DATA EVALUATION AND METHODS RESEARCH

Series 2
Number 45

Reporting Health Events in Household Interviews:

Effects of Reinforcement, Question Length, and Reinterviews

A methodological study designed to test the effectiveness of certain questionnaire designs and interviewing techniques used in the collection of data on health events in a health interview.

DHEW Publication No. (HRA) 74-1028

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Public Health Service

Health Resources Administration
National Center for Health Statistics
Rockville, Maryland

Vital and Health Statistics-Series 2, No. 45
Reprinted as DHEW Publication No. (HRA) 74-1028
May 1974

NATIONAL CENTER FOR HEALTH STATISTICS

Vital and Health Statistics-Series 2-No. 45

ACKNOWLEDGMENT

This study was conducted by the Survey Research Center, Institute for Social Research, The University of Michigan, under contract with the National Center for Health Statistics.

Technical assistance was provided by Dr. Thomas Tharakan, Mrs. Barbara King, Mrs. Marion Wirick, Mrs. Carol Berlin, and Mrs. Diane Schoeff. The staff of the Community Health Association in Detroit supplied information on the sample for the study. Dr. George LeRoy, Director of Metropolitan Hospital, Mr. John Kerr, Assistant Director, and the medical staff extended special assistance.

FOREWORD

This study, conducted by contractual arrangement with the Survey Research Center, Institute for Social Research, The University of Michigan, is the second in a series of three studies designed to investigate the effects of some experimental interviewing techniques on the amount and quality of information obtained during a health interview. (The other two studies are described in Series 2, Numbers 41 and 49.) The plan for this series was motivated by the findings of an earlier study on interviewer-respondent behavior also completed by the Survey Research Center. The basic study, which is described in *Vital and Health Statistics,* Series 2, Number 26, indicated that reporting in an interview can be more effectively improved by increasing the behavioral interaction of the respondent and the interviewer during the interview than by changing the basic attitudes of the respondent or increasing his levels of information.

In view of this finding, it seemed that improved reporting might be obtained by the introduction of techniques used by the interviewer to encourage respondent reaction during the interview which would stimulate maximum recall. This approach, however, varied substantially from the usual practice of training interviewers to behave in a standardized manner during an interview. The standardized manner, which was restricted to the asking of questions and recording of responses, was an attempt to reduce the known biasing influence on survey data that has been attributed to interviewer performance.

The design of this series of studies has taken advantage of the fact that interviewers can influence respondents, and it has attempted to bring the potentially biasing behavior cues under control--in effect, to incorporate them as a part of the "standardized" behavior. Through the interaction between the interviewer and the respondent it was expected that the systematic changing of the interviewer's technique would change the activity level of the respondent, thereby increasing the amount and quality of reported health information.

Because of the complex relationship between methods of interviewing, the performance of interviewers, and the reporting of respondents, the problem of obtaining accurate data in a household interview is not a simple one. The findings from this investigation of experimental interviewing techniques indicate that verbal "reinforcement" of the respondent (i.e., appreciative comments by the interviewer following fruitful recall efforts by the respondent), question length, direct memory probing, an intensive interview, a diary procedure, and a reinterview can have important effects on survey interview data. More investigation is needed to determine the appropriateness of specific techniques for the collection of certain types of health information and to evaluate their effectiveness in terms of the validity, reliability, and amount of data reported.

Elijah L. White
Director
Division of Health Interview Statistics

CONTENTS

CONTENTS—Con.

REPORTING HEALTH EVENTS IN HOUSEHOLD INTERVIEWS:

EFFECTS OF REINFORCEMENT, QUESTION LENGTH, AND REINTERVIEWS

Kent H. Marquis, Ph.D., Charles F. Cannell, Ph.D., and André Laurent, Doct. (Sorbonne),
Survey Research Center, Institute for Social Research, The University of Michigan

INTRODUCTION

Since its inception, the National Center for Health Statistics (NCHS) has undertaken a continuing research program designed to assess the quality of its data and, more basically, to understand the survey interview process.

