testing, reported difficulties subjects had in providing codeable responses might have been as a result of the open cognitive procedures. In an attempt to overcome this, Rothgeb et al. (2000) recommends that survey sponsors be given opportunities to observe the testing process [81].

For CATI questions it is essential to integrate the delivery of the questions over the telephone into the laboratory setting [12, 70, 83]. As an example, in the ABS cognitive laboratory in Canberra, telephone interviews from adjacent rooms can incorporate CATI interviews.

Summary

Focus Groups

Used to assess the question and answering process

- what are the underlying assumptions in the question?
- do the questions cover the topic?
- do the respondents understand what the questions are covering?

Cognitive Testing

Assessment of interviewing in a laboratory by trained cognitive interviewers

- in 'think aloud' interviews respondents report their thought processes while answering questions
- use of probing and follow-up questions
- going through questions a second time to understand the thought processes used by the respondent while answering questions
- hypothetical vignettes can be used to test situations that can be seen as likely problem areas in a topic.

Question Design

Fowler (1998) presents five basic characteristics of questions and answers that are fundamental to a good self report measurement process, namely:

1. questions need to be consistently understood;
2. questions need to be consistently administered or communicated to respondents;
3. what constitutes an adequate answer should be consistently communicated;
4. unless measuring knowledge is the goal of the question, all respondents should have access to the information needed to answer the question accurately; and
5. respondents must be willing to provide the answers called for in the question [36].

One important difference between computer assisted interviewing and paper-based surveys, is that there is no questionnaire in the usual sense in CATI surveys. The 'questionnaire' is a computer program which creates a unique questionnaire for each interview that contains only those questions that apply to the interview. To ensure that the presented construction and relationship of questions throughout the interview are in the desired order, flow charts of the different question responses are of great assistance [50, 53]. In addition, flow charts can assist in ensuring that errors in the preparation of the 'questionnaire' are reduced with questions having the appropriate links and skips in the interview.

Another key advantage of computer assisted interviewing is the ability to provide additional information for the interviewer to answer additional anticipated queries with the preparation of set introduction/selection procedures and the provision of fallback statements for use by the interviewers [47, 60]. Auxiliary information can be placed on the screen as part of the question or as a special area reserved for instructions. Unlike printed questionnaires this additional support for the interviewer is not limited by the layout or space in the 'questionnaire'.

However, respondents sometimes do just enough to satisfy the survey request, but no more. This has been termed 'satisficing', giving minimal acceptable answers, rather than optimal answers. To avoid many of the undesirable effects of 'satisficing', such as response order effects and acquiescence bias, question design must maximise respondent motivation, minimise task difficulty and minimise response effects.

Validity and Reliability

As an integral part of the testing and design stages, there needs to be specific attention to the validity and reliability of questions in the CATI environment [2].

Validity

Validity refers to the extent to which a question (or set of questions) measures what it is intended to measure and does not measure what it is not intended to measure [94]. In addition, because health measures can be used for different purposes, their validity needs to be evaluated separately for each purpose. There are a number of different aspects of the validity of questions with the three broad types being content, criterion and construct validity.

Reliability

Reliability of a question is the consistency and stability of the question from one survey to another. When the question is repeated and gives identical or very similar results, then it is said to be reliable [102].

Test-retest and internal consistency are the key reliability measures in question development. Questions with low reliability are ones in which the answers respondents give vary widely as a function of when the questions are asked, who asks them and the fact that the particular questions chosen from a set of items seem to be asking the same thing, but are not [1].

The aim of CATI question development is to standardise the delivery of the question by the interviewer to the respondent. Therefore, inter-rater reliability does not have a significant impact on CATI health survey questions, particularly those that are developed along the systematic lines outlined in this document.

Summary

Content Validity

- are all relevant concepts represented in the questions?

Criterion Validity

- how do the questions compare to a “true” value or “gold standard”?

Construct Validity

- how do the questions match to other related measures?

Reliability

- are responses to the questions consistent and stable from one survey to another?

Field Testing

Testing the questions in the field is a vital stage in the development and assessment of survey questions, along with the development and testing of associated procedures and instructions [75]. No matter how much development work has been done, the question (or set of questions) has to be tested under field conditions. In particular, the questions have to be tested in the mode they will be used; specifically CATI questions require field testing in a CATI survey environment.

Field testing can cover an evaluation of procedures by collection from a small sample of respondents, the identification of any production mistakes and learning whether respondents understand the questions [32, 73]. The pretesting policy of the US Census Bureau highlights that [101]:

“Changes in survey procedures, including the questions asked, may affect the continuity of time series data. Seemingly minor changes in question wording or sequence sometimes can affect survey responses in important and unexpected ways. When a time series measure may be affected by a questionnaire or procedural change, the Census Bureau recommends that an experimental field test be conducted to evaluate the effects of the proposed changes on survey estimates.”

It is therefore important, particularly for the collection of health surveillance data where changes over time are a key component, that as part of the evaluation of questions, there is measurement of the impact of changes to existing questions. If existing questions are to be replaced, then there is a need to provide estimates of how the replacement questions will affect the ongoing measures.

The ABS User’s Guide to Sample Surveys provides direction to the background for survey testing, stating that testing is used to:

- a. assess the suitability of the chosen sampling methodology;
- b. estimate sampling error and variability in the target population;
- c. estimate likely response rates;
- d. identify weaknesses in sample framework, questionnaire design, and the method of data collection;
- e. assess field work procedures, and processing procedures; and
- f. estimate costs [7].

Other forms of field testing include various versions of **pilot testing**. Pilot testing involves formally testing the questions with a small sample of respondents in the same way that the final survey will be conducted. A pilot test should incorporate a test-retest component to establish the question reliability. Where substantial revisions are required as a result of the pilot test then a subsequent pilot test should be conducted. Pilot testing is used to:

- a. identify problems with the questions;
- b. provide validation, reliability and usability assessment of the questions;
- c. assess the adequacy of instructions to interviewers and supervisors; and
- d. ascertain the time required to ask the questions.

