

**SAFER • HEALTHIER • PEOPLE™**

# Working Paper Series

**COGNITIVE METHODS STAFF**  
Working Paper Series, No. 7

**Cognitive Interviewing and  
Questionnaire Design:  
A Training Manual**

Gordon B. Willis, Ph.D.

Office of Research and Methodology  
National Center for Health Statistics

March, 1994



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES  
Centers for Disease Control and Prevention  
National Center for Health Statistics

# **COGNITIVE INTERVIEWING AND QUESTIONNAIRE DESIGN: A TRAINING MANUAL**

**Gordon B. Willis, Ph.D.  
National Center for Health Statistics**

---

---

**OVERVIEW:** The methods described are used to test draft survey questionnaires by performing cognitive interviews of volunteer laboratory subjects. In these interviews, we focus on the mental processes subjects use when answering the survey questions. Understanding the nature of the response process helps us to re-design the questions, so that survey respondents will be better able to answer them accurately. Also discussed are practical aspects of operating a questionnaire design research laboratory.

---

---

# Contents

Preface and acknowledgements.....	ii
I. Introduction: Approaches to questionnaire design and development.....	1
II. Cognitive theory.....	4
III. Cognitive interviewing methods.....	6
Think-aloud techniques.....	6
Probing techniques.....	9
IV. Examples from cognitive interviewing.....	13
V. Detection of "non-cognitive" problems in survey questions.....	25
VI. Sequence of Laboratory Activities.....	26
VII. Practical aspects of laboratory operation.....	28
Interview length.....	28
Individuals who make good interviewers.....	28
Interviewer training.....	29
Other considerations for interviewing.....	30
Features of lab operation.....	31
VIII. What happens after the interview.....	34
IX. Limitations/criticisms of the laboratory approach.....	38
X. References.....	41
Appendix 1: Details of day-to-day lab operations.....	43
Appendix 2: Examples of newspaper ad/flyers.....	46
Appendix 3: Subject data sheet.....	48
Appendix 4: Examples of interviewing notes taken by a cognitive interviewer.....	49

Preface and acknowledgements. This document describes the cognitive interviewing techniques used in questionnaire development and testing by the staff of the Questionnaire Design Research Laboratory, National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, located in Hyattsville, Maryland. Typically, NCHS questionnaires are used in large-scale household surveys, in which considerable resources are devoted to development and pretesting activities (a survey may include 50,000 households, and the pretest itself may be over 300 household in size). Therefore, the activities described apply best to surveys that are large in scope, and involve significant developmental activities. However, in order to made the document as widely useful as possible, I have attempted to indicate how the techniques used can be modified and adapted for use on a smaller scale.

I do not purport to describe "the" way of conducting cognitive interviews, as there are probably points made that others who use these techniques would disagree with. However, QDRL staff members have often been asked for a written document that describes, in detail, procedures that we have found to be effective in addressing sources of response error in survey questionnaires. This document is intended to fulfill this request.

Many current and former NCHS staff members contributed to the production of this document. I would like to thank these individuals for their specific contributions, for their patient and careful review, and especially for their efforts in developing the techniques described. In particular, I thank Monroe Sirken, Susan Schechter, Jared Jobe, Douglas Herrmann, Paul Beatty, Mary Ann Guadagno, Barbara Wilson, Patricia Royston, and Deborah Trunzo.

# I. INTRODUCTION: APPROACHES TO QUESTIONNAIRE DESIGN AND DEVELOPMENT

## A) THE TRADITIONAL APPROACH TO QUESTIONNAIRE DEVELOPMENT: THE USE OF DESIGN RULES

Traditionally, survey questionnaires have been developed largely through the application of questionnaire design rules (for example, see Payne, 1951). For example, we are cautioned to avoid questions that are very long, that use difficult terms, that use biased phrasing, or that are double-barrelled (contain two-questions-in-one).

These rules are certainly helpful. However, there are limits to their effectiveness:

- 1) Rules are often not specific enough to be helpful when actually writing particular survey questions. For example, although design rules tell us to avoid questions that are very long, they do not determine how long a question should be. They also do not tell one whether a certain term will, in practice, actually be difficult to understand.
- 2) Rules are often not successful in predicting respondent ability to accurately answer survey questions. These abilities can vary markedly, depending on the population subgroup surveyed.
- 3) The fact remains that, even if design rules are applied stringently by experts, we still observe significant levels of response error in questionnaire-based data (Jabine, Straf, Tanur, and Tourangeau, 1984). The advances obtained in survey sampling precision have not been accompanied by similar gains in reduction of response errors; in fact, I suggest that elimination of this error stands as the major impediment to the improvement of data gathered by questionnaire.

## B) EMPIRICAL ("PRODUCT TESTING") APPROACHES TO QUESTIONNAIRE DESIGN

### 1) THE FIELD PRETEST:

Because we cannot always predict the problems that will be associated with particular survey questions, survey-takers have also relied on empirical, "product testing" techniques. Traditionally, we have used the *field pretest*, a small-scale survey conducted before the main survey, under conditions as similar as possible to those to be used in the full-scale implementation. Federal surveys such as the

National Health Interview Survey (NHIS) may involve a field pretest exceeding 300 cases.

However, there are also limitations to the field pretest approach:

- a) The field pretest tends to focus on the entire survey data collection process, rather than solely on the questionnaire itself. Thus attention to the questionnaire is diffused, as there are many other important issues that demand attention (such as formatting, issues related to interviewer training, adequacy of respondent rules, and so on).
- b) The pretest focuses on overt, rather than covert, problems; problems must be clearly observable to an interviewer or other observer, given the normal question asking-and-answering process.
- c) Field pretests often occur late in the development process; we often don't have time to make significant alterations afterward.

Because of limitations of the field test, survey designers and psychologists have developed another empirical approach to questionnaire testing:

## 2) COGNITIVE INTERVIEWING TECHNIQUES:

The cognitive interviewing approach was developed during the 1980's by survey methodologists and psychologists, and is the approach emphasized in the Questionnaire Design Research Laboratory (QDRL) at NCHS.

Some general features of this approach, as implemented in the QDRL, are as follows:

- a) It focuses mainly on the questionnaire instrument, rather than on the entire survey process.
- b) It pays explicit attention to the mental processes that respondents use to answer survey questions; therefore, covert as well as overt problems can be detected.
- c) It tests survey questions at multiple points in the design process.
- d) Generally, paid volunteer laboratory subjects are recruited, and are interviewed in a special setting devoted to the development of survey questionnaires (I use the term subject to refer to an individual who is tested in the laboratory, and respondent to mean someone who is interviewed in a fielded survey).

- e) Recruitment of laboratory subjects targets persons with specific characteristics of interest.

The general cognitive approach to questionnaire design has generated much research (see Jobe and Mingay, 1991, Jobe, Tourangeau, and Smith, 1993, and Lessler and Sirken, 1985), and in the Federal statistical system, is currently carried out not only at NCHS (Royston, Bercini, Sirken, and Mingay, 1986; Willis, Royston, and Bercini, 1991), but also at the Bureau of Labor Statistics (Dippo, 1989; Esposito, and Hess, 1992), and at the Census Bureau (Campanelli, Rothgeb, and Martin, 1989; Campanelli, Martin, and Rothgeb, 1991; DeMaio, Martin, and Sigman, 1987). A review of the work done at these agencies is contained in Jobe and Mingay (1991).

## II. COGNITIVE THEORY

I will first present a short introduction to the basic theoretical background to cognitive interviewing. The specific cognitive model I apply consists of several elements:

### 1) COMPREHENSION OF THE QUESTION:

- a) What does the respondent think the question is asking?
- b) What do specific words and phrases in the question mean to the respondent?

### 2) RETRIEVAL FROM MEMORY OF RELEVANT INFORMATION:

- a) What types of information does the respondent need to recall in order to answer the question?
- b) What type of strategy does the respondent rely on when retrieving information? For example, does the respondent tend to count events by recalling each one individually, or does he/she use an estimation strategy?

### 3) DECISION PROCESSES:

- a) Does the respondent devote enough mental effort to answer the question accurately and thoughtfully?
- b) Does the respondent want to tell the truth, or does he/she say something that makes him/her look "better"?

The question-answering process is clearly complex, and may involve a number of cognitive steps. These steps are not set in an invariant sequence, but may be analogous to computer "subroutines" that are implemented as necessary. Further, some of the processes may be "conscious", but some are automatic, so that the respondent is not aware of their operation. Further, the cognitive processes used to answer survey questions may vary, depending on the type of question asked. Autobiographical questions, for example, may place a heavy burden on retrieval processes. Asking questions that are sensitive, on the other hand (for example; "Have you ever smoked marijuana?"), may place more demands on the respondent's decision processes. The presentation above is very skeletal in nature; for more information on cognitive modelling of the survey response process, see Cannell, Miller, and Oksenberg (1981), Forsyth and Lessler (1991), Tourangeau and Rasinski (1988), and Willis, Royston, and



Bercini (1991).

Those of us who apply cognitive interviewing techniques recognize that we cannot know with certainty what is going on in someone's mind as he or she answers survey questions. Rather, our goal is to simply prompt the individual to tell us things that give us valuable clues about the types of processes mentioned above. The way we go about this is described in the next section.

### III. COGNITIVE INTERVIEWING METHODS

I have defined the basic approach as cognitive in nature, which simply means that it involves the study of mental information processing. The exact technique to be used can be selected by the individual or organization using it. The methods one might use, and what we feel to be the relative advantages and disadvantages of each, are discussed below. This discussion reflects our experience with each, and is meant largely as a set of issues for consideration, rather than strict rules; others who have used these techniques may have made observations that differ somewhat from ours.

#### A) "THINK-ALOUD" INTERVIEWING

##### 1) DESCRIPTION:

This has been the general term often used to describe the techniques used in the cognitive laboratory. We apply the term to describe a very specific type of activity, in which subjects are instructed to "think aloud" as they answer the survey questions. The interviewer reads each question to the subject, and then records and/or otherwise notes the processes the subject uses in arriving at an answer to the question. The interviewer interjects little else, except to say "tell me what you're thinking" when the subject pauses. For example, a portion of a think-aloud interview may consist of the following:

INTERVIEWER: How many times have you been to a dentist in the last 12 months?

SUBJECT: Well, let's see- I think I went one time for a general type of cleaning and checkup- that was just a month ago or so. I really can't remember any other times I would have gone- You said in the last 12 months, so that would take us back to July of last year- I'm not sure but around that time I went in for pain in my jaw that I was worried about, but they couldn't find anything. I'm not sure what month that was, though.

From this "think-aloud protocol", the interviewer observes that the individual answers this by trying to recall each visit individually, rather than by estimating. Further, the interviewer might conclude that the individual has trouble determining whether a visit was really in the last 12 months. If, after interviewing several subjects, it becomes clear that none could really "think through" with confidence the number of times they had been to the dentist, one might conclude that the reference period is simply too long to provide adequate answers.

## 2) TRAINING THE SUBJECT TO THINK-ALOUD:

The interviewer must be adept at teaching the subject how to perform the think-aloud procedure. This generally requires careful practice at the start of an interview, especially among individuals who are not especially verbally articulate.

One training approach that has been reported to work (by David Mingay, National Opinion Research Center), is the following:

"Try to visualize the place where you live, and think about how many windows there are in that place. As you count up the windows, tell me what you are seeing and thinking about."

Depending on how well the subject responds to this task, other exercises may be needed prior to beginning the interview.