Previous research has indicated that respondents make errors when relating factual information in interviews, and that such errors may be influenced by at least three classes of variables: (1) the nature of the information to be reported, (2) the method by which the information is requested, and (3) the response of the interviewer to the reported data. This study is concerned with the second and third variables.

Several NCHS studies indicate some kinds of errors respondents tend to make when describing their health. In recent research of this type, Madow[1] measured response error in reporting chronic illnesses and conditions by persons known to have at least one chronic condition. The records of a large group health insurance plan were used to establish the fact of chronic illness. The data indicated that approximately 45 percent of those conditions shown in the medical records were not reported in the interview. Various approaches to questionnaire construction had no great overall effect on the validity of these data. One of the major predictors of underreporting was a lack of impact of the condition on the respondent. High impact events included worry, long duration of the illness, imposed work limits, and medication, among others. Conditions accompanied by these high impact events were much more likely to be reported than conditions with low impact.

The Survey Research Center (SRC) of The University of Michigan in coöperation with NCHS has undertaken a new series of studies designed to study variables other than the nature of the information reported which might be responsible for variations in response error.

One major study in the new series by Cannell, Fowler, and Marquis[2] was designed to investigate the effects of some psychological characteristics (attitudes, opinions, motives, perceptions, interest, health information, and abilities) of interviewers and respondents on the number of events reported.[a] The investigators were unable to find any significant association between these

[a]Several earlier validity studies compared respondents' declared hospitalizations or visits to doctors with known records. These studies revealed that those respondents who were more accurate in their reporting tended also to report a larger number of health events. Since higher reporting does seem to suggest more accurate reporting, the number of reported events has been used as a dependent variable in later studies where better validity data have not been available.

psychological-cognitive variables and the amount reported. Needless to say, this was contrary to expectations derived from both common sense and accepted interview theory.

The research also involved an effort to collect detailed data on the kinds of behavior exhibited by questioner and respondent during the interview.[b] Their behavior was noted on special forms by an observer who took no other part in the household interview. When the behavioral data were related to the number of items reported, after correcting as much as possible for factors causing spurious correlations, it was obvious that the greater the rate of behavioral activity on the part of the respondent, the greater the number of items she reported. This held true regardless of whether the behavior was task oriented or irrelevant to the interview. Because of this correlation between behavior and number of items reported, the inferences were made that the validity of the data had been improved and that the improvements in the reporting were a function of things that went on during the interview rather than the more remote psychological and demographic characteristics of the respondent and interviewer.

The study was descriptive rather than experimental, so it was impossible to isolate variables which caused or accounted for the general activity level of the respondent. However, there was a very high correlation between interviewer behavior activity level and that of the respondent. This relationship suggested that systematically changing the interviewer's technique would change the respondent's activity level, thereby increasing both the amount and quality of reported health information.

---

[b]The kinds of behavior noted were interpersonal and interview- or task-oriented behavior of both respondent and interviewer. Respondent interpersonal behavior might include laughing, joking, and furnishing unnecessary information, while the interviewer interpersonal behavior included similar items such as asking nonhealth questions of the respondent, flattery, and praise. The respondent's task-oriented behaviors mentioned were such things as elaborations and asking for clarifications. The interviewer's task-oriented behaviors noted were directive and nondirective probes, clarification of questions, and so forth.

The overall pattern of findings from this first study led to hypothesizing a "cue-search" model of the household interview. It was assumed that a household interview was so out of the ordinary that respondents could not generalize a well-defined set of previously held attitudes, feelings, or expectations toward it. Therefore, the respondent looked to the interviewer as a source of cues about how to behave.

Thus, in a subsequent study in the SRC-NCHS series by Marquis and Cannell,[3] the investigators attempted to manipulate systematically two sets of these cues and to assess their effects on the number of health items reported. The first set of cues consisted of positive reinforcement statements given by interviewers in response to information supplied by the respondent. Although this feedback is generally overlooked in considering the interview process, it was felt that it might be a major source of uncontrolled variation. Bringing these cues under some systematic control could greatly influence the respondent and the quality of the information he provided subsequently. Results indicated that a technique in which the interviewer used a positive reinforcing statement after the respondent mentioned a health item resulted in a 25 percent increase in the number of such items reported. This increase occurred for most classes of items—recent and past, medically attended and personally treated, embarrassing and neutral—as well as for items reported on a "yes-no" list or in response to less structured questions.