As part of the pilot testing, debriefing of the interviewers is required to identify possible problems and anomalies with the questions [27, 34, 101]. From the pilot testing, it is also possible to identify whether the questions are generating high item non-response, as well as providing estimates of the likely overall survey response rate [32].

Usability Testing

Usability is the concept of putting the question into the survey environment away from the artificial setting of question development, and includes issues such as the ability of the interviewer to ask the questions as intended and the ability of the respondent to answer the questions. There is a need to evaluate the 'usability' of the CATI instrument to ascertain how easy or difficult it is for both interviewers and respondents to interact and to determine the accuracy of responses. Usability testing focuses on the interviewer's interaction with the CATI system and survey instrument, shifting the focus from system feasibility and functionality to design of instruments from the user's perspective [45].

A number of additional design issues are introduced with CATI procedures, particularly the many ways in which the technology used affects the interviewers and respondents. The main focus of usability testing is directed toward the CATI software. However, there are major question design issues, particularly related to the interviewer/respondent interaction that can be resolved through usability testing of questions. In the CATI environment the interview flow is segmented, where the interviewer reads the question, gets an answer, enters the response, presses [enter] and moves to the next question, mostly on a new screen. With this segmented interaction it is sometimes difficult for interviewers to maintain consistent performance [39]. It should be noted that this is unlikely to be a major problem with comprehensive interviewer training and with familiarisation of both the computer system and survey itself.

Usability problems arise, in part, because the design of the user interface imposes both cognitive and physical demands that interviewers and respondents are sometimes unable or unwilling to meet [71]. With the use of usability testing of CATI questions, increased confidence can be gained that the delivery and reception of the questions will be in accordance with the original aims of the questions. In addition, question problems that are identified through usability testing can provide valuable input into what definitions, items and summaries are required in interviewer 'help screens' [48].

Reporting

The wide dissemination of the details of the question development process is required to ensure that extensive value can be gained from the resources used and advances that have been achieved. Not only does this support the input from the key informants but it facilitates the wide use of comparable questions across a number of surveys.

A report covering material that includes reporting on the qualitative testing, validity and reliability findings, should be publicly available. In addition, the combining of sets of questions into a comprehensive manual with support instructions and explanations for interviewers is recommended.

There may be value in including the developed questions in the National Health Data Dictionary (NHDD) where appropriate [9]. Although the evaluation processes required for inclusion in the NHDD are additional to the stages presented in this overview of question development principles and practices, there would be even broader utilisation of standard (comparable) questions, procedures and instructions across data collections once they were an integral part of the NHDD. This would assist in the comparison of findings from different survey collections.

Interviewers and Supervisors

The training of interviewers and supervisors is an area of survey quality that has received scant attention compared to other survey areas, such as sampling errors. One of the strengths of CATI is the centralised supervision and support for interviewers [16].

When examining the quality of responses to questions used across a number of surveys, Cannell & Oksenberg (1988) reported that in the worse situation nearly a quarter of interviewers made major wording changes and over half used unacceptable probes [20]. It is therefore necessary to support the question (or set of questions) with additional material that can be included in interviewer training. In addition, interviewers with less familiarity with computing technology require additional training to raise their confidence and precision in the CATI environment [80]. When there are a small number of interviewers, each having a large workload then poor performance from a single interviewer can have a major effect on the results of the collection. Improvements in ongoing supervision is particularly important to monitor each interviewer [16].

With a greater emphasis on question development there is generally a reduction in the amount of probing required by interviewers. However, there is a need for interviewer training to cover the concepts that are inherent in the questions so that interviewers can assess the completeness of respondents' answers and provide "non-directive probing that is based on a sound understanding of the question's objectives rather than the interviewers' own speculation of what constitutes a complete answer" [11].

It is important that questions are supported with instructions for interviewers, both as a part of the CATI software and as part of interviewer training. Interviewers need material, such as instructions and definitions, that they can draw upon to correctly respond to respondent queries or misunderstandings [43, 84]. For more complicated questions both scripted and para-phrased additional wording clarification increases the accuracy of responses, but also increases the length of interviews [85].

Voluntary informed consent implies that as part of the introduction to the survey the respondent is advised of the questions to be asked and given opportunities to ask for more information about the questions [88]. To assist in responding to these inquiries interviewers need relevant information both in training and readily at hand.

It is also beneficial to provide interviewers with a thorough orientation of the question (or set of questions) as well as seeking their involvement in the survey development process. “Interviewers jump at the opportunity to help refine a questionnaire” [70].

Although interviewer training is generally accepted as being essential either when the interviewer is first employed or at the introduction of a new data collection, there can be substantial value in scheduling regular training.

The ability to monitor interviews is one of the major features of CATI surveys, aiding in the improvement of costs, productivity and efficiency [25]. A large number of CATI systems provide the facility to have online monitoring of interviews with the system able to reproduce any interviewer’s screen at the supervisor’s terminal where audio monitoring can also be linked [74]. Under such settings, not only can the interviewer be monitored but also the performance of the question, in particular how well it is delivered and received by respondents, can be assessed. “When appropriately used, (monitoring) permits reinforcement of training guidelines, presumably leading to a reduction in interviewer-induced error” [25].

Additional Issues

Although some of the additional issues presented below do not relate directly to the development of questions, the issues do impact on the conduct and delivery of questions in the data collections.

Ethical and Related Issues

Ethical issues are fundamental in social research and are not just an add-on that is only addressed in the planning stages of a project. The design of questions needs to assume that participation is voluntary and that informed consent safeguards the rights of respondents. If there are important ethical issues associated with a question (or set of questions), appropriate additional material needs to be prepared and included in interviewers' instructions.

Concerns are being expressed regarding the intrusion of information collection, particularly via telephone. These relate mainly to the unethical practices of market research and tele-marketing, such as 'push-polls' (political propaganda disguised as legitimate polling, but using biased question wording solely to expose respondents to a highly partisan viewpoint), 'FRUGing' (fund-raising under the guise of surveying), and 'SUGing' (selling under the guise of surveying).