## 3) ADVANTAGES OF THE THINK-ALOUD TECHNIQUE:

- a) *Freedom from interviewer-imposed bias:* Because the interviewer contributes little other than the reading of the survey question, he or she interjects little else that may serve to bias responses from the subject.
- b) *Low burden on interviewer:* The interviewer can spend most of the time simply listening to the respondent talk, and take notes.
- c) *Open-format:* Because the subject is not guided in any way, he or she may elicit information that is completely unanticipated by the interviewer. Think-aloud interviewing is especially valuable when the subject is outgoing, articulate, and has had significant experience with the topics covered by the questionnaire.

## 4) DISADVANTAGES OF THE THINK-ALOUD TECHNIQUE:

- a) *Artificiality:* Some lab subjects find this to be a very unnatural activity, (that is, even more artificial than is answering survey questions).
- b) *Training:* Because the activity is somewhat unusual, the technique can require a significant amount of training of lab subjects. The time this requires can substantially reduce the time one has to actually test survey questions.
- c) *Subject limitations:* Many of our subjects are not at all adept at the think-aloud activity: We have found that, even with training, some subjects tend to simply answer the question that is asked.

d) *Subject burden*: The think-aloud activity places the main burden on the subject. We prefer to place more of the relative burden on the interviewer.

e) *Staying on task*: With think-aloud, the subject controls the nature of much of the discourse. Therefore, it is very easy for a "free associating" subject to get completely off-track, and spend a significant amount of unproductive time on one question. In general, we find that using the think-aloud technique results in fewer survey questions being tested within a particular amount of time.

f) *Depth of subject information processing*: In a truly successful think-aloud, subjects are investing a good deal of mental effort into their answers (it's much harder to justify an answer through the think-aloud process than to simply say "yes" or "no"). This additional processing may bias the subject toward more careful responding than is the case in a usual survey interview. Therefore, it is possible that the act of producing "think-aloud" speech may serve to burden or contaminate the cognitive processes that are used in answering the question. This issue is clearly still open to debate, however.

## B) A SECOND COGNITIVE METHOD: USE OF PROBING TECHNIQUES

### 1) DESCRIPTION:

The use of **probing** is the basic technique favored at the QDRL. After the interviewer asks the survey question, and usually after the subject has answered it, the interviewer asks for other, specific information of the subject (we "probe" further into the basis for the response).

The following examples demonstrate a number of different "cognitive probes":

Examples:

<i>Comprehension/ Interpretation probe:</i>	What does the term "dental sealant" mean to you?
<i>Paraphrase:</i>	Can you repeat the question in your own words?
<i>Confidence judgment:</i>	How sure are you that your health insurance covers...
<i>Recall probe:</i>	How do you remember that you went to the dentist 3 times?
<i>Specific probe:</i>	Why do you think that cancer is the most serious health problem?
<i>General probes:</i>	How did you arrive at that answer?  Was that easy or hard to answer?  I noticed that you hesitated before you answered - what were you thinking about?

### 2) ADVANTAGES OF THE PROBING TECHNIQUE:

a) *Control of interview.* The use of probing to guide the interview tailors the interchange in a way that is controlled by the interviewer. This avoids a good deal of think-aloud that may be irrelevant and non-productive. Further, the interviewer can focus on particular areas that appear to be sources of potential response error.

b) *Subject training.* It is fairly easy to get subjects into the spirit of answering probe questions. In fact, subjects will often begin to expect probes and to offer their own spontaneous thoughts and critiques, an activity that becomes somewhat similar to what is desired in thinking-aloud.

### 3) DISADVANTAGES OF PROBING TECHNIQUES:

a) *Artificiality.* Some researchers feel that use of probing is inappropriate because the interjection of probes by interviewers produces a situation quite unlike the normal survey interview (although it is certainly no more so than the think-aloud). We find this not to be a problem, though, because the purpose is very different (to analyze questions, rather than to collect data). Further, one can counter this "unreal" element by relying less on probes during later stages of questionnaire testing, in order to produce an interview situation that is more realistic. Alternatively, one can make use of retrospective probing (see below).

b) *Bias.* A related criticism is that the use of probes may "lead" the respondent to particular types of responses. This is possible, but can be minimized by the careful use of "non-leading" probing techniques. Note that other professionals who attempt to extract information from individuals, such as police detectives, do not refrain from carrying out activities similar to probing, but are expected to do so in a relatively unbiased manner. For example, in conducting probing, rather than suggesting to the subject one possibility ("Did you think the question was asking X"), it is preferable to list all reasonable possibilities ("Did you think the question was asking X, Y, or Z"). In other words, we try to use unbiased phrasing, in the same manner that one does when writing survey questions in the first place.

### 4) WHEN SHOULD PROBES BE USED? - CONCURRENT VERSUS RETROSPECTIVE PROBING:

Probes can be asked at any time during the interview. Two general approaches are: a) *concurrent probing*, and b) *retrospective probing*. In concurrent probing, the probes are asked at the time the subject answers the questions. In retrospective probing, the subject is asked the probe questions during a debriefing session, after the entire interview has been administered. We generally rely on concurrent probing, mainly because the information is fresh in the subject's mind at the time of the probing. It may seem more "realistic" to wait and to debrief the subject by probing after the questions have been administered, but we find that there is a significant danger of creating a situation in which the person can no longer remember the basis for an answer, and instead fabricates an

explanation. The times that retrospective probing appears to be useful are:

- a) *When testing self-administered questionnaires.* Retrospective probing is useful when the purpose of testing is mainly to determine the subject's ability to complete the instrument unaided, and especially to follow sequencing instructions. We have probed "after the fact" on short, self-administered questionnaires of both teenagers and adults, with apparent success.
- b) *When in later stages of questionnaire development.* When a questionnaire is in later stages of development, and one wants to simulate a more "realistic" type of presentation, it makes sense to administer the questionnaire "straight", and to perform some probing afterward.

##### 5) HOW ARE THE SPECIFIC PROBES DEVELOPED?

There are two basic ways to develop and use probe questions:

- a) *Standard probes:* For use by all interviewers; these are developed prior to the interview.
- b) *Individual probes:* Used by a particular interviewer; these are developed either prior to, or during, the interview.

Standard probes are meant for use by all interviewers, and are developed prior to the interview by either a questionnaire development group or by a lead individual. For example, if it is anticipated that a particular term may not be universally understood, all interviewers can be instructed to apply the probe: "What does (TERM) mean to you?" These probes are often typed directly into the questionnaire draft.

Standard probes are practical and useful when:

- a) there is sufficient time to prepare for interviews.
- b) resources exist to plan and execute a fairly standardized testing approach.
- c) Some interviewers are relatively inexperienced and would benefit from the guidance provided by a structured protocol.

Individual probes consist of:

- i) **Prepared probes:** These are developed prior to the interview for personal use, and are based on an individual review of the questionnaire.

ii) **Spontaneous probes:** Probes that the interviewer develops during the course of the interview, based on either the subject's verbal report or non-verbal behavior. For example, a subject may answer a question, but then hesitate, which should prompt a probe from the interviewer. Or, the interviewer may think of a potentially useful probe "on the spot".

Individual probes are useful, and most practical, when there is very limited time to prepare for interviews (if, for example, a draft becomes available only a short time before a scheduled interview).

Prepared versus Spontaneous probes. Admittedly, the "spontaneous" approach to probing has the look of being "unscientific" or haphazard, especially because there is no coordination of probing across interviewers. However, there are particular advantages to this approach. We have found that the most interesting forms of probing often develop through the course of the interview, as a product of the particular relationship between the interviewer, subject, and survey instrument. These developments cannot be anticipated in advance of the interview. We have also found that, over time, interviewers become very adept at the use of this type of spontaneous probing. Further, the answer to a particular probe may well lead the interviewer to use other probes, and to follow-up on what issues emerge as the most interesting and important.

Combination of probe types. We have found that the most effective interviews actually consist of a combination of prepared and spontaneous probes described above, rather than either type alone. By way of analogy, we believe that the cognitive interview is similar to a session within a clinical psychological practice; the "therapist" has certain guiding principles, and perhaps specific questions or comments, to apply during a session with the patient. However, much of the interchange emerges spontaneously during the course of therapy. The clinical session may be approached in ways similar to other sessions, and be somewhat "scripted", but every interview is different, entails its own developmental sequence, and makes a unique contribution.

Again, because we focus on probing rather than the think-aloud at the NCHS lab, this will be the focus reflected in the remainder of the manual.



## IV. EXAMPLES FROM COGNITIVE INTERVIEWING

In order to make the above discussion of probing techniques more "concrete", I have assembled the following examples of questions that have been tested in the NCHS Questionnaire Design Research Laboratory. For each question, I have listed:

- 1) The question in its original form.
- 2) A list of several probes that would be appropriate to use in testing that question.
- 3) A short description of the problems we found, in testing these questions, using probes of the types suggested. Each of the examples is classed generally according to whether the problems found are representative of the cognitive categories defined in Section 2 (comprehension, retrieval, decision). Some questions may have more than one type of problem, and in some cases it is arguable what class of problem is really being reflected. This type of ambiguity is not harmful, as long as it is clear how to resolve the problem in order to produce a better "next generation" survey question.
- 4) Finally, a suggested resolution to the problem is presented, based on our testing results.

**EXAMPLE 1:****1) Original form of survey question:**

**Has anyone in the household ever received vocational rehabilitation services from-**

**... The State Vocational Rehabilitation program?**

**... another vocational rehabilitation program?**

**2) Probes:**

- a) Can you repeat the question in your own words?  
(To test how well the subject comprehends the question.)
- b) What, to you, is a "vocational rehabilitation program"?  
(To test comprehension of a particular term.)
- c) How sure are you that (person) got this type of service?  
(To determine the subject's ability to recall information confidently.)

**3) Results:**

Comprehension problems: Subjects found it difficult to understand the question, because of its length and technical nature. Further, the meaning of "vocational rehabilitation" was not at all clear; some subjects thought this just meant any type of physical therapy. Because of the comprehension problems in the original form, we suggested the following change:

**4) Suggested revision:**

**Has anyone in the household ever received job rehabilitation services?**

**If YES, ask WHO, and:**

**Was (person's) rehabilitation from the state, or from another job rehabilitation program?**

Note: The question is "decomposed", or divided up, to make it easier to understand. The term "vocational" is also changed to the more understandable form "job".

**EXAMPLE 2:****1) Original form:**

How long has (name) used the (cane, wheelchair, walker...)?

**2) Probes:**

- a) How did you get the answer of (x) years?  
(To determine the overall cognitive strategy used.)
- b) When did (x) first use the (device)?  
(To test comprehension/interpretation of the question.)
- c) How well do you remember this?  
(To test recall of the relevant information.)

**3) Results:**

We found that for target individuals whose use was intermittent over a long period of time, the question was interpreted in two distinctly different ways:

- 1) "How long has it been since (person) first used the (device)? For example, the subject may say: "since 1960, so about 30 years".
- 2) "For how long, overall, has (person) actually used the device since first having it? The subject counts up periods of use within a longer time- for example: "For two five-year periods since 1960, so 10 years".

Note that the problem identified can be considered a type of "comprehension" problem, but doesn't involve a failure of comprehension of a key term, as did the last example. Rather, subjects simply have alternate, but reasonable, interpretations of the question intent.

**4) Suggested revision:**

This required consultation with the client, in order to clarify the objective of the question. It became clear that the desired expression was:

How long ago did (person) first use a (device)?