The second set of cues was inserted at the beginning of the experimental interviews in the form of a symptom list, a list of items containing a wide variety of health symptoms and conditions (serious and minor, embarrassing and neutral, frequent and rare), which was read to the respondent. The interviewer asked her to respond "yes" to each item she had ever experienced, and "no" to the others. The purposes of this experimental procedure were to accustom the respondent to thinking about her health and to convey the idea that a broad range of health items would be covered. However, the number of reported items was not increased by this preparatory procedure.

The present research was designed to provide some clarification of the results of the pre-

vious reinforcement experiments. One hypothesis tested was that an interview using reinforcement reduces response error to a greater extent than an interview not employing reinforcement. An adequate test of this hypothesis required a departure from two important procedures used in previous research. The reinforcement technique mentioned in connection with a previous study[3] involved at least three variables. In addition to verbal reinforcement, the interviewer smiled and looked up at the respondents in the experimental group but neither smiled nor looked at the control group respondents. Also, the questionnaire for the experimental group contained long questions and section introductions augmented by clichés and redundancies; the control questionnaires did not. In the present study the question length variable was manipulated independently of reinforcement, and the interviewer used eye contact and smiles for respondents of both groups.

An alternative hypothesis was also advanced. It stated that reinforcement merely produces more "yes" answers to list-type questions than when reinforcement is not used and that the net effect of this reinforcement on the reporting of listed chronic conditions is not only to reduce the underreporting but also to increase overreporting substantially. Previous validity studies on health reporting had been designed mainly to test the extent of underreporting. The present study was deliberately designed to examine overreporting as well. With the resulting data it therefore was possible to measure changes in both over- and underreporting rates and any "tradeoff" occurring between them.

Research has suggested that question length may be an important determinant of reporting behavior. A recent series of studies, carried out under the direction of Matarazzo and Saslow, is of special interest. They investigated the effects of interviewer behavior on respondent behavior during different types of interviews. Among these studies is a subset dealing with the relationship between the length of time the interviewers spoke and the length of time the respondents answered.[4] Briefly, the results indicated that when interviewers talked more, so did respondents. (The findings seem similar to those from the behavior observation health studies described above showing the amount of total interviewer and respondent behavior to be highly correlated.) The data also suggested that respondents might talk more in answer to longer questions, thereby giving more health information. In the present study the investigators also explored the possibility that increasing interviewer speech duration (by manipulating question length) would yield additional, valid health information from the respondent.

Finally, while not directly derived from prior theory or research, the effects of a reinterview on the validity of chronic condition reporting were tested in this research. It is often assumed that the reinterview, sometimes conducted to provide continuing data in panel studies and sometimes done to check the accuracy of interviewer's performance, produces more valid information than that obtained in the original interview. The little research in the area, most of which has been done by Ferber,[5] does not show clear-cut findings. The classic study in the field by Neter and Waksberg[6] suggests that reinterviews obtain less accurate data than do original interviews.[c] This present study was designed to explore the comparative validity of original interviews and reinterviews and to provide data on effects, if any, which the experimental variables might have on respondent performance when interviewed a second time.

---

[c]Neter and Waksberg[6] uncovered what was termed a "conditioning effect" of reinterviews. This effect is not to be confused with the effect of verbal reinforcement or verbal conditioning tested in this study as the two are distinct phenomena. Neter and Waksberg found a 9 percent loss in the number of jobs reported between the second and third interview. Both interviews had asked for the same information.

SUMMARY OF MAIN RESULTS

Previous published research and pilot studies indicated that variables having the greatest effects on response accuracy in personal interviews are probably the variables most closely associated with the interview process itself rather than with more remote phenomena such as participant attitudes and demographic and social characteristics.

This study was designed to test the effects of interviewer reinforcement, length of question, reinterviews, and respondent education on the validity, reliability, and quantity of respondent reports of health information in a household interview. A special respondent sample not representing a national cross section was used. A measure of validity of respondent reports was approximated by comparing interviewee information on chronic conditions with information obtained from physicians.