The interviewing of children and adolescents raises a number of issues such as consent of their parents/guardians plus their own assent to participate. In addition, a number of key factors relevant to interviewing children include:

1. they have limited psychological, as well as legal, capacity to give informed consent;
2. they may be cognitively, socially, and emotionally immature;
3. there are external constraints on their self-determination and independent decision making;
4. they have unequal power in relation to authorities, such as parents, teachers, and researchers;
5. their parents and certain institutions, as well as the youngsters themselves, have an interest in their participation; and
6. national priorities for research on children and adolescents include research on drug use, the problems of runaways, pregnancy among teenagers, and other sensitive topics, compounding the ethical and legal surrounding research on minors [88].

However, it would seem patronising to exclude children and adolescents from answering a range of questions on their own behalf, provided the questions have been appropriately designed and consent is obtained (as discussed above).

Core Questions

One issue associated with the development of sets of questions for a topic is the determination of which questions should be considered as part of a core component of an ongoing series of surveys and which questions can be considered as expanded/additional questions. To a major extent this depends on the priority

given to the topic and the requirement to monitor this topic. This determination would be made by the body responsible for the survey following consultation with appropriate reference groups.

Survey Coverage

Coverage of CATI surveys is a major concern, particularly when the characteristics of those not covered by the population being studied can be expected to be substantially different from those covered by the survey methodology. In particular, it has been shown that telephone ownership remains lower in households with young adult males, households with a non-English speaking background and those with low household income who live in public housing [4, 8, 92].

There are also key groups who are generally excluded from CATI health surveys, namely those who are:

- homeless;
- live in sparsely settled areas;
- in the armed forces,
- in long stay hospital, nursing homes, hostels;
- in prison/detention; and
- in other long-term communal accommodation.

It could be expected that for most of these groups, and for particular health issues, there could be substantial and important differences.

Survey Introduction

Although there has been little research conducted on the nature and extent of the survey introduction in the CATI environment, Meegama & Blair (1999) observe that “the goal of an effective introduction appears to be not only to provide minimum information, but the right kind of information – eliciting cooperation of the potential respondent and ... enhancing rapport during the interview” [67]. This supports the need for the inclusion of additional background information on the question (or set of questions) for both interviewer training and as part of the CATI help screens and definitions that are readily accessible at the beginning of the interview.

Order Effect

Care is required to ensure that the question (or set of questions) is placed in a logical position in the data collection and the order of response categories in a question follows a logical sequence. The questions (or set of questions) are part of a context with other questions, as well as part of a continuous flow of questions [38]. Accordingly, the context in which the question (or set of questions) is placed can affect an individual’s response, with the cognitive cues required to respond to prior questions influencing answers to subsequent questions [98]. In addition, there is also the response order effect where the order of response categories presented for a question influences the selection made by the respondent [52].

One particular concern arises when expanded questions are included in a survey, and what their subsequent impact might be on responses to later questions. It may be that the expanded questions should be included after core questions, with appropriate statements to explain to the respondent that some additional questions are being asked related to earlier questions.

Re-interviewing and Monitoring

As part of an ongoing assessment of the questions, there should be a component of re-interviews as part of the final survey. As mentioned earlier this is required for the establishment of reliability of the questions. It could be expected that some questions might need to be further developed and the incorporation of a re-interview component would provide an early warning whether the questions were in need of further development, for example with a particular population.

In large or ongoing surveys, question development and testing should be seen as a continuous and integrated process rather than as a set of independent activities [106]. To this end, the use of re-interviews to monitor the reliability of questions should be incorporated as an integral component of each round of data collection.

Bibliography

- 1 Aday LA. Designing and conducting health surveys. 2nd Edition. San Francisco: Jossey-Bass, 1996.
- 2 Akkerboom H, Dehue F. The Dutch model of data collection development for official surveys. *International Journal of Public Opinion Research* 1997; 9: 126-145.
- 3 Akkerboom H, Luiten A. Selecting pretesting tools according to a model of questionnaire development, with illustrations concerning patient satisfaction with medical care. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1996; 911-916.
- 4 Allister J, Brennan D, Davies M, Hoskin G, Roberts-Thomson K, Spencer J, Slade G, Stewart J, Szuster F. A research database on dental care in Australia. Adelaide: The University of Adelaide, 1993.
- 5 Australian Bureau of Statistics. The measurement of ethnicity in the Australian census of population and housing. Catalogue No. 2172.0. Canberra: ABS, 1984.
- 6 Australian Bureau of Statistics. 1995 national health survey data reference package. Canberra: ABS, 1996.
- 7 Australian Bureau of Statistics. An introduction to sample surveys: a user's guide. Catalogue No. 1299.0. Canberra: ABS, 1999.
- 8 Australian Bureau of Statistics. Population survey monitor. Unpublished tables, 2000.
- 9 Australian Institute of Health and Welfare. National Health Data Dictionary. Version 9. Catalogue No. HWI 24. Canberra: AIHW, 2000.
- 10 Baker T. (ed.) Key performance indicators manual. South Melbourne: Pearson Professional, 1995.
- 11 Beatty P. Understanding the standardized/non-standardized interviewing controversy. *Journal of Official Statistics* 1995; 11: 147-160.
- 12 Beatty P, Schechter S. An examination of mode effects in cognitive laboratory research. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1994; 1275-1280.
- 13 Beatty P, Schechter S, Whitaker K. Evaluating subjective health questions: cognitive and methodological investigations. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1996; 956-961.
- 14 Beatty P, Schechter S, Whitaker K. Variation in cognitive interviewer behavior – extent and consequences. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1997; 1064-1068.