**EXAMPLE 3:****1) Original form:**

**In the last year have you been bothered by pain in the abdomen?**

**2) Probes:**

- a) Why do you say (no)(yes)? [ General probe ]
- b) What, to you, is your abdomen?  
[ Comprehension/Interpretation probe ]
- d) What does it mean to be "bothered by pain in the abdomen?"  
[ Comprehension/Interpretation probe ]
- c) What period of time are you thinking of here, specifically?  
[ Comprehension/Interpretation probe ]

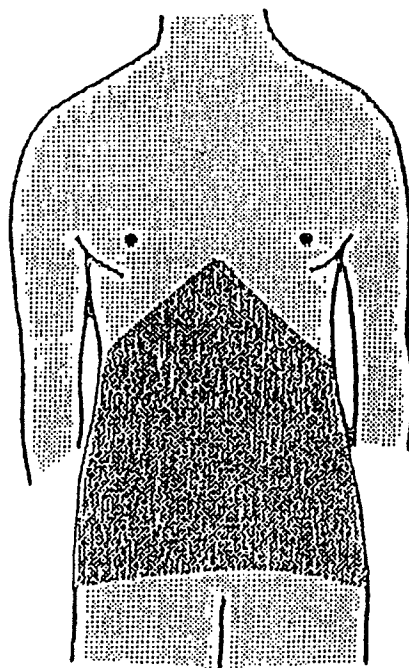
**3) Results:**

- a) To some people, "the last year" means Jan. 1 to the present. Others think that "the last year" started 365 days ago. Comprehension of this phrase is therefore somewhat varied.
- b) The phrase "bothered by pain ..." was also interpreted differently by different people. Some subjects experienced pain, but were not particularly "bothered" by it, and so responded with "No". Given that the purpose of the question was to measure the experience of pain, and not the psychological reaction to it, we saw this as a defect.
- c) Finally, as one might suspect, notions of where one's abdomen is differ greatly among individuals (almost no one chose the correct anatomical region). Again, problems related to comprehension of a key term were revealed here.

#### 4) Suggested revision:

Because the problems related mainly to comprehension, the solution we proposed focused on this process as well.

In the past 12 months, have you had pain in the abdomen?  
By abdomen, we mean the region shaded on this chart  
(show hand card).



**EXAMPLE 4:****1) Original form:**

About how many miles from here is the home (child) lived in before (he/she) moved to this home?

(the following is printed on the questionnaire, but not read):

- ☐ less than 1 mile
- ☐ 1-50 miles
- ☐ 50+ miles

**2) Probes:**

a) How sure are you of your answer?  
(to determine overall level of confidence)

b) How hard was this to answer?  
(to determine level of difficulty, and likelihood of estimation/guessing)

**3) Results:**

No one had difficulty understanding the question as posed. However, we observed that some subjects needed to think for a fairly long time before giving an answer. Further, some subjects struggled needlessly with the level of specificity they thought was required (for example, deciding whether the distance was closer to 20 or to 25 miles, when this information was ultimately irrelevant, as the interviewer would mark "1-50 miles" in either case).

The problem can be described as one involving a difficult recall task, as opposed to comprehension. A rephrasing of the question that incorporated response alternatives was necessary to make clear to subjects the degree of precision that was necessary in their answer.

**4) Suggested revision:**

About how far from here is the home \_\_\_\_\_ lived in before (he/she) moved to this home- less than a mile, 1 to 50 miles, or more than 50 miles?

**EXAMPLE 5:****1) Original form:**

**We are interested in your lifetime exercise patterns.  
First, when you were 14 to 19 years old:**

**How many hours a week of brisk walking did you do?**

**How many hours a week of vigorous exercise such as  
running, cycling, swimming, or aerobics did you do?**

**How many hours a week of activities that required you  
to be on your feet (excluding running or walking) such  
as dancing, hiking, .... did you do?**

**2) Probes:**

- a) Was this hard or easy to answer?  
(to determine comprehension, and overall ability to recall)
- b) How do you remember this?  
(to study recall strategy)
- c) How sure are you of your answer?  
(confidence probe)
- d) What, to you, is "vigorous exercise?"  
(comprehension/interpretation of a specific term)

**3) Results:**

Subjects found it very difficult to remember back to the time period specified, at the required level of detail. In fact, it seemed that some subjects really could not even answer this with respect to their current behavior, let alone their behavior many years ago. Recall of information (assuming it was ever "learned" in the first place) seemed to be the dominant problem.

As for the previous example, we needed to confer with the sponsor/client to clarify question objectives. We were able to determine that use of a broad scale of level of activity, comparing past and present behavior, would satisfy the data objectives:

## 4) Suggested revision:

We are interested in your lifetime exercise patterns. When you were 14 to 19 years old, were you more active than you are now, less active than now, or about as active as now?



## EXAMPLE 6:

## 1) Original:

During a typical work day at your job as an (occupation) for (employer), how much time do you spend doing strenuous physical activities such as lifting, pushing, or pulling?

- \_\_\_ None
- \_\_\_ Less than 1 hour
- \_\_\_ 1-4 hours
- \_\_\_ 4+ hours

## 2) Probes:

- a) What type of work do you do? Describe a typical workday.
- b) How did you arrive at the answer of X hours?

## 3) Results:

Careful probing revealed that people who gave reports of 1-4 hours often were office workers who did little or no heavy physical work. This appeared to be due to biasing characteristics of the question; saying "none" makes one appear to be "non-physical", and is therefore somewhat socially undesirable. This problem was seen as related to respondent decision processes, rather than to comprehension or recall. A resolution was needed to make it "easier" for someone to report little work-related physical activity:

## 4) Suggested revision:

The next questions are about your job as a \_\_\_\_ for \_\_\_\_.

Does your job require you to do repeated strenuous physical activities such as lifting, pushing, or pulling heavy objects?

(IF YES:)

During a typical work day, how many minutes or hours altogether do you spend doing strenuous physical activities?

**EXAMPLE 7:****1) Original:**

<b>Do you believe that prolonged exposure to high levels of radon gas can cause:</b>			
	<b>YES</b>	<b>NO</b>	<b>Don't Know</b>
<b>Headaches?</b>	—	—	—
<b>Asthma?</b>	—	—	—
<b>Arthritis?</b>	—	—	—
<b>Lung Cancer?</b>	—	—	—
<b>Other cancers?</b>	—	—	—

**2) Probes:**

- a) Why do you believe this?
- b) How sure are you of this?
- c) Is it difficult to answer these?

**3) Results:**

Simple observation of subjects made it clear that this question is difficult to answer. Subjects took a long time to respond to each item, and tended to be unsure about several of the items. Further, probing revealed that the format encouraged a "guessing" strategy, rather than actual retrieval of information. Finally, for people who do not believe that exposure to radon is harmful, it became very tedious, and sometimes even offensive, to repeatedly ask about the specific harmful effects of radon.

We felt that in this case, the subject's decision processes were again excessively burdened by the phrasing of the question.

## 4) Suggested revision:

Do you believe that prolonged exposure to radon is unhealthy, or do you believe that it has little or no effect on health?

(IF radon believed unhealthy:)

[ SHOW HAND CARD ] Which, if any, of these conditions do you believe can be caused by radon exposure?

<input type="checkbox"/> Headaches	<input type="checkbox"/> Lung cancer
<input type="checkbox"/> Asthma	<input type="checkbox"/> Other cancers
<input type="checkbox"/> Arthritis	<input type="checkbox"/> Don't Know

The revised phrasing gives the respondent a way to say, one time, that he or she does not believe that radon is harmful. Then, if he/she does believe it to be harmful, the next question simply allows him/her to "pick and choose" the items that seem appropriate. The burden on decision processes appeared to be reduced, using this alternative.

**EXAMPLE 8:****1) Original:**

What is the primary reason you have not tested your home for radon?

**2) Probes:**

- a) Is it hard to think of the main reason?
- b) Can you think of any other reasons?
- c) How much have you thought about having your home tested?

**3) Results:**

Although the question is easily enough understood, it was very difficult for subjects to produce a reasonable answer, especially for subjects who had never given the issue much thought. Instead of simply saying "I never thought about it", or "I haven't gotten around to it", subjects tried to think of more "appropriate" answers, that appear to be more defensible. Here both recall and decision processes appeared to be operating.

**4) Suggested solution:     --- DELETE QUESTION ---**

The sponsor/client agreed that it was not especially useful to ask the reason that someone had not carried out this activity.

This example demonstrates an important point worth emphasizing; sometimes, there is no obvious "correction" to a survey question. Especially when subjects simply don't have information that we want, it is better to acknowledge that we may not want to ask that question. Thus, one effect of lab testing is to test the boundaries of "what can be asked and what can't."

## V. DETECTION OF "NON-COGNITIVE" PROBLEMS IN SURVEY QUESTIONS

The discussion above has focused almost completely on cognitive problems in questionnaires; that is, problems involving the comprehension, recall, or decision processes necessary to adequately answer the question. However, we have found that lab interviewing has several overall positive effects, in addition to the understanding of particular cognitive processes:

- a) We explore the nature of the underlying concepts to be measured in the survey, and the specific subject matter (by relying on lab subjects as "experts").
- b) We learn about non-cognitive defects in the questionnaire.

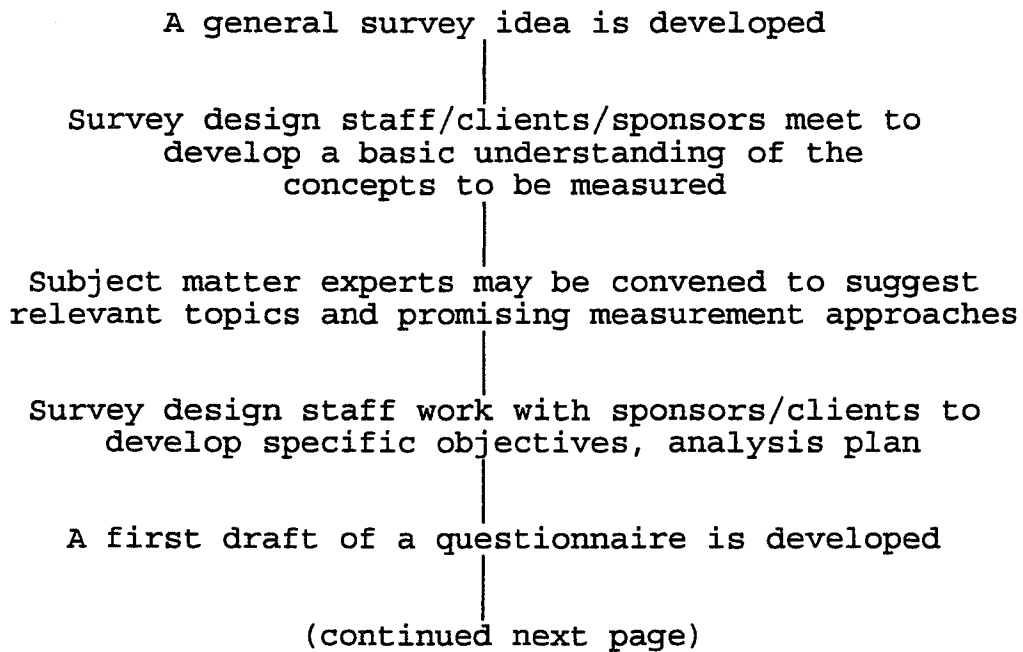
Non-cognitive problems. Item b) is especially worthy of clarification. We have found that an important, beneficial effect of lab testing is to detect structural, or logical problems, not normally viewed as a part of the cognitive approach. By structural problems, we mean those features of the questionnaire such as erroneous skip patterns, unclear layout, and other elements that do not clearly involve the cognitive processes of the respondent. We also include in this category logical problems inherent in the statement of the question that do not clearly involve a cognitive element. For example, if I ask the question: "How long have you owned your house?" the subject may simply respond that he is a renter. Here, it should not be strictly necessary to study cognitive processes to make the discovery that the question is flawed, because simple knowledge of the appropriate logical relationships ("some people own, some people rent") should have been sufficient to avoid such a problem. However, survey designers often fail to take into account all of these logical truths when constructing a questionnaire, and the laboratory-based interview allows the subject spontaneously point out flaws.