The results obtained were not as anticipated initially. They indicated that the effects of different ways of conducting interviews were mediated by the level of education of the respondent.[d] Procedures which increased accuracy and reduced bias for one education group had opposite effects for the other education group.

Care should be taken in generalizing the results of this study since the sample was not typical of the U.S. population. In addition, many of the findings are based on trends in the data, the components of which are not statistically significant. Nevertheless, it appears to be a sound conclusion that small variations in either the asking of questions or reinforcement procedures can produce variations in the quality of data.

The effects of the experimental procedures are summarized below.

---

[d]The initial report from this study did not include education of the respondent as a variable in the analysis. However, further analysis of the same data, which was supported in part by a general purpose research grant from the National Center for Health Services Research and Development (PHS Grant No. HS00252), indicated the importance of education as a variable in explaining the effects of the different interviewing procedures. Subsequently, the initial report underwent modifications to introduce the effects of the respondent's education.

Reinforcement

The interviewer's use of a positive reinforcing verbal statement following reports of morbidity increased the accuracy and reduced the bias of chronic condition reporting only for the group of respondents who had not completed high school. It had the opposite effect of decreasing accuracy and increasing bias for respondents who had graduated from high school. Reasons for these effects are not clear, although it is possible to rule out one plausible hypothesis. It might be expected that less educated respondents were not as able to report their chronic conditions as were more educated respondents and that, therefore, reinforcement had the effect of increasing the ability of less educated respondents to recall and report their chronic sickness. The high education respondents already possessed the ability to recall and report accurately, and hence, would not benefit from the reinforcement procedure. However, it is possible to infer from the data that the two education groups did not differ in ability to report chronic conditions accurately. Therefore, the different reinforcement effects appear not to be due to their influence on different initial levels of ability to report.

Several untested hypotheses are offered to account for the mediating effects of education on the relationship between reinforcement and reporting accuracy. Less educated respondents interviewed by more educated (and, consequently, high status) interviewers "naturally" establish a task-oriented interpersonal atmosphere rather than a personal or social relationship. Reinforcement for adequate performance in this task-oriented context has the effects predicted by theory because reinforcement is perceived by the respondent as appropriate and necessary.

For the more educated respondent, the interviewer-respondent relationship is hypothesized to be more interpersonally oriented and the respondent less concerned about adequate task performance, possibly because of the apparent simplicity of the reporting requirements. Reinforcement, because it is friendly and supportive, may accentuate the tendency to perceive the relationship as personal rather than professional or it may be perceived as inappro-

priate behavior for an equal status relationship. Either or both perceptions may result in biased or inaccurate reporting.

Question Length

Increasing the length of questions about chronic conditions by introducing redundancies and irrelevant material (not by furnishing more pertinent information such as clarifying phrases or examples) significantly increased reporting accuracy and reduced bias for the more educated respondent group. In contrast, these long questions tended to decrease accuracy and increase bias for the less educated respondent group.

The positive and negative effects of long questions as mediated by respondent education cannot be fully explained. It is hypothesized that the additional presentation of key elements of a question and the additional time a long question allows for the consideration of an answer are useful to the higher educated respondent who may need his attention directed more exclusively to the information-giving aspects of the interview situation. Why long questions produced less accuracy and greater bias for the low education group is more conjectural. Possibly the long questions, including the irrelevant statements, merely confused the less educated respondent.

Reinterviews

Accuracy and bias of reporting did not differ to any great extent in reinterviews and initial interviews, but again, the observed differences were mediated by respondent education. Generally, but with some exceptions, the more educated group reported more accurately during the reinterview than during the initial interview. As in the original interview, the use of long questions without reinforcement produced the highest average accuracy for this group. The less educated group showed no improvement during the reinterview.

It may be that the more educated respondents benefited from a second opportunity to think about their chronic conditions and were thereby able to overcome initial tendencies to respond "no" to a question when they were unsure of their answer.