- 15 Belson WA. The design and understanding of survey questions. London: Gower, 1981.
- 16 Bennett DJ, Steel D. An evaluation of a large-scale CATI household survey using random digit dialing. *Australian and New Zealand Journal of Statistics* 2000; 42: 255-270.
- 17 Bickman L, Rog DJ. (eds.) *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage, 1998
- 18 Biemer PP. Nonresponse bias and measurement bias in a comparison of face to face and telephone interviewing. *Journal of Official Statistics* 2001; 17: 295-320.
- 19 Bischooping K, Dykema J. Toward a social psychological programme for improving focus group methods of developing questionnaires. *Journal of Official Statistics* 1999; 15: 495-516.
- 20 Cannell C, Oksenberg L. Observation of behavior in telephone interviews. In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J. (eds.) *Telephone Survey Methodology*. NY: John Wiley & Sons, 1988.
- 21 Computer Assisted Telephone Interview Technical Reference Group. Report of the Module Development Project. National Public Health Partnership, 2000.
- 22 Conrad F, Blair J. From impressions to data: increasing the objectivity of cognitive interviews. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1996; 1-9.
- 23 Conrad FG, Schober MF. Clarifying question meaning in a household telephone survey. *Public Opinion Q* 2000; 64: 1-28.
- 24 Couper MP, Baker RP, Bethlehem J, Clark CZF, Martin J, Nicholls WL II, O'Reilly JM. (eds.) *Computer Assisted Survey Information Collection*. NY: John Wiley & Sons, 1998.
- 25 Couper MP, Holland L, Groves RM. Developing systematic procedures for monitoring in a centralized telephone facility. *Journal of Official Statistics* 1992; 8: 63-76.
- 26 Couper MP, Nicholls WL II. The history and development of computer assisted survey information collection methods. In: Couper MP, Baker RP, Bethlehem J, Clark CZF, Martin J, Nicholls WL II, O'Reilly JM. (eds.) *Computer Assisted Survey Information Collection*. NY: John Wiley & Sons, 1998.
- 27 Czaja R. Questionnaire pretesting comes of age. *Marketing Bulletin* 1998; 9: 52-66.
- 28 DeLeeuw ED. The effect of computer-assisted interviewing on data quality: a review of the evidence. Cologne: 5th Social Science Methodology Conference, 2000.

- 29 DeMaio TJ, Rothgeb J, Hess J. Improving survey quality through pretesting. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 50-58.
- 30 Deming WE. On errors in surveys. *American Sociological Review* 1944; 9: 359-369.
- 31 Desvousges WH, Frey JH. Integrating focus groups and surveys: examples from environmental risk studies. *Journal of Official Statistics* 1989; 5: 349-363.
- 32 Dillman DA. *Mail and Internet Surveys. The Tailored Design Method*. 2nd Edition. NY: John Wiley & Sons, 2000.
- 33 Eisenhower DL. Design-oriented focus groups and cognitive laboratories: a comparison. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1994; 1374-9.
- 34 Esposito JL, Campanelli PC, Rothgeb J, Polivka AE. Determining which questions are best: methodologies for evaluating survey questions. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1991; 46-55.
- 35 Forsyth BH, Hubbard ML. A method for identifying cognitive properties of survey items. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1992; 470-475.
- 36 Fowler FJ Jr. Design and evaluation of survey questions. In: Bickman L, Rog DJ. (eds.) *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage, 1998: 343-374.
- 37 Frankfort-Nachmias C, Nachmias D. *Research methods in the social sciences*. 4th Edition. London: St Martin's press, 1992.
- 38 Frey JH. *Survey research by telephone*. 2nd Edition. Newbury Park: Sage, 1989.
- 39 Fuchs M. Screen design and question order in a CAI instrument effects on interviewers and respondents. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1999; 1013-1018.
- 40 Gerber ER. Hidden assumptions: the use of vignettes in cognitive interviewing. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1994; 1269-1274.
- 41 Gower AR, Bélanger B, Williams M-J. Using focus groups with respondents and interviewers to evaluate the questionnaire and interviewing procedures after the survey has taken place. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 404-409.
- 42 Greenbaum TL. *The Handbook for Focus Group Research*. 2nd Edition. Thousand Oaks: Sage Publications, 1998.

- 43 Hansen SE. The impact of design on interaction in the computer assisted survey interview. Proceedings of the Section on Survey Research Methods, American Statistical Association 1999; 1019-1024.
- 44 Hansen SE, Couper MP, Fuchs M. Usability evaluation of the NHIS CAPI instrument. Proceedings of the Section on Survey Research Methods, American Statistical Association 1998; 928-933.
- 45 Hansen SE, Fuchs M, Couper MP. CAI instrument usability testing. Proceedings of the Section on Survey Research Methods, American Statistical Association 1997; 1023-1028.
- 46 Hilton JS, Bushery JM. Automating the focused reinterview. Proceedings of the Section on Survey Research Methods, American Statistical Association 1998; 421-424.
- 47 House CC. Questionnaire design with computer assisted telephone interviewing. Journal of Official Statistics 1985; 1: 209-219.
- 48 House CC, Nicholls WL II. Questionnaire design for CATI: design objectives and methods. In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J. (eds.) Telephone Survey Methodology. NY: John Wiley & Sons, 1988.
- 49 Humphries B. Unsettling ethics in social research. 5th Social Science Methodology Conference, Cologne, 2000.
- 50 Jabine TB. Flow charts: a tool for developing and understanding survey questionnaires. Journal of Official Statistics 1985; 1: 189-207.
- 51 Jobe JB, Mingay DJ. Cognition and survey measurement: history and overview. Applied Cognitive Psychology 1991; 5: 175-92.
- 52 Johnson T, O'Rourke D, Severns E. Effects of question context and response order on attitude questions. Proceedings of the Section on Survey Research Methods, American Statistical Association 1998; 857-860.
- 53 Katz I, Stinson L, Conrad F. Questionnaire designers versus instrument authors: bottlenecks in the development of computer-administered questionnaires. Proceedings of the Section on Survey Research Methods, American Statistical Association 1997; 1029-1034.
- 54 Kerr C, Taylor R, Heard G. (eds.) Handbook of Public Health Methods. Sydney: McGraw-Hill, 1998.
- 55 Kinsey SH, Jewell DM. A systematic approach to instrument development in CAI. In: Couper MP, Baker RP, Bethlehem J, Clark CZF, Martin J, Nicholls WL II, O'Reilly JM. (eds.) Computer Assisted Survey Information Collection. NY: John Wiley & Sons, 1998.
- 56 Khurshid A, Sahai H. A bibliography on telephone survey methodology. Journal of Official Statistics 1995; 11: 325-367.