It is of course true that most of the structural, non-cognitive problems I have referred to could be detected by either a careful expert review, or in the field pretest. However, from a practical point of view, the expert review may never get done, and it can be imperfect. The field pretest generally occurs late in the process; it is much better to detect the problems earlier rather than later, and the lab serves this purpose well. Note that it takes no special "techniques" to detect the types of problems mentioned above, beyond simply attending to the possibility that they can occur.

## VI. SEQUENCE OF LABORATORY ACTIVITIES

The discussion above has provided the background necessary to understand the main techniques used in the NCHS Questionnaire Laboratory. The following sections will place these techniques into the broader context of actually conducting this testing within a "real-life" survey development process. To appreciate the overall picture regarding this process, it is useful to first consider an overview of the general sequence of events that occurs when a questionnaire is designed. I provide below a schematic diagram of some basic steps that can be taken to develop a questionnaire, incorporating laboratory cognitive interviewing techniques into the developmental sequence:

### INITIAL QUESTIONNAIRE DEVELOPMENT\*



\*See Aday (1989) for a more complete description of these preliminary steps

## LAB TESTING PHASE

Laboratory staff determine staff allocation to:

- a) recruitment,
- b) conducting of interviews
- c) communication with clients/sponsors

### PREPARATION FOR INTERVIEWING

Lab staff perform an initial review of the questionnaire, and may make suggestions for modifications prior to testing

Staff develop basic probes to use in first round of interviewing

### RECRUITMENT

Lab staff determine the types of subjects to be interviewed

A recruitment strategy is developed and implemented (advertisements, flyers, ...)

Subjects are recruited, scheduled

### LABORATORY INTERVIEWING ROUNDS

> Individual interviews are conducted with 5-10 subjects  
(In early rounds, emphasis is on general concepts)  
(In later rounds, emphasis is on question wording, format)

Interviewers write up comments on each interview done

Interviewers meet (as a group, or with sponsors/clients) to discuss results, propose potential modifications

The questionnaire is revised by designers

< A decision is made concerning whether another round of interviewing can/should be conducted

**YES**

**NO**

## FIELD PRETESTING PHASE

OPTIONAL: Return questionnaire to lab for additional round(s) of laboratory testing

**SURVEY IS ADMINISTERED**

## VII. PRACTICAL ASPECTS OF LABORATORY INTERVIEWING

Because this manual focuses mainly on the *Lab Testing Phase* of development, the remainder of the discussion will concern activities relevant to that phase. I will therefore address, in turn, several basic issues that we are frequently asked about with respect to conducting laboratory cognitive interviews.

### A) HOW LONG SHOULD A LABORATORY INTERVIEW BE?

We have found that one-hour interviews are optimal; longer interviews make excessive demands on laboratory subjects. In general, we recommend that the interview process be as flexible as possible, and not require interviewers to cover a certain number of pages of a questionnaire. Questionnaires often have skip patterns that result in widely varying actual questionnaire lengths for different individuals, and subjects vary in their overall speed and the degree to which they respond in detailed ways to either the survey questions, or to probe questions.

Note that even though the interview itself may take only an hour, the interviewing process requires considerably more time. In all, preparation, interviewing, and writing up results of the interview usually take a total of about three hours. Because of this, and because cognitive interviewing can be a taxing activity, we recommend that any individual do no more than three interviews in a single day (and preferably fewer).

### B) WHAT TYPES OF INDIVIDUALS MAKE GOOD INTERVIEWERS?

It is unnecessary to have an advanced degree in psychology to be a good cognitive interviewer (although a behavioral sciences background appears to be helpful). We have found that good interviewers are those people who:

- a) Have experience in questionnaire design, and are knowledgeable about both survey design and about the purpose of the questionnaire to be tested. These skills are essential when the time comes to apply the results of the interviews in revising the questionnaire.
- b) Have learned the basic premises of cognitive interviewing, and are familiar with the ways in which fundamental cognitive processes may influence the survey response.



c) Have been exposed to social science research concepts such as bias, context effects, scale effects, and so on.

d) Perhaps most importantly, have good inter-personal skills, are capable of putting a subject at ease, and remaining non-judgmental in approach. There is still an open question, however, concerning how "professional" versus "friendly" the interviewer should be during the interview itself, in order to get the best quality data (we tend to take the "friendly" approach).

An oft-asked question is whether field interviewers can be taught to perform laboratory cognitive interviews. We believe that this is possible if interviewers can be induced to "unlearn" some habits that are very valuable for field interviewing, but that may be counterproductive for cognitive interviewing. In particular:

a) Field interviewers have learned over time simply "to make a question work", for example, by repeating it slowly, so that a confused respondent will ultimately provide a codeable response. It must be emphasized that our task in the lab is the reverse; to find flaws in the questions.

b) Interviewers tend to work as fast as possible in the field, usually in order to complete a very long interview before the respondent becomes uncooperative. Interviewers must be reminded to work at an unhurried pace in the lab.

c) Field interviewers often focus their attention on very detailed formatting and other structural features such as skip pattern errors and redundant questions. They must be instructed that the format of the questionnaire may be very rough, and that it is the content that is of primary concern in lab testing.

d) Field Interviewers are taught to ask questions exactly as worded, and not to deviate from the instructions contained in the instrument. In the laboratory, cognitive interviewers must be comfortable departing from the questionnaire flow when this appears to be called for. They also must be able to adjust to a situation in which sequencing instructions are incorrect or missing, which often occurs in the testing of a draft questionnaire.

## **C) HOW IS COGNITIVE INTERVIEWER TRAINING DONE?**

Cognitive interviewing is an acquired skill, consisting of a number of separate skills. Optimally, good interviewers can serve as "detectives" who can find problems in survey questions, and as "engineers" who can work toward developing workable solutions to the problems defined. We have found that the former skill is obtained more quickly than the latter, and that the attainment of mastery is very gradual. Interviewers can be taught in an incremental, step-wise fashion, consisting of the following steps:

- a) Trainee interviewers should conduct technical reviews of questionnaires to make determinations of structural and potential cognitive problems. They also attend early questionnaire design meetings, as well as meetings where cognitive interviewers discuss the results of lab testing.
- b) Trainees familiarize themselves with material on the philosophy and purposes of the cognitive aspects of survey methodology and intensive laboratory interviewing.
- c) They are taught the specific probing methods for use in the laboratory.
- d) They are shown examples of the way that probing is used to detect problems in survey questions. This can be in both written form, and through the use of audio- and video-taped recordings of previous interviews.
- e) Trainees observe experienced interviewers performing actual interviews. Unless a topic is very sensitive, subjects generally have no objection to being observed by an individual who is described as "in training".
- f) Trainees perform one or more interviews while being observed by a practiced interviewer, or compile tape recording of the interviews for review by other staff. The trainee can then be given feedback.
- g) Trainees attend questionnaire review meetings, subsequent to the interviews, and attempt to make specific recommendations for solution of the observed problems.

For organizations that do not have existing cognitive interviewers available to teach new staff, this manual is itself intended to serve as "training". There is no substitute for experience, however, and interviewers should begin interviewing as soon as they have a fairly good idea of what is involved.

## D) OTHER CONSIDERATIONS FOR INTERVIEWING:

There are several features of laboratory interviewing that are important for cognitive interviewers to understand, and that are useful to express to the subject, before beginning a laboratory interview:

- a) Stress to subjects that we are not primarily collecting survey data on them, but rather testing a questionnaire that has questions that may be difficult to understand, hard to answer, or that make little sense.
- b) Make clear that although we are asking the subject to answer the survey questions as carefully as possible, *we are primarily interested in the ways that they arrived at those answers, and the problems they encountered.* Therefore, any

detailed help they can give us is of interest, even if it seems irrelevant or trivial.

c) If think-aloud responding is desired, tell subjects, at the least, to "think out loud to the extent possible, so we can tell what you are thinking about when you answer the questions". Be warned that this introduction generally does not produce a great amount of thinking-aloud, however! To get more of a true "think-aloud", you need to have subjects practice the technique.

d) It also is somewhat helpful to add: "I didn't write these questions, so don't worry about hurting my feelings if you criticize them- my job is to find out what's wrong with them". This helps to "bring out" subjects who may otherwise be sensitive about being overly critical.

## E) FEATURES OF LAB OPERATION

(See Appendix 1, entitled: "Details of day-to-day lab operations" for more complete information.)

### 1) RECRUITMENT:

We identify and recruit volunteers from appropriate sub-populations for testing the survey questionnaire:

a) Subjects either have characteristics of interest for the survey (a particular status with respect to health, work, age, sex characteristics), or they may be "general" subjects, for questionnaires that are asked of the general population. However, even for a questionnaire that is intended for special populations, it is worth testing the initial screening sections on people who don't have the characteristic of interest. This allows the interviewers to make sure that the questions don't create problems in the majority of cases in which the questionnaire will be administered (where the respondent does not have the characteristic of interest). Because this is an important point, I provide a specific example. A questionnaire that is intended to identify individuals with Podiatric conditions might be tested only on individuals who answer an advertisement for "people with foot problems". However, failure to test the screening questions on individuals who do not have foot problems could be catastrophic. If for example, virtually everyone answers initial, screening questions (in effect asking: "do you have any foot problems") in the affirmative, a large number of inappropriate respondents might wind up "passing" the screener and be subjected to a series of completely irrelevant follow-up questions. As a general rule, questionnaires that seek to "identify" a particular population should be tested to determine that they adequately 1) screen in people having the characteristic of interest, and 2) screen out those who do not.

b) Subjects are recruited through newspapers, fliers, service agencies, and support groups. An example of a newspaper ad and a flyer are provided in Appendix 2.

c) Statistical sampling methods are NOT normally used in obtaining laboratory subjects. At most, we use a "quota" sample, in which we try to obtain a range of ages, genders, and socio-economic levels, if possible.

## 2) PAYMENT:

As of 1994, we pay subjects \$30 for a one-hour interview. This amount is sufficient to pay for the subjects' travel time and for basic inconvenience involved in travelling to the laboratory. Further, this payment is enough that it is not simply a "token" remuneration; this way, we are less likely to only recruit individuals who are "practiced volunteers" and have a particular interest in volunteering mainly out of interest in health topics, and who are therefore very different from the usual household survey respondent. However, the amount of payment should be determined by considering a number of issues, such as the general demographic status of the types of subjects required, difficulty of getting to the interview location, difficulty of the task, and so on.

## 3) ADMINISTRATION MODE: FACE-TO-FACE VERSUS PHONE:

We have performed most cognitive interviews "live" in the laboratory, on a one-on-one basis. However, it is also possible to conduct these interviews over the phone, once an interview has been scheduled. We have never called "out of the blue" to someone selected randomly, as in a Random-Digit-Dial telephone survey). We have found telephone interviews to be useful for several specific purposes:

- a) To test questionnaires intended for telephone administration.
- b) To interview subjects (elderly, disabled) who are unable to travel to the lab.

Generally, we prefer in-person interviews when possible, because this allows observation of non-verbal cues, and provides a more natural type of interchange between the subject and the interviewer. However, we advocate the imaginative use of many different testing modes (for example, we have even conducted telephone interviews in the laboratory, where the interviewer and subject are placed in different rooms).

#### 4) PHYSICAL LAYOUT OF LAB:

Any quiet room, such as a conference room or empty office, can serve as a "laboratory" in which to conduct interviews. Optimally, a special interviewing room can be set up, with one-way mirrors, though this is not vital. Interviews can even be conducted "off site", such as at a location where health services are given, or where the target group will be found (for example, elderly centers, schools, and so on).