PILOT STUDIES

In order to test some basic assumptions and the feasibility of the method and techniques to be used in the main study, several pilot studies were conducted. Two of these preliminary studies are described briefly here. The first study was designed to see how the use of long questions affected the length of respondent answers and the amount of information given by respondents. The purpose of the second pilot test was to examine the effects of both long questions and reinforcement on answers given by respondents.

Effects of Question Length on Reporting—Pilot Study I

To explore the possibility that increasing the duration of interviewer speech by manipulating question length would yield more health information from the respondent, a pilot study was conducted, using the services of two interviewers on the Survey Research Center staff. The interviewers received about 8 hours of training on the use of the questionnaires and the special procedures for selecting households and respondents. They were instructed to ask the questions exactly as worded and to give clarification only when it was absolutely necessary and then only by repeating the relevant part of the question. Furthermore, they were told to accept the respondent's first answer rather than to probe for more information and were instructed to give no feedback to the respondent. Respondents were chosen from blocks selected at random from two census tracts in Jackson, Michigan. The tracts contained intact white families, of moderate income, with a low proportion of persons over 65 years of age and a high proportion of native-born citizens.

Interviewer speech duration was varied by the use of long and short questions. An interview with 28 questions was created, covering a wide variety of health topics and a wide variety of types of questions (for example, open and closed, specific and general). In an attempt to rule out the possibility that respondents gave longer answers to long questions merely because these asked for more information, length was added to these questions in three ways: (1) by redundancy or asking the same thing twice, with variation in the grammatical structure of each request, (2) by inserting clichés, such as "The next item is...," and (3) by introducing extraneous information, for example, "We ask this of all respondents" or "The health service wants to know about...." The exact words used in a short question always appeared intact in a long one, usually at the end of the question.

The pilot study design employed three questioning procedures. Twenty-seven interviews were obtained, nine in each of these three procedural groups (A, B, and C). The basic questionnaire was divided into four roughly equal parts. In procedure A the first and third parts of the questionnaire contained long questions, while the second and fourth parts asked short questions. For procedure B short questions were used in the first and third parts, and long questions in the other two parts. In procedure C short questions were used in all parts. A possible fourth procedure, employing only long questions, was not tested as it was assumed to be potentially detrimental to useful respondent performance.

Two dependent variables, each thought to be affected by the length of the question asked, were tested: (1) the length of time respondents talked, defined as the number of seconds from the end of the question to the end of the response minus any interviewer speech which intervened; and (2) the percent of questions for which at least one item of information was reported.

Table 1 shows the average length of time the respondent took to answer a question in relation to the question length. In the interviews using both short and long questions, question length did not seem to affect the time duration of the respondent's answer. Interviews using only short questions, however, did obtain an average answer length which was slightly lower than that obtained in the other interviews, but this difference is considered to be inconsequential. This pattern of findings is not consistent with the results of other research, but no reason for the discrepancy can be given at this time.

Table 2 describes the effect question length had upon the number of questions for which one or more health items were reported. Again, the

Table 1. Number of respondents and average duration (in seconds) of answers per question, by question length, pilot study I

Question length	Number of respondents	Average duration of answers (in seconds)
Long questions in interviews with both long and short questions---	18	5.58
Short questions in interviews with both long and short questions---		5.69
Short questions in interviews with short questions only---	9	5.31

Table 2. Number of respondents and percent of questions for which any health information was reported, by question length, pilot study I

Question length	Number of respondents	Percent of questions for which information was reported
Long questions in interviews with both long and short questions---	18	40
Short questions in interviews with both long and short questions---		38
Short questions in interviews with short questions only---	9	29

results were somewhat different from those expected. In interviews using both long and short questions respondents gave affirmative answers [e] to 40 percent of the long questions and 38 percent of the short questions. Therefore, within an interview using questions of varying length, the length of a specific question did not appear to have any effect on reporting. However, interviews using only short questions obtained affirmative answers to only 29 percent of the questions. This amounts to a drop in the affirmative answer rate of about 25 percent.

Therefore, the pilot study suggested that while question length per se may have some effect on respondent answering behavior, it is the combination of long and short questions that yields the most health information. Within an interview using different question lengths, the length of an individual question has little or no influence on the answers as long as the question has been preceded by a series of long questions.