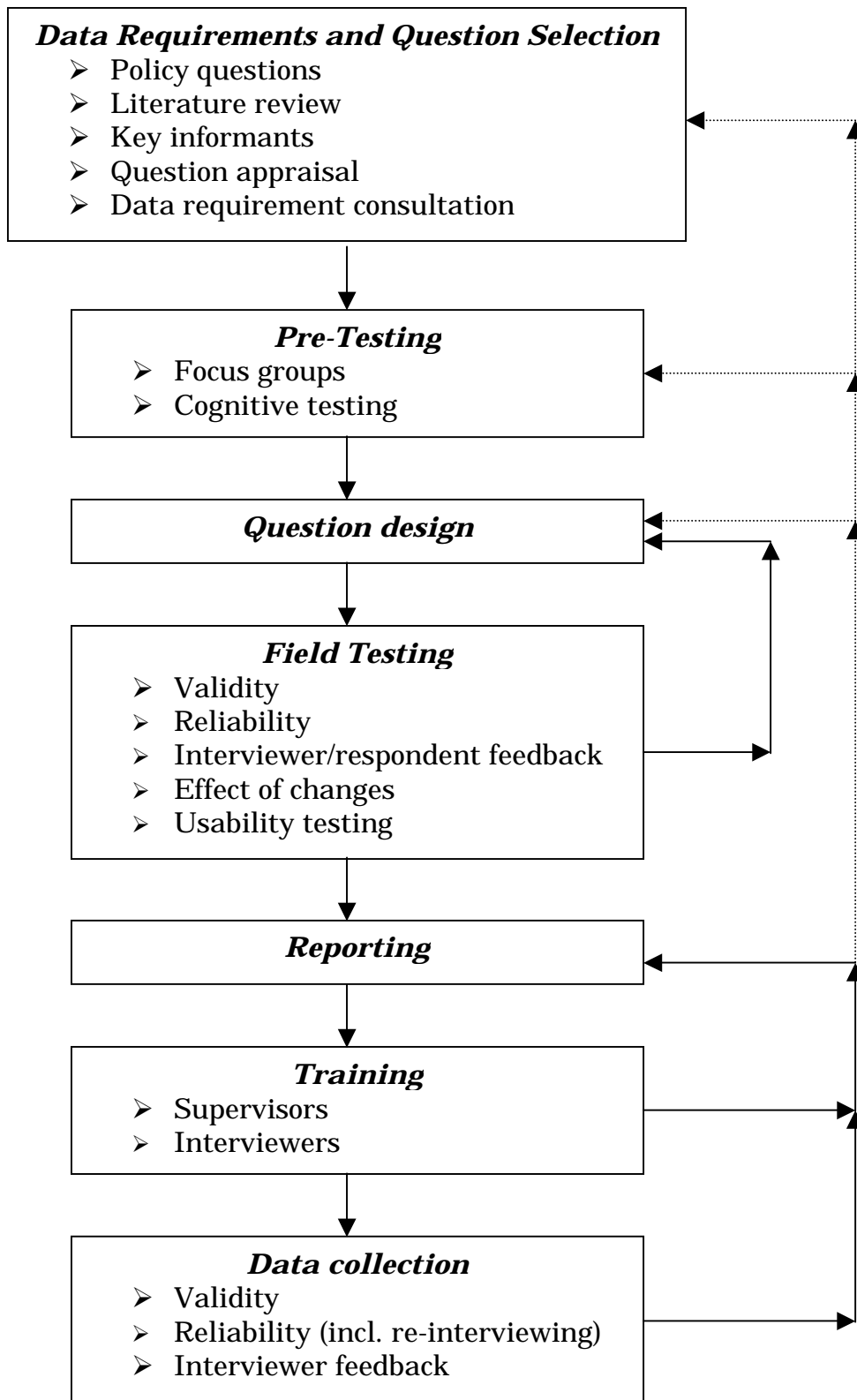
- 57 Krosnick J. Survey research. *Annual Review of Psychology* 1999; 50: 537-567
- 58 Krosnick J. The threat of satisficing in surveys: the shortcuts respondents take in answering questions. *Survey Methods Newsletter*. London: National Centre for Social Research 2000; 20 (1): 4-8.
- 59 Lavrakas PJ. Telephone survey methods. Sampling, selection, and supervision. Newbury Park: Sage, 1987.
- 60 Lavrakas PJ. Methods for sampling and interviewing in telephone surveys. In: Bickman L, Rog DJ. (eds.) *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage, 1998: 429-472.
- 61 Lepkowski JM, Couper MP, Hansen SE, Landers W, McGonagle KA, Schlegel J. CAPI instrument evaluation: behavior coding, trace files, and usability methods. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 917-922.
- 62 Lessler JT, Kalsbeek WD. *Nonsampling Error in Surveys*. NY: John Wiley & Sons, 1992.
- 63 Lindstrom K, Dallner M, Elo A-L, Gamberale F, Knardahl S, Skogstad A, Orhede E (eds.) *Review of psychological and social factors at work and suggestions for the general Nordic questionnaire*. Copenhagen: Nordic Council of Ministers, 1997.
- 64 Marquis K, Nichols E, Tedesco H. Human-computer interface usability in a survey organization: getting started at the Census Bureau. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 416-420.
- 65 Martin E, Polivka AE. Diagnostics for redesigning survey questionnaires. *Measuring work in the current population survey*. *Public Opinion Quarterly* 1995; 59: 547-567.
- 66 McDowell I, Newell C. *Measuring Health. A Guide to Rating Scales and Questionnaires*. 2nd Edition. NY: Oxford University Press, 1996
- 67 Meegama N, Blair J. The effects of telephone introductions on cooperation: an experimental comparison. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1999; 394-397.
- 68 Morgan DL. Focus groups. *Annual Review of Sociology* 1996; 22: 129-152.
- 69 Morgan DL, Krueger RA. *The Focus Group Kit*. Thousand Oaks: Sage Publications, 1998.
- 70 Mullin PA. Feasibility of using CAPI/CATI interviewers to conduct cognitive interviews during survey field tests. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1999; 38-43.