#### 5) STAFFING:

It is helpful to have permanent staff members who have a "history" of cognitive interviewing experience. As mentioned above, interviewing skill is an acquired capacity, and it helps to avoid the need to constantly train new staff members. It also helps to have a particular staff member who can be responsible for lab management; making phone calls, scheduling, and generally monitoring laboratory operations. The details of lab management are discussed in more detail in Appendix 1, Details of day-to-day lab operations. Further, staff should have experience in relating to clients or sponsors of questionnaires, in order to communicate the findings from laboratory interviews, and the questionnaire design experience necessary to translate laboratory findings into realistic and practical solutions.

#### 6) EQUIPMENT:

All one really needs is a tape-recorder, as it is helpful to have recorded interviews (most subjects don't mind being recorded, as long as a consent form is used). Video-taping is another possibility that we have used to a limited extent. We recommend that if respondents are to be videotaped, a means be found for hiding the camera or making it minimally obtrusive (though of course informed consent from the subject for taping is obtained). A camera that is in full view may strongly affect the nature of the interaction. Some organizations also make use of one-way mirrors for observation; these might also affect the interchange, however, especially when the questions that are asked are sensitive or potentially embarrassing.

## VIII. WHAT HAPPENS AFTER THE INTERVIEW?

### 1) RECORD KEEPING:

We find that the most efficient way to process "data" is to abstract from written comments that interviewers make during the interview. Each interviewer can simply enter comments below the appropriate question on a computerized version of the questionnaire (see the sample comment sheet in Appendix 4). These comments can then be aggregated, over interviewer, and over interview, for a complete review of a particular draft of the questionnaire.

### 2) ANALYSIS OF TAPED INTERVIEWS:

Some researchers prefer to rely on standardized analysis of tape recordings of interviews. We have found, however, that this is a very time-consuming activity, and the appropriateness of this activity depends on the nature of the testing. For "production" work, in which revisions are made at a fairly quick rate, it is often not possible to devote the resources necessary to transcribe or analyze taped interviews. In this case, reliance on written notes is sufficient. Tape-recording is still valuable, however, where a sponsor or client may want to listen to the tape to get a "first-hand" impression of how the questionnaire is working. Transcription or analysis of these tapes can also be valuable for purposes of research, in addition to pure questionnaire development.

Generally, we do not make use of much quantitative data; mainly, we look for:

- a) Dominant trends across interviews (problems that seem to emerge repeatedly).
- b) "Discoveries": Even if they occur in only a single interview, there are some problems that prove to be very important, because they can severely threaten data quality in a few cases, or because these problems are expected to be fairly frequent in the actual survey.

We rely strongly on interviewer judgment, in determining the implications of the laboratory results, as these have ramifications for the fielded survey. For example, one might conclude that a particular interview was very idiosyncratic, and should be ignored. Or, it may be felt that the set of subjects tested was much more sophisticated than the population to be surveyed, so even a hint of comprehension problems might serve to motivate the designers to attempt a simplification of the questionnaire. The general point is that it is usually dangerous to conclude, for example, that if problems are found in 30% of lab interviews, then they are to be expected in 30% of field interviews. We must always be careful to apply a type of subjective "correction factor" to the lab

findings, based on our knowledge of the likely differences that exist between the lab and field environments.

### 3) MEETINGS AND SUBSEQUENT MODIFICATION:

Because the focus of cognitive interviewing is the detection of questionnaire problems, there is often a tendency to "get into the lab quickly", and then deal with the problems that emerge. It is imperative, however, that initial meetings be conducted prior to interviewing, to make clear the objectives of the questionnaire, and that interviewers conduct a technical review of the first draft. In fact, experienced cognitive interviewers can often anticipate the types of difficulties that may be expected, prior to interviewing. Once an initial review, and perhaps a modification, has been conducted, interviewing can start in earnest. After a sufficient number of interviews are completed, and interviewer notes are compiled, one can convene a group meeting to discuss findings. The determination of what a "sufficient" number of interviews is depends on several factors:

- a) If it becomes obvious after several interviews that there are major problems to be rectified, then there is little use in conducting more interviews before modifications are made to the questionnaire. Especially in the very early stages of development, as few as four interviews may be sufficient to constitute a "round" of interviewing.
- b) Even if it appears that more interviews should be done, we have seldom found it necessary, or helpful, to conduct more than 12 - 15 interviews before meeting or delivering comments concerning that round of interview results.

At the meeting, interviewers should discuss their findings in detail with any questionnaire designer who has not participated in the interviewing process. *As a general rule, it is very beneficial that all everyone who is actively involved in the questionnaire design process, including clients, participate in laboratory testing, even if simply as observers.* Clients or sponsors should also be encouraged to observe interviews, or to listen to tape recordings; the impact of a question that is confusing or long-winded is very difficult to ignore when such evidence is used. Very often, where abstract discussions concerning the flaws contained in a questionnaire are unconvincing to a sponsor or client, the evidence from only a few laboratory interviews can have a positive impact. This is a point that we stress; beyond its strength in *identifying* problems, a major positive feature of the cognitive laboratory approach is in the *persuasiveness* of the information it collects.

Meetings should be used both to point out identified problems and to suggest resolutions to these problems. An advantage of the cognitive approach is that,

if one understands the basis for the failure of a particular question, a resolution to the problem may be readily suggested. For example, if a term is clearly not understood, the group may search for an easier-to-understand substitute. Likewise, if it is found that a reference period for a question is far too long for subjects to recall information with any confidence, it is indicated that the use of a shorter interval is in order.

#### 4) THE NEXT LAB TESTING ROUND:

After the questionnaire has been revised, based on the comments from the group meeting, and based on any discussions with clients or sponsors, a new round of interviewing can be conducted to test the changes made, and to provide additional testing of questionnaire segments that were not yet changed. Several issues are pertinent at this stage:

a) *Number of interviewing rounds to conduct.* In one sense, a questionnaire could be tested forever, and still have problems (there is no such thing as a perfect question). Optimally, one wants to test until all the major problems have been detected and satisfactorily addressed. Usually, however, surveys are subject to strict developmental time-frames, and there is limited time for testing. Thus, the usual scenario involves conducting as many rounds as possible (three to four), prior to the pretest, or to actual administration. Though limited lab testing may not produce a perfect questionnaire, the final product should be markedly better than if not subjected to any testing of this type (although this is always subject to the client or sponsors' willingness to make recommended changes).

b) *Changes in the nature of interviewing.* As noted earlier, the nature of the interviewing rounds tends to change over the course of development of a questionnaire. Early in the process, findings relate not only to particular question wordings, but to more global issues, such as the appropriateness of the survey measurement of major concepts that the questionnaire is attempting to cover. For example, it may be determined that a class of information is simply not available through reliance on respondent knowledge and memory (for example, we have found that fathers especially tend to have appallingly bad knowledge of their toddlers' immunization histories). Or, it may be determined that a concept is much too complicated to be measured in a few short questions, and that a long series would actually be required to adequately cover this level of complexity.

Once major conceptual problems have been ironed out, later rounds of interviewing tend to be focused more exclusively on the appropriateness of individual questions (as in the examples in Section IV). Still, the unit of analysis is not the particular survey question, apart from its context;



relevant findings may cover a series of questions, and relate to clarity, appropriateness of the series, biases due to ordering, and so on. One of the challenges of engaging in a useful cycle of testing and development activities is that we must be cognizant of all of these levels, both small- and large-scale, simultaneously.

## IX. LIMITATIONS/CRITICISMS OF THE LABORATORY APPROACH TO QUESTIONNAIRE DEVELOPMENT

Here, I address specific potential criticisms and limitations of the cognitive laboratory approach:

### 1) LABORATORY SUBJECTS DIFFER FROM SURVEY RESPONDENTS:

Laboratory volunteers are self-selected into participation, and are therefore clearly not representative of the survey population as a whole. Most importantly, we find that laboratory volunteers tend to be higher in education than the average survey respondent. This could have important ramifications, in that we might miss problems that occur in "real life", and the laboratory findings therefore underestimate of the severity of problems. However, note also that if a question does not work in the lab, with our more "able" subjects, it will almost certainly be expected to cause problems in the field (and we generally find no shortage of problems in the lab). We suggest that relying on the field pretest, on additional forms of pretesting (such as behavior coding of the interaction between interviewer and respondents), on a subjective assessment of the abilities of our lab subjects, and on a concerted effort to use multiple recruitment sources, be used to counter this problem as much as possible.

### 2) THE LABORATORY CONTEXT IS DIFFERENT FROM THE FIELD:

This is another reason why the lab is not a substitute for the field test. To see how the questionnaire works in "real-life" circumstances, it has to be tested under field conditions, and this is worth doing, even for a small survey, with a few test respondents. We believe that the extent to which the differences in question-answering contexts between lab and field matters depends greatly on the type of question. For example, comprehension processes appear not to differ greatly between the lab and the household; if someone does not know the location of his or her abdomen in the lab, it is doubtful that he or she will know this when at home. Retrieval processes, similarly, will be different between lab and field to the extent that the home environment provides cues that affect the nature of the retrieval process. This again does not appear to be a great problem, in our experience. The case is much different, however, for survey questions that ask about sensitive topics. Here the environment appears to be critical, and in fact often overshadows the other, more basic cognitive processes. Therefore, we never use the laboratory to attempt to directly assess how likely people will be to answer survey questions about such activities as drug use or sexual behavior. Rather, we use the lab only as a context for more indirect,

experimental studies, in which we interview individuals about their understanding of questions, or about conditions they believe would be more or less likely to prompt them to answer truthfully in a fielded survey.

### 3) SAMPLE SIZES FROM COGNITIVE INTERVIEWING ARE SMALL:

It is sometimes argued that the laboratory is deficient, compared to the field pretest, because the samples used are too small to make reasonable conclusions. There are at least three faults in this argument:

a) *The purpose of laboratory interviews is not statistical estimation.* We do not desire sample sizes large enough to supply precision in statistical estimates. Rather, we strive to interview a variety of individuals.

b) *The nature of laboratory interviews is qualitative, not quantitative.* We do not evaluate problems in survey questions simply by counting the number of interviews in which a problem occurs. Of course, if every interviewer reports a problem with a particular question in every interview, that is significant. However, a finding can be based on one interview; that is, an interviewer may say that "I had a person with a particular disease for which this question does not work...". This points out a potential problem which does not need to be verified by finding a large number of other individuals with the same situation; the problem is there, and needs to be addressed.

c) *The apparent increased sample size of the field pretest is often illusory.* As discussed previously, questionnaires often contain initial screening questions, and then long follow-up sections that apply only if one "passes" the screener. However, in cases where it is relatively infrequent that respondents receive the follow-up questions, the general-population-based field pretest tends to provide only a few cases in which these follow-up questions are actually tested. For example, we found that one pretest of 300 households passed less than 10 individuals on to an important section on "use of assistive devices" (canes, wheelchairs, etc.). On the other hand, prior laboratory testing of the same questionnaire had specifically incorporated recruitment of subjects who would naturally screen-in to the follow-up questions, and so the "effective" size of the lab sample turned out to be significantly larger.