- 71 Murphy E, Marquis K, Hoffman R, Saner L, Tedesco H, Harris C, Roske-Hofstrand R. Improving electronic data collection and dissemination through usability testing. Paper SM00/01, US Bureau of the Census, 2000.
- 72 National Public Health Partnership. Preventing Chronic Disease: A Strategic Framework Background. Melbourne: NPHP, 2001.
- 73 Nelson DD. Informal testing as a means of questionnaire development. *Journal of Official Statistics* 1985; 1: 179-188.
- 74 Nicholls WL II. Computer-assisted telephone interviewing: a general introduction. In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J. (eds.) *Telephone Survey Methodology*. NY: John Wiley & Sons, 1988.
- 75 Oksenberg L, Cannell C, Kalton G. New strategies for pretesting survey questions. *Journal of Official Statistics* 1991; 7: 349-365.
- 76 Oppenheim AN. *Questionnaire Design and Attitude Measurement*. London, Heinemann, 1966.
- 77 Organisation for Economic Co-operation and Development. *The environmental goods and services industry. Manual for data collection and analysis*. Paris: OECD, 1999.
- 78 Payne SL. *The art of asking questions*. Princeton: Princeton University Press, 1951.
- 79 Platek R. Some important issues in questionnaire development. *Journal of Official Statistics* 1985; 1: 119-136.
- 80 Potter DEB, Sperry S, Edwards B, Dulaney R. Can information from and about interviewers predict their CAPI behavior? *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 987-991.
- 81 Rothgeb JM, Loomis LS, Hess JC. Challenges in gaining acceptance of the results of cognitive methodological research. Cologne: 5th Social Science Methodology Conference, 2000.
- 82 Schaeffer NC. A decade of questions. *Journal of Official Statistics* 1995; 11: 79-92.
- 83 Schechter S, Blair J, van de Hey J. Conducting cognitive interviews to test self-administered and telephone surveys: which methods should we use? *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1996; 10-17.
- 84 Schober MF, Conrad FG. Response accuracy when interviewers stray from standardization. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1998; 940-945.

- 85 Schober MF, Conrad FG, Fricker SS. When and how should survey interviewers clarify question meaning? Proceedings of the Section on Survey Research Methods, American Statistical Association 1999; 985-990.
- 86 Schuman H, Presser S. Questions and Answers in Attitude Surveys. Experiments on Question Form, Wording, and Context. NY: Academic Press, 1981.
- 87 Schwarz N, Strack F, Hippler H-J, Bishop G. The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology* 1991; 5: 193-212.
- 88 Sieber JE. Planning ethically responsible research. In: Bickman L, Rog DJ. (eds.) *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage, 1998: 127-156.
- 89 Sirken M, Herrmann D. Relationships between cognitive psychology and survey research. Proceedings of the Section on Survey Research Methods, American Statistical Association 1996; 245-249.
- 90 Snijkers G, Hox J & DeLeeuw ED. Interviewers' tactics for fighting survey nonresponse. *Journal of Official Statistics* 1999; 15: 185-198.
- 91 Starr GJ, Dal Grande E, Taylor AW, Wilson DH. Reliability of self-reported behavioural health risk factors in a South Australian telephone survey. *Australian and New Zealand Journal of Public Health* 1999; 23: 528-530.
- 92 Steel D, Boal P. Accessibility by telephone in Australia: implications for telephone surveys. *Journal of Official Statistics* 1988; 4: 285-297.
- 93 Steel D, Vella J, Harrington P. Quality issues in telephone surveys: coverage, non-response and quota sampling. *Australian Journal of Statistics* 1996; 38: 15-34.
- 94 Stewart AL, Hays RD, Ware JE Jr. Methods of validating MOS health measures. In: Stewart AL, Ware JE Jr. (eds.) *Measuring functioning and well-being*. Durham: Duke University Press, 1992.
- 95 Stewart DW, Shamdasani PN. Focus group research: exploration and discovery. In: Bickman L, Rog DJ. (eds.) *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage, 1998: 505-526.
- 96 Sykes W, Morton-Williams J. Evaluating survey questions. *Journal of Official Statistics* 1987; 3: 191-207.
- 97 Tanur JM, Fienberg SE. Cognitive aspects of surveys: yesterday, today, and tomorrow. *Journal of Official Statistics* 1992; 8: 5-17.
- 98 Todorov A. Context effects in national health surveys. *Public Opinion Quarterly* 2000; 64: 65-76.
- 99 Trewin D. How do we reduce non-sampling errors? *Journal of Official Statistics* 1987; 3: 343-347.