### 4) DEMONSTRATION OF THE EFFECTIVENESS OF COGNITIVE INTERVIEWING:

It has been pointed out that there are few comparative experimental studies that demonstrate the superiority of cognitive interviewing to other pretesting methods (however, see Presser and Blair (1993) for a study containing some features of such a study). We have addressed the question of utility of cognitive

interviewing elsewhere (Willis, Royston, and Bercini, 1991). We have used a range of pretesting activities, and our belief is that cognitive interviewing should not be directly "stacked up" against any other particular type of evaluation activity, such as the field pretest, the expert review, behavior coding, and so on. Each type of activity appears to be useful in its own right, and provides information that is unique to that activity. Different processes seem to be most applicable at different points in the developmental sequence— one does not conduct a field pretest on an embryonic questionnaire, and on the other hand, the final "dress rehearsal" of a forms-designed questionnaire is not usually tested with a round of cognitive interviews. The challenge to pretesting is to utilize a cohesive developmental plan that takes advantage of the strengths of each method.

## References

- Aday, L.A. Designing and conducting health surveys (1989). San Francisco: Jossey-Bass Publishers.
- Campanelli, P., Martin, E., & Rothgeb, J.M. (1991). The use of respondent and interviewer debriefing studies as a way to study response error in survey data. The Statistician, 40, 253-264.
- Campanelli, P., Rothgeb, J.M., & Martin, E. (1989). The role of respondent comprehension and interviewer knowledge in CPS labor force classification. Proceedings of the Section on Survey Methods Research, American Statistical Association, pp. 425-429.
- Cannell, C.F., Miller, P.V., & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (ed.), Sociological methodology (pp. 389-437). San Francisco: Jossey-Bass.
- Dippo, C. (1989). The use of cognitive laboratory techniques for investigating memory retrieval errors in retrospective surveys. Proceedings of the International Association of Survey Statisticians, International Statistical Institute, pp. 323-342.
- DeMaio, T.J., Martin, E.A., & Sigman, E.P. (1987). Improving the design of the decennial census questionnaire. Proceedings of the Section on Survey Methods Research, American Statistical Association, pp. 256-261.
- Esposito, J., & Hess, J. (May, 1992). The use of interviewer debriefings to identify problematic questions on alternative questionnaires. Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL.
- Forsyth, B., and Lessler, J.T. (1991). Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (eds.), Measurement errors in surveys. New York: Wiley.
- Jabine, T.B., Straf, M.L., Tanur, J.M., & Tourangeau, R. (eds.). (1984). Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington, DC: National Academy Press.
- Jobe, J.B., & Mingay, D.J. (1991). Cognition and survey measurement: History and overview. (1991). Applied Cognitive Psychology, 5, 175-192.
- Jobe, J. B., Tourangeau, R., & Smith A. F. (1993). Contributions of survey research to the understanding of memory. Applied Cognitive Psychology, 7, 567-584.

Lessler, J.T., & Sirken, M.G. (1985). Laboratory-based research on the cognitive aspects of survey methodology: The goals and methods of the National Center for Health Statistics study. Milbank Memorial Fund Quarterly/Health and Society, 63, 565-581.

Payne, S.L. (1951). The art of asking questions. Princeton, NJ: Princeton University Press.

Presser, S., & Blair, J. (1993). An evaluation of alternative pretest methods. Contract Final Report, submitted to the Division of Health Interview Statistics, National Center for Health Statistics.

Royston, P., Bercini, D., Sirken, M., & Mingay, D. (1986). Questionnaire Design Research Laboratory. Proceedings of the Section on Survey Methods Research, American Statistical Association, pp. 703-707.

Tourangeau, R., & Rasinski, K.A. (1988). Cognitive processes underlying context effects in attitude measurement. Psychological Bulletin, 103, 299-314.

Willis, G.B., Royston, P., & Bercini, D. (1991). The use of verbal report methods in the development and testing of survey questionnaires. Applied Cognitive Psychology, 5, 251-267.

## APPENDIX 1: DETAILS OF DAY-TO-DAY LAB OPERATIONS

This section is mainly of benefit to those who are interested in more detail on recruitment and scheduling. We have found it highly beneficial to establish a dedicated laboratory-manager position, and an integrated computerized control system to monitor the operations of our laboratory. This section will be of interest to those who envision the adaptation of cognitive interviewing techniques on a fairly large scale, as we have done.

### I. RECRUITMENT METHODS:

Clearly, efficient recruitment methods are essential to running a questionnaire laboratory. Flexible and timely methods are needed to handle the wide range of subject characteristics that are often needed in a short period of time. The methods we use to recruit subjects include newspaper ads, flyers, interviews scheduled through outside organizations, and re-contacts with previous subjects. We also reach a fair number of subjects through word-of-mouth. A lot of thought goes into deciding the best approach to get volunteers to test any given questionnaire. Sometimes recruitment involves only one method, while other times several methods are used. The recruitment method used for a particular questionnaire depends on several factors, such as the testing schedule, the subject characteristics we require, and how many subjects we need.

#### 1) Newspaper ads:

Appendix 2 contains an example of an ad we placed in the Washington Post to recruit subjects for a questionnaire on smoking. In newspaper ads, we list any requirements such as age, health conditions, or gender; we also specify that we pay \$30 for a one hour interview. For "general purpose" subjects, we often advertise in local county newspapers. When we need people with very specific health conditions (use of an implanted medical device, such as a pacemaker), we advertise in the weekly health magazine of the Washington Post. Because we receive many responses to most newspaper ads, we use an answering machine to receive the calls, and transcribe the calls later. When using the answering machine, we have found that we can ask people for about three pieces of information without their forgetting one. For example, we will ask for their name, daytime phone number, and one other criterion such as age.

#### 2) Advertising through flyers:

Appendix 2 contains an example of a flyer we used to recruit subjects for the survey on disability. Flyers cost far less than newspaper ads, and can be better for reaching specific groups. Further, we use flyers when we need only a few subjects, because flyers are much less effective in locating large numbers of subjects. We are able to place flyers directly in places or areas where people with the desired characteristics are likely to see them. In the past we have placed flyers in churches, libraries, and Y.M.C.A.s, as well as special interest organizations such as Easter Seals or the Polio Society. To reach people who are unlikely to see our ad in the newspaper, flyers are especially useful. For example, we have had more success recruiting low-income and lower-educated people through flyers than with newspaper ads.

When we want to place flyers in a particular organization, we need to obtain the organization's consent, which can sometimes take several weeks. Phone calls in response to flyers tend to trickle in rather than come all at once as with newspaper ads, so flyers require a longer lead-time than newspaper ads. Answering calls in person, rather than using the answering machine, is advisable when flyers are used, or when the number of calls expected is fairly low.

### 3) Interviewing at outside organizations:

Another way we reach subjects is through the staff of outside organizations that let us conduct interviews in their facility. These off-site interviews are necessary for people that are not able or willing to come to our laboratory. These have included elderly persons, high school students, and the severely disabled who cannot travel easily, as well as drug clinic patients who are reluctant to go to a government office to talk about drug use. Setting up interviews in outside organizations requires a considerable amount of time and effort. When we interview off-site we need organizations to provide us with a private place to interview, and we need help with scheduling. For example, to conduct interviews in high schools, we had to get the approval of the school board. This required a considerable amount of correspondence with various levels of the organization. So far, the organizations we have contacted have been very receptive to our projects.

### 4) Use of "old" subjects:

Finally, a useful way of obtaining subjects is to use the same subject for more than one project; this saves considerable time and effort. We keep a computerized database of all subjects who are interviewed, and their specific characteristics. When we re-use subjects, we can select the "best" ones for our purposes - both in terms of ability to take part in a cognitive interview and desired characteristics. Because of the danger of "over-educating" the subjects, we limit them to three interviews. We do not want subjects to become so familiar with our process that they try too hard to give us suggestions and fail to react spontaneously.

## II. EFFECTIVENESS OF RECRUITMENT METHODS.

Newspaper ads generate the most subjects. Washington Post ads that list specific criteria such as people with disabilities will yield over 100 responses in a few days, and those for general purpose subjects may return as many as 400 responses. County or college papers tend to yield far fewer responses. An important point to note is that when we increased our payment from \$20 to \$30 for a one-hour interview, this greatly increased the number of responses.

## III. SUBJECT CONTROL INFORMATION.

### 1) Subject data sheets.

Appendix 3 contains an example of a subject sheet that we use to collect information about each volunteer. We use this form to obtain the information at the time of the interview; the information is entered into the computerized database later. We keep track of the day, time, questionnaire version, and the interviewer. We also note the specific criteria pertinent to the subject, if any, such as the type of disability. At the bottom of the sheet we list the specific instruments that were tested and the usefulness of the subject.

### 2) Database programs.

All of the information contained on the subject sheet is entered into the computer in a set of user-friendly programs created in a database management package (DBASE IV). These programs help keep track of and organize most of the day-to-day operations of the lab. Computerized databases are maintained both for every subject who is interviewed, and for every questionnaire that is tested. The subject database keeps track of the demographic variables, and also keeps records of people who are scheduled for interviews but fail to arrive, the number of times a subject is interviewed, and the



overall usefulness of each subject. Information contained on the questionnaire database includes the questionnaire, the version number, and the number of interviews per version. Additional tasks that the database system performs include keeping track of interviewers' schedules and of amount of subject payment money available, generating various reports, selecting subjects with various criteria, as well as other routines.

### 3) Timing of subject scheduling.

Timetables for recruiting and scheduling subjects are important. We do not start new recruitments unless a new project will begin within the next month. Although it may seem effective to have a large pool of people "in reserve", we have found that this is not realistic, because subjects have a limited "shelf-life". If a subject is contacted more than a month after he or she responds to an ad, the probability increases that the person is no longer available; he/she may have changed jobs or had some other schedule change, or has simply lost enthusiasm for the project. Additionally, as more time passes, it is more likely that a person has a different residence or phone number.

We schedule subjects no more than a few days before the interview. If the subject is scheduled any earlier, he or she is more likely to be a no-show. If we must schedule more than a week in advance, a reminder call is necessary. Interviews last about one hour, which is about how long subjects can maintain their attention in the lab before becoming fatigued or losing interest. We usually schedule no more than 12 interviews per round, depending on how many people are going to be doing the interviews. This is usually enough to find flaws and suggest changes before conducting another round of interviews. It is inefficient to conduct too many interviews per round. Once a flaw is found, there is little point in conducting many more interviews until attempts are made to solve the problem.

In summary, in order for a recruitment and management system to work, it needs to be flexible, and rely on a variety of techniques as they appear warranted.

## APPENDIX 2: EXAMPLES OF NEWSPAPER ADVERTISEMENT AND FLYER

Newspaper advertisement used to recruit subjects for a survey on health promotion practices:

### EARN \$30 IN ONE HOUR

The U.S. Public Health  
Service in Hyattsville, MD  
needs paid volunteers,  
AGES 16 - 64  
to spend one hour  
answering survey questions  
on a variety of health topics.  
Anyone may participate, but we  
especially need people  
AGE 40 OR ABOVE  
OR  
WHO SMOKE CIGARETTES.

For more information, please call  
301-436-7460

Centers for Disease Control and Prevention/  
National Center for Health Statistics

Flyer used to recruit subjects for a disability survey:

(ORGANIZATION LETTERHEAD APPEARS AT TOP OF FLYER)

# **WE NEED YOUR HELP!**

**THE U.S. PUBLIC HEALTH SERVICE is planning an important national survey on DISABILITY. We need volunteers to participate in interviews so that we can test the survey questions.**

## **VOLUNTEERS WILL BE PAID \$30.**

**You may help with this study if you, your child, or another family member living in your household has a disability. We need volunteers with many different kinds of disabilities. Interviews will be held one block from Prince George's Plaza, at the National Center for Health Statistics in Hyattsville, Maryland. Interviews take about one hour and are conducted during regular business hours.**

**For more information, or to volunteers, please call 301-436-7460, weekdays between 9 AM and 4 PM.**

**APPENDIX 3: SUBJECT DATA SHEET**

SUBNO: S- _____	<input type="checkbox"/> Entered	<input type="checkbox"/> Closed	<input type="checkbox"/> No Show	<input type="checkbox"/> Cancelled
STUDY _____	CRITERIA: _____ _____ _____ _____			
DAY _____				
TIME _____				
INTERVIEWER _____				

**INTERVIEWERS PLEASE FILL IN BELOW:**

Name: \_\_\_\_\_ Date: \_\_\_\_\_

Home phone: \_\_\_\_\_ Employed Y/N \_\_\_\_\_ Work# \_\_\_\_\_

Where did you see our ad/flyer? ☐ PG Sentinel ☐ PG Journal  
☐ Shoppers Food ☐ Giant ☐ Unemployment Office  
☐ P.G. Social Services ☐ Friend

Sex: ☐ Male ☐ Female Birthdate: \_\_\_\_/\_\_\_\_/\_\_\_\_ Age: \_\_\_\_ Education: \_\_\_\_

Marital Status: ☐ Married ☐ Divorced ☐ Widowed ☐ Separated  
☐ Never been married

Household Income: ☐ 20K or less ☐ over 20K  
☐ 30K or less ☐ over 30K

Racial/Ethnic Group:

☐ White ☐ Black ☐ Asian/Pacific Islander ☐ Hispanic ☐ Eskimo, Aleut, or American Indian

Address \_\_\_\_\_

INSTRUMENTS TESTED:				
_____				
_____				
_____				
Interviewer rating of overall subject usefulness:				
1 (poor)	2	3	4	5 (excellent)

Comments: \_\_\_\_\_

## **APPENDIX 4: EXAMPLE OF INTERVIEWING NOTES TAKEN BY A COGNITIVE INTERVIEWER**

### **NOTES:**

- 1) The example represents notes taken by one interviewer from three cognitive interviews, on one version of a questionnaire draft.
- 2) The use of PROBE refers to prepared probe questions embedded in the draft of the questionnaire, for use in all interviews.
- 3) Comments preceded by "S#" were findings specific to that subject (for example, S#1).
- 4) General comments (not pertaining to any particular subject) are not preceded by a subject number, and are simply observations about the questionnaire made by the interviewer either during or subsequent to interviewing.
- 5) The interviewing notes contains only those questions that were found to contain problems or issues to be resolved. That is, the original draft is "abstracted" to create the report, but the original question numbers and section headings are retained for reference.

# PRELIMINARY 1994 Health Promotion and Disease Prevention Questionnaire:

Draft 1: 5/17/93

Subject #1 = female, 51  
Subject #2 = male, 77  
Subject #3 = female, 21

## SECTION A: ENVIRONMENTAL HEALTH

These next questions are about this home.

2. How many floors or levels does this home have? Do not include unfinished basements or attics.

QUESTION DOESN'T WORK FOR APARTMENT BUILDING- SUBJECT #3 REPORTS ALL LEVELS IN BUILDING

4. Do you have at least one smoke detector on each floor of your home not including unfinished attic or unfinished basement?

THIS CAN BE COMPLEX IF WE WANT TO MEASURE THIS EXACTLY: S#2 HAS A 4-LEVEL SPLIT-LEVEL. THERE ARE 3 SMOKE DETECTORS, EACH ONE IN THE STAIRWAY BETWEEN LEVELS. SO, YES, THEY ARE COVERED BUT DON'T LITERALLY HAVE "ONE DETECTOR PER FLOOR".

- 9b. How many PEOPLE smoke cigarettes, cigars, or pipes anywhere inside this home?

PROBE: WHO SMOKES IN THIS HOME?

\_\_\_\_\_  
Number

S#1: WANTS TO KNOW WHETHER TO INCLUDE VISITORS. DO WE MEAN "EVER" OR "USUALLY" HERE?

- 9d. On the average, on the days that there is smoking, about how CIGARETTES are smoked per day ANYWHERE INSIDE this home?

PROBE: IS THIS EASY OR DIFFICULT TO SAY?

00 [ ] Less than cigarette per day/Rarely  
Cigarettes per day  
09 [ ] DK

S#1: SHOULD WE ADD A RESPONSE CATEGORY FOR PACKS? (THIS SEEMS TO BE A NATURAL TYPE OF RESPONSE GIVEN)

**SECTION B: SMOKING**

5. During the past 12 months, have you quit smoking for one day or longer?

- 1 ☐ Yes  
2 ☐ No

PROBE: FOR HOW LONG HAVE YOU QUIT SMOKING? HOW MANY TIMES?

SUBJECT #1 THINKS THAT "STOPPED SMOKING" IS A BETTER PHRASE.

---

6b. On average, when you smoked DURING THE PAST 30 DAYS, about how many cigarettes did you smoke EACH day?

NOTE: I CHANGED THE WORDING ON THIS ONE (ADDED IN CAPS) -  
THIS SEEMS TO WORK OK.

---

7. Would you like to completely stop smoking cigarettes?

- 1 ☐ Yes  
2 ☐ No  
9 ☐ DK

MAKE IT "QUIT SMOKING"

---

8. Have you ever used snuff such as Skoal, Skoal Bandits, or Copenhagen?

- 1 ☐ Yes  
2 ☐ No  
9 ☐ DK

PROBE: HAVE YOU EVER TRIED IT?

IF NO, SKIP TO 11.

---

11. Have you ever used chewing tobacco, such as Redman, Levi Garrett, or Beechnut?

- 1 ☐ Yes  
2 ☐ No (next page)  
9 ☐ DK (next page)
- 

**SECTION C: EXERCISE**

1. Do you exercise or play sports regularly?

- 1 ☐ Yes  
2 ☐ No  
9 ☐ DK

PROBE: WHAT DOES "EXERCISE REGULARLY" MEAN TO YOU?

S#1 THINKS THIS MEANS 7 DAYS A WEEK. NOT WHAT WE MEAN- TOO VAGUE.

S#2: HE SAYS NO. AFTER PROBING, I FIND OUT THAT HE WALKS 2-3 MILES EVERY DAY AFTER DINNER. THE CONCEPT OF "REGULARLY" IS TOO VAGUE AND RESPONDENT-DEFINED.

S#3 IT'S NOT CLEAR TO HIM WHAT "REGULARLY" IS.

---

**SECTION D: OCCUPATIONAL SAFETY AND HEALTH**

ITEM T1:

During the past 2 weeks, did you work at any time at a job or business not counting work around the house? (Include unpaid work in the family [farm/business].)

**THIS IS A REALLY TECHNICAL AND COMPLEX WAY TO ASK WHETHER THEY HAVE WORKED IN THE PAST TWO WEEKS**

---

a. Altogether, does your employer have 50 or more employees?

**S#3: HOW IS >1 JOB HANDLED?**

---

2a. Does your employer have an official policy that restricts smoking in any way?

**S#3: THE EMPLOYER DOESN'T, BUT THE BUILDING OWNER DOES.**

---

3. Does your employer offer a quit smoking program or any other help to employees who want to quit smoking?

- 1 ☐ Yes
- 2 ☐ No
- 9 ☐ DK

**S#1: DO WE MEAN A PROGRAM PAID FOR BY THE EMPLOYER HERE? THIS WAS UNCLEAR.**

---

5. Which of these exercise programs are made available to you by your employer?

**S#1: THE TERM "MADE AVAILABLE" IS VAGUE.**

- 1 ☐ Walking group
  - 2 ☐ Jogging/Running group
  - 3 ☐ Biking/Cycling group
  - 4 ☐ Aerobics classes
  - 5 ☐ Swimming classes
  - 6 ☐ Non-aerobic exercise classes
  - 7 ☐ Weight lifting classes
  - 8 ☐ Fully paid membership in a health/fitness club
  - 9 ☐ Partially paid membership in health/fitness club
  - 10 ☐ Physical activity or exercise competitions
  - 98 ☐ Other - Specify
- 
- 00 ☐ No programs
  - 99 ☐ DK

**S#3 IS INVOLVED IN A "WELLNESS PROGRAM". WHERE DOES THIS FIT IN?**

**IF NO PROGRAMS OR DK, GO TO 6A.**

---



5c. About how often do you participate in (exercise program) that was made available by your employer?

PROBE: TELL ME ABOUT THE PROGRAMS YOU'VE PARTICIPATED IN.

S#3 SAYS "ONCE"- MEANING ONE PERIOD IN WHICH SHE ACTUALLY EXERCISED SEVERAL TIMES A WEEK.

6. In the past 12 months, did you participate in a quit smoking program made available by this employer?

IF THEY DON'T SMOKE, DON'T ASK THIS.

6a. Does your employer make available screening tests for:

	Yes	No	DK
(1) blood pressure?	1 [ ]	2 [ ]	9 [ ]
(2) cholesterol?	1 [ ]	2 [ ]	9 [ ]
(3) cancer?	1 [ ]	2 [ ]	9 [ ]

S#3: FOR FREE? SHE WANTS TO KNOW.

NOTE: SHOULDN'T THIS BE "IN THE PAST 12 MONTHS"?

7a. Does your employer make available brochures, programs, talks, or counseling concerning ANY OF THESE (CARD T5):

1 [ ]	Weight control	
2 [ ]	Nutrition information	
3 [ ]	Prenatal education	SIMPLIFY THIS TERM
4 [ ]	Medical self-care	
5 [ ]	Stress reduction and management	
6 [ ]	Alcohol and other drugs	
7 [ ]	Sexually transmitted diseases (including HIV or AIDS)	
8 [ ]	Job hazards and injury prevention	
9 [ ]	Back care and prevention of back injury	
10 [ ]	Preventing off-the-job accidents	

SUBJECT #3 WANTS TO KNOW WHAT "MEDICAL SELF-CARE" (#4) IS. GOOD QUESTION.

7b. In the past 12 months, have you participated in any of these activities or used any the information made available by this employer?

1 [ ]	Yes
2 [ ]	No
9 [ ]	DK

S#1 SAYS NO, EVEN THOUGH SHE HAS READ THE FLYERS. WE NEED TO MAKE MORE CLEAR WHAT WE MEAN BY "USED ANY OF THE INFORMATION". MAYBE WE EVEN NEED ONE QUESTION ON "PARTICIPATING IN ACTIVITIES" AND ANOTHER ON "READING ANY INFORMATION".

SUBJECT #2 SAYS NO, BUT ACTUALLY DOES READ THE BROCHURES. THE CONCEPT TO "USING THE INFORMATION" IS NOT COMING ACROSS.