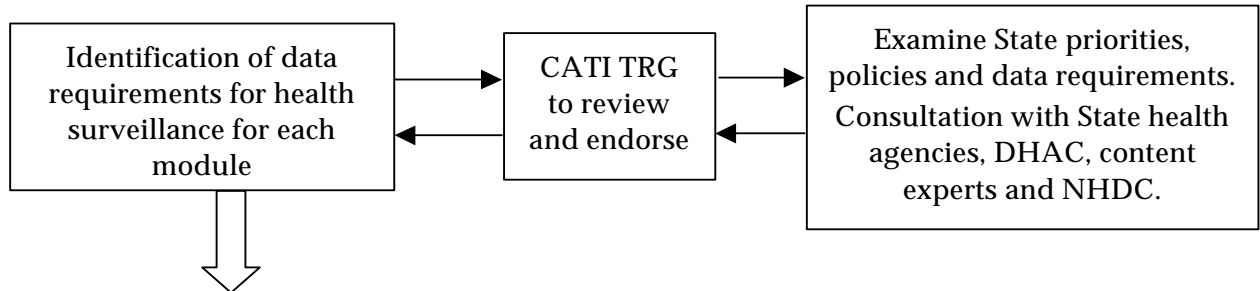
- 100 Trewin D, Lee G. International comparisons of telephone coverage. In: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls WL II, Waksberg J. (eds.) *Telephone Survey Methodology*. NY: John Wiley & Sons, 1988.
- 101 United States Census Bureau. *Pretesting policy and options: demographic surveys at the Census Bureau*. Washington, DC: US Department of Commerce, 1998.
- 102 Vogt WP. *Dictionary of Statistical and Methodology*. 2nd Edition. Thousand Oaks: Sage, 1999.
- 103 Weeks MF. Computer-assisted survey information collection: a review of CATI methods and their implications for survey operations. *Journal of Official Statistics* 1992; 8: 445-465.
- 104 Willis GB, Royston P, Bercini D. The use of verbal report methods in the development and testing of survey questionnaires. *Applied Cognitive Psychology* 1991; 5: 251-267.
- 105 Willis GB, Schechter S, Whitaker K. A comparison of cognitive interviewing, expert review, and behavior coding: what do they tell us? *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1999; 28-37.
- 106 Willis GB, Trunzo D, Stussman BJ. The use of novel pretesting methods in the development of survey questionnaires. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1992; 824-828.
- 107 Wilson BF, Peterson LS. Using the NCHS Cognitive Lab to help design Cycle VI of the National Survey of Family Growth. *Proceedings of the Section on Survey Research Methods, American Statistical Association* 1999; 997-1002.
- 108 Wilson DH, Walsh PG, Sanchez L, Davis AC, Taylor AW, Tucker G, Meagher I. The epidemiology of hearing impairment in an Australian adult population. *International Journal of Epidemiology* 1999; 28: 247-252.
- 109 Wojcik MS, Hunt E. Training field interviewers to use computers: past, present, and future trends. In: Couper MP, Baker RP, Bethlehem J, Clark CZF, Martin J, Nicholls WL II, O'Reilly JM. (eds.) *Computer Assisted Survey Information Collection*. NY: John Wiley & Sons, 1998.

Appendix A: Stages in question development

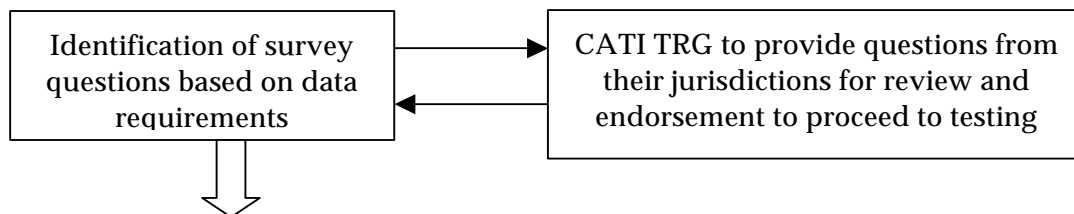


Appendix B: CATI module development consultation process

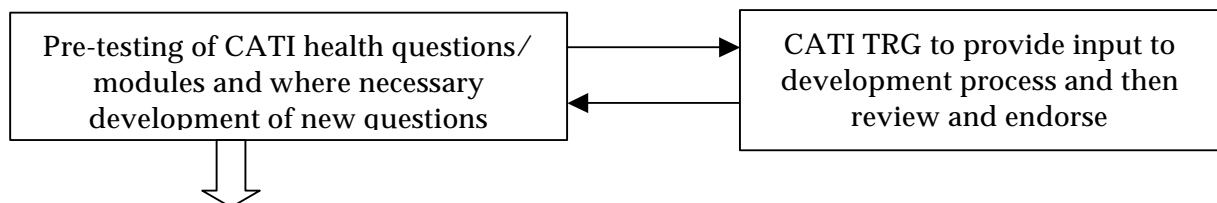
Stage 1 Data requirements for health surveillance



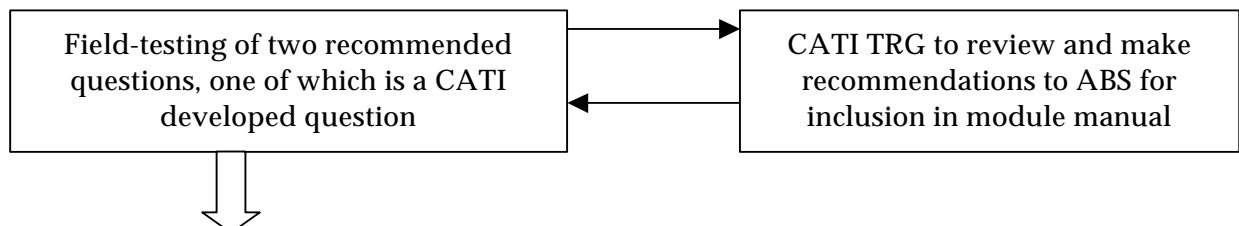
Stage 2 Identification of health surveillance/survey questions