- 8b. Which of these best describes your employer's smoking policy for indoor public or common areas, such as lobbies, rest rooms, and lunch rooms?  
(CARD T2)

**THIS IS COMPLEX AND DIFFICULT TO UNDERSTAND. S#2 DOESN'T REALLY LISTEN- JUST LOOKS AT THE CARD AND FIGURES OUT WHICH RESPONSE APPLIES.**

- 1 ☐ Not allowed in ANY indoor or common public areas
- 2 ☐ Allowed in SOME public areas, including designated smoking areas
- 3 ☐ Allowed in ALL indoor or common public areas
- 9 ☐ DK

#### **SECTION E: CLINICAL PREVENTIVE SERVICES**

1. About how long has it been since your last routine check-up by a medical doctor or other health professional?

**S#3 REPORTS ON A LIMITED FOLLOW-UP VISIT, RATHER THAN AN ACTUAL "CHECK-UP", SO THE FOLLOWING QUESTIONS ARE NOT MEANINGFUL. MAYBE WE NEED TO ASK EXPLICITLY IF THE LAST VISIT WAS A GENERAL CHECKUP, A FOLLOW-UP VISIT, OR A VISIT FOR A PARTICULAR HEALTH PROBLEM. ONLY IF IT'S THE FIRST DO WE REALLY WANT TO ASK Q2, 3, ETC.**

2. During this last check-up, were you asked about:

	Yes	No	DK
a. Your diet and eating habits?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>
b. The amount of physical activity or exercise you get?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>
c. Whether you smoke cigarettes or use other forms of tobacco?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>
d. How much and how often you drink alcohol?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>
e. Whether you use marijuana, cocaine, or other drugs?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>
f. Sexually transmitted diseases?	1 <input type="checkbox"/>	2 <input type="checkbox"/>	9 <input type="checkbox"/>

**S#2: SHOULD WE SKIP E AND F IF OVER A CERTAIN AGE? SEEMS KIND OF RIDICULOUS FOR HIM.**

Refer to age 1 ☐ SP (Sample Person) is 65 + (4)  
8 ☐ Other (6)

- 4a. During this last check-up, were you asked about the symptoms of a transient ischemic attack (TIA)? 1 ☐ Yes      2 ☐ No      9 ☒ X

**S#2: THIS IS VERY TECHNICAL. THE SUBJECT SAYS "YES" BECAUSE HE WAS ASKED ABOUT SHORTNESS OF BREATH. IS THIS A SYMPTOM OF TIA? MAYBE WE SHOULD ASK ABOUT WHETHER THEY WERE ASKED ABOUT THE SPECIFIC SYMPTOMS EXPLICITLY, RATHER THAN ABOUT THE CONDITION.**

- b. During this last check-up, were you asked about whether you have difficulty taking care of yourself, including dressing, using the toilet, bathing, eating, or getting around inside your home without help? 1 [ ] Yes 2 [ ] No 9 [ ] ~~DK~~
- c. During this last check-up, were you asked about whether you have difficulty doing every day activities and chores, including preparing your meals, managing your money, using the telephones, doing light housework, and shopping? 1 [ ] Yes 2 [ ] No 9 [ ] ~~DK~~

**S#2: B AND C ARE REALLY LONG TO READ. I REALLY DOUBT THAT RESPONDENTS CAN KEEP IN MIND ALL THAT WE'RE ASKING HERE.**

5. During this last check-up, did you have:

**S#2: THE SUBJECT MISSES THE REQUIREMENT THAT THIS BE "THE LAST CHECK-UP". WE NEED TO SPECIFY THIS BETTER. THIS NEEDS TO BE EMPHASIZED MORE. MAYBE "DURING THE ROUTINE CHECK UP THAT YOU JUST TOLD ME ABOUT".**

	Yes	No	DK
a. A vision test to see how well you see?	1 [ ]	2 [ ]	9 [ ]

**"SEE HOW WELL YOU SEE" IS KIND OF AWKWARD.**

b. A hearing test?	1 [ ]	2 [ ]	9 [ ]
c. A urine test?	1 [ ]	2 [ ]	9 [ ]
d. A blood test to check your thyroid function?	1 [ ]	2 [ ]	9 [ ]

**S#2: HE HAD A BLOOD TEST: HOWEVER, HE DOESN'T KNOW THAT IT WAS SPECIFICALLY TO TEST THYROID FUNCTION.**

e. A stool test to check for blood in the stool?	1 [ ]	2 [ ]	9 [ ]
--	-------	-------	-------

12. A breast physical exam is when the breast is felt for lumps by a physician or other health care professional. About how long has it been since you had a breast physical exam done by a doctor or other health care professional? Was it within the past year, 1-2 years ago, or over 2 years ago?

**LOTS OF WORDS TO READ HERE.**

#### **SECTION F: ESTROGEN REPLACEMENT**

13. Have you ever been counseled by a physician or other health care professional about the benefits of taking estrogen pills after menopause in order to prevent bone loss?

**"COUNSELING" IS PROBABLY TOO MUCH HERE (SUBJECT #1 THINKS IT IMPLIES "IN-DEPTH" DISCUSSION. MAYBE WE JUST WANT TO ASK IF THE PHYSICIAN TALKED ABOUT THIS WITH THEM.**

**ALSO, IF PREVENTING BONE LOSS THE ONLY REASON TO TAKE ESTROGEN PILLS? IF NOT, THE PERSON MIGHT HAVE BEEN TOLD ABOUT ESTROGEN REPLACEMENT, BUT NOT NECESSARILY WITH RESPECT TO BONE LOSS.**

**SECTION G: FAMILY**

1. In the past month, have you had any discussions with family members in your household about health issues related to:

**PROBE: WHAT DO YOU THINK OF WHEN WE ASK ABOUT "DISCUSSIONS WITH FAMILY MEMBERS"?**

☐ Drinking beer, wine, liquor, and other alcoholic beverages?

**MAKE IT "WINE, LIQUOR, OR OTHER ALCOHOLIC..."**

☐ Nutrition and healthy eating habits?

☐ Exercise, sports or other physical activities?

**PROBE: WHAT DO YOU THINK OF WHEN WE SAY "PHYSICAL ACTIVITIES"**

☐ Sexual behavior, such as risk of infection with HIV or other sexually transmitted diseases or unwanted pregnancy?

☐ Smoking cigarettes or use of other forms of tobacco, including chewing tobacco and snuff?

☐ Using illicit drugs, including steroids?

**"ILLICIT" DRUGS IS TOO TECHNICAL.**

☐ Safety and things that you can do to prevent injuries, including using seat belts?

**S#1 THINKS THAT "HAD DISCUSSIONS" IS TOO MUCH HERE- JUST "TALKED ABOUT" IS BETTER FOR HER.**

**S#2 THINKS THAT "TALKED ABOUT" IS BETTER THAN "DISCUSSIONS"**

**S#3: THE CONCEPT OF "DISCUSSIONS" IS VERY VAGUE AND IT'S DIFFICULT FOR HER TO SAY FOR SEVERAL OF THESE. IF SHE TELLS HER MOTHER TO WEAR HER SEAT BELT, IS THIS A "DISCUSSION"? "TALKING ABOUT" THESE THINGS MAY BE SOMEWHAT BETTER (BUT MAY STILL HAVE PROBLEMS).**

---

# National Center for Health Statistics

## Cognitive Methods Staff

### Working Paper Series

- Number 1: *Questionnaire Design Research Laboratory: Focus groups on family health issues*, by Deborah Trunzo and Susan Schechter (November, 1993)
- Number 2: *Investigations into the cognitive processes of answering self-assessed health status questions*, by Susan Schechter (November, 1993)
- Number 3: *Questionnaire Design Research Laboratory: Mini-pretest of the 1994 Year 2000 and AIDS Supplements*, by Gordon Willis (November, 1993)
- Number 4: *Questionnaire Design Research Laboratory: Cognitive testing of the 1993 NHIS AIDS Knowledge and Attitudes Supplement*, by Jeffrey Kerwin (November, 1993)
- Number 5: *Questionnaire Design Research Laboratory: Cognitive testing of the 1993 Family Resources Supplement*, by Jeffrey Kerwin (November, 1993)
- Number 6: *Questionnaire Design Research Laboratory: Cognitive laboratory testing of the 1993 Teenage Attitudes and Practices Survey II*, by Barbara Stussman (March, 1994)
- Number 7: *Cognitive interviewing and questionnaire design: A training manual*, by Gordon Willis (March, 1994)
- Number 8: *Conducting cognitive laboratory tests by telephone*, by Susan Schechter and Paul Beatty (April, 1994)
- Number 9: *The cognitive aspects of responses to sensitive survey questions*, by Gordon Willis, Monroe Sirken, and Gad Nathan (May 1994)
- Number 10: *Proceedings of the 1993 NCHS Conference on the Cognitive Aspects of Self-Reported Health Status*, Edited by Susan Schechter (August 1994)
- Number 11: *A cognitive subtask model of statistical map reading*, by Douglas Herrmann and Linda Williams Pickle (August, 1994)
- Number 12: *The use of strategic processes by survey respondents*, by Gordon Willis (November, 1994)
- Number 13: *NASA questionnaire development project: Developing an employee health survey from secondary sources*, by Paul Beatty (November, 1994)
- Number 14: *The contributions of the NCHS Collaborative Research Program to memory research*, by Douglas Herrmann (December, 1994)
- Number 15: *The Mini-Pretest: Small-scale testing of a large-scale questionnaire*, by Gordon B. Willis (December, 1994)
- Number 16: *NCI questionnaire development and testing: DES, mammography decisional balance, and clinical breast exams*, by Paul Beatty (January, 1995)
- Number 17: *Laboratory testing of a customer satisfaction survey: The Social Security Initial Awards Questionnaire*, by Paul Beatty (April, 1995)
- Number 18: *Cognitive aspects of statistical mapping*, by Linda Williams Pickle and Douglas J. Herrmann (June, 1995)
- Number 19: *Cognitive testing of NCI occupational exposure questionnaires*, by Paul Beatty, Barbara Wilson, and Mary Ann Guadagno (August, 1996)
- Number 20: *The use of the psychological laboratory to study sensitive survey topics: Can cognitive techniques supplement field study?*, by Gordon Willis (August, 1996)
- Number 21: *Construction of a battery of survey-based instruments to assess cognitive abilities*, by Robert Sternberg, Martin Dennis, and Paul Beatty (October, 1996)

**National Center for Health Statistics  
Cognitive Methods Staff  
Working Paper Series**

- Number 22: *Public perception of health risk: A literature review*, by Jack Elinson and Nurit Guttman (November, 1997)
- Number 23: *Public health geographic information systems (GIS) news and information: 1994-1997*, by Charles Croner (January 1998)
- Number 24: *Cognitive Research on Sensitive Survey Questions*, by Gordon Willis, Kenneth A. Rasinski, Alison K. Baldwin, and Wenchi Yeh ( May 1998)
- Number 25: *Evaluating Discrepancies in Print Reading Disability Statistics*, by Paul Beatty and Wendy Davis (June 1998)
- Number 26: *Questionnaire Evaluation and Testing in Support of the Behavioral Risk Factor Surveillance System (BRFSS), 1992-1998* by Paul Beatty and Susan Schechter (December 1998)
- Number 27: *NHANES 1999+ Methodology Research: Approaches for Estimating Quantities of Foods Reported in a Dietary Interview Setting*. Edgar Chambers, Sandria Goodwin and Erika Edwards (July 1999)
- Number 28: *Survey Research at the Intersection of Statistics and Cognitive Psychology*, edited by Monroe G. Sirken (January 2000)
- Number 29: *Development of a Brochure to Counter Mistrust-based Survey Refusals: Putting Tuskegee in Perspective*, by Joel Kennet, Barbara F. Wilson, Alfredo Calvillo, Karen R. Whitaker, Michele Garber, & Glenn Pinder (August, 2000).
- Number 30: *Interdisciplinary Survey Research Involving the Computer and Cognitive Sciences*, edited by Monroe G. Sirken, (January, 2001).
- Number 31: *Cognitive Testing of a Children's Mental Health Assessment Instrument*, by Kristen Miller and Paul Beatty (March 2001).
- Number 32: *Cognitive Analysis of Sexual Identity, Attraction and Behavior Questions*, by Kristen Miller (June 2002).
- Number 33: *Cognitive Evaluation of Population Activity Limitations Screeners (PALS) [as part of the cognitive testing for the US/Canada Joint Health Survey]* , by Kristen Miller (September 2002).

Copies of these publications are available free of charge from:  
Karen R. Whitaker  
Room 3101, NCHS, 3311 Toledo Rd., Hyattsville, MD 20782  
(301) 458-4569; INTERNET: KWhitaker@cdc.gov