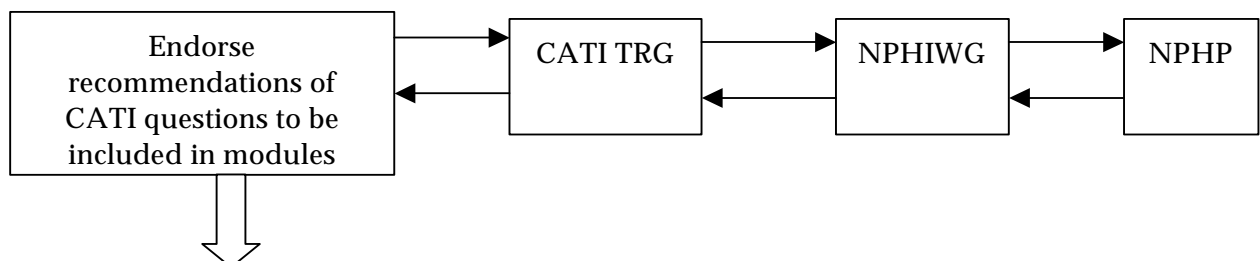
Stage 3 Pre-testing of CATI health surveillance modules



Stage 4 Field-testing of CATI health surveillance modules



Stage 5 Recommended health surveillance modules



Stage 6 Inclusion of modules in National Health Data Dictionary



Appendix C: Additional topic references

For the broader context of CATI question development within the conduct and quality of conducting health surveys, the following texts will provide further information:

- Aday LA. Designing and Conducting Health Surveys. [1]
- ABS. An Introduction to Sample Surveys, Cat. No. 1299.0. [7]
- Bickman L & Rog DJ. Handbook of Applied Social Research Methods. [17]
- Couper MP et al. (eds.) Computer Assisted Survey Information Collection. [23]
- Dillman DA. Mail and Internet Surveys. The Tailored Design Method. [31]
- Lessler JT & Kalsbeek WD. Nonsampling Error in Surveys. [61]
- McDowell I & Newell C. Measuring Health. A Guide to Rating Scales and Questionnaires. [65]

There are a number of resources that can be utilised in the gathering together of existing questions. In particular, there have been a number of State/Territory CATI Health Surveys that have covered a wide range of health topics, for example see:

- NSW <http://www.health.nsw.gov.au/public-health>
- Victoria <http://www.dhs.vic.gov.au/phd/0009074/index.htm>
- WA <http://www.health.wa.gov.au/publications/cwhs>
- SA <http://www.health.sa.gov.au/pehs/CPSE.html>
- Tasmania www.dhhs.tas.gov.au/services/healthy_living/pages/survey.html
- ACT www.health.act.gov.au/epidem/hps.html

The NHS questions can also be a valuable resource in the identification of existing questions. The questions used in the 1995 NHS are published in the ABS reference package [6].

There are also four particular international resources that provide extensive detail for health questions in the US, Canada and the UK, namely:

- BRFSS: the Behavioral Risk Factor Surveillance Survey conducted in each State of the US by the Centers for Disease Control and Prevention, see <http://www.cdc.gov/nccdphp/brfss>
- CCHS: the Canadian Community Health Survey conducted as part of the Canadian Roadmap Initiative, see <http://www.cihi.ca/Roadmap/rdindex.shtml> [Roadmap Initiative]
<http://www.statcan.ca/english/concepts/quest.htm> [Questionnaires and data dictionaries]
<http://www.statcan.ca/english/concepts/health/content.htm> [CCHS questionnaire]
- CASS: the Centre for Applied Social Survey in the UK, see <http://www.scpr.ac.uk/cass/docs/casshome.htm> [CASS]

<http://qb.soc.surrey.ac.uk> [the Question Bank]

- Health Poll Search: Presented by the Roper Center at the University of Connecticut and the Kaiser Family Foundation, this feature at the Kaiser website is a searchable archive of public opinion questions on health issues. The database contains about 40,000 selected questions dating back to 1935, which are drawn from the Roper Center's comprehensive database of more than 350,000 questions.

http://www.kaisernetwork.org/health_poll

The following texts will provide further information on focus groups, in particular, providing details on their conduct and the analysis of their results:

- Greenbaum TL. The Handbook for Focus Group Research. [41]
- Morgan DL. Focus groups. [67]
- Morgan DL, Krueger RA. The Focus Group Kit. [68]
- Stewart DW, Shamdasani PN. Focus group research: exploration and discovery. [94]

There are a large number of references that outline details of question design. The following will provide greater detail and direction:

- Belson WA. The design and understanding of survey questions. [15]
- Oppenheim AN. Questionnaire Design and Attitude Measurement. [75]
- Payne SL. The art of asking questions. [77]
- Schuman H, Presser S. Questions and Answers in Attitude Surveys. Experiments on Question Form, Wording, and Context. [85]

Appendix D: Stage One Template Example for the Collation of CATI Questions

Objective	Question Used	Source	Use of the information	Expected Information Outcome	Comments	Any results – basic demographic results
To identify if marital status may predispose people to health problems.	AA2 WHAT IS YOUR CURRENT MARITAL STATUS? (Tick one only) MARRIED DE FACTO SEPARATED BUT NOT DIVORCED DIVORCED WIDOWED NEVER MARRIED	WA Health Survey 1995	To see if marital status category predisposes people to health problems.	Certain life style behaviors were expected to be significantly related to marital status.	The results from this question were quite valuable in some of the lifestyle and health status interpretation	WA 1995 Married (50.9%) Defacto (4.2%) Separated (3.4%) Divorced (5.6%) Widowed (8%) Never Married (28%)

Source: CATI Technical Reference Group, 2000

Population Health Monitoring and Surveillance:

- **CATI Information Question and Module Development Principles and Practices**
- **Question Development Background Papers**
 - Alcohol Consumption in Australia
 - Asthma in Australia
 - Cardiovascular Disease in Australia
 - Demographic Characteristics
 - Diabetes in Australia
 - Nutritional Food Behaviour in Australia
 - Physical Activity in Australia
 - Tobacco Consumption in Australia

Information on the above papers can be obtained from:

Secretariat
National CATI TRG
c/- Population Health Data & Information Services Unit
Australian Institute of Health & Welfare
GPO Box 9848 (MDP 16)
CANBERRA ACT 2601

Tel: 02 6244 1000
Fax: 02 6244 1299
Email: info@aihw.gov.au



Computer
Assisted
Telephone
Interviewing
Technical
Reference
Group

www.nphp.gov.au/catitrg