[e] "Affirmative answer" is defined as the reporting of at least one health item.

Effects of Reinforcement and Question Length on Reporting—Pilot Study II

In a second pilot study consideration was given to the effect of question length with and without reinforcement on the number of health items reported. Three procedures were tested: (1) short questions with reinforcement, (2) long questions [f] with reinforcement, and (3) long questions without reinforcement. The dependent variable was the number of questions for which at least one health item was reported.

Two experienced interviewers queried a total of 48 respondents residing in Jackson, Michigan. Respondents were sampled by a procedure essentially similar to that described for the first pilot study, while the reinforcement procedures used were similar to those described below for

[f] Actually, following the results of the first pilot study, a mixture of long and short questions was used. This procedure is given the abbreviated designation "long questions" for convenience of presentation.

the main study. However, only about 1 hour of training was devoted to the use of the reinforcement procedures, in contrast to the several days of instruction given for the main study.

Twelve questions were asked of all respondents. Since answer duration was not measured, the interviewers were instructed to probe and clarify when necessary to obtain answers which met the objectives of the questions.

The percent of questions for which at least one health item was obtained by the three interviewing procedures is shown in table 3. The data indicated that interviews in which long questions were used with or without reinforcement obtained more affirmative responses than did interviews with short, reinforced questions. The main finding was that the effects of reinforcement and long questions were additive—they could be combined into one interview procedure and would thus produce more affirmative answers than either reinforcement or long questions used alone.

Table 3. Percent of questions for which health information was reported, by interview procedure, pilot study II

Interview procedure	Percent of questions for which information was reported
Short questions with reinforcement	66
Long questions[1] with reinforcement	84
Long questions without reinforcement	72

[1]Consisted of a mixture of long and short questions.

METHODS AND PROCEDURES

This section describes in detail the design and procedures used in the main experimental study.

Experimental and Data Analysis Design

The research was designed as a 2 x 2 factorial experiment with one interview for half the respondents and two interviews for the other half (see figure 1). The data were analyzed separately for two education groups.

The independent variables used were reinforcement and question length. Reinterviews and respondent education served as mediating variables. The dependent variables included indexes of accuracy (under- and overreporting, total error, and net bias) and amount of reported health information. More detailed descriptions of these variables are given later in this section.

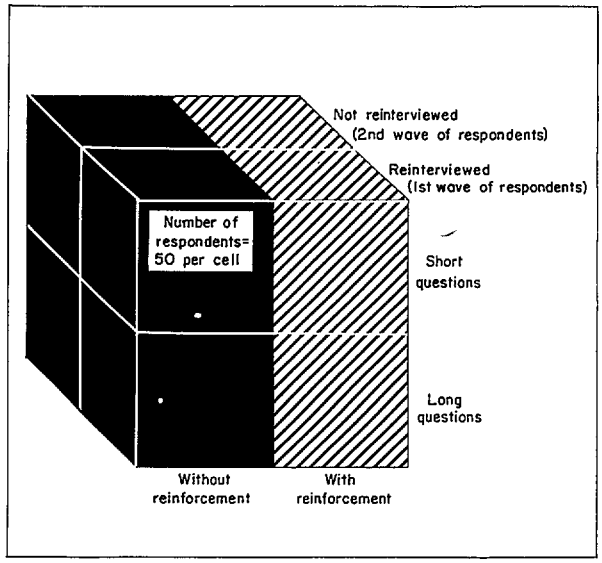

Figure 1. Experimental design.

Sample Design

Sample size.—It was determined from estimates of the frequency of occurrence of 13 selected chronic conditions in the U.S. population[7] and estimates of over- and underreporting rates[1,8] that approximately 600 respondents were needed to establish the reliability at the .05 level of a 25 percent difference on the dependent variable (due to the main effects of reinforcement and question length). It was then assumed that the number of respondents could be reduced somewhat because of the experimental controls and sample weighting described below. Since a sample size of 400 for the original interviews was within budget limitations, this number became the final sample size.

Respondents.—Respondents were selected from a population of persons who belonged to a prepaid health insurance plan and who came to one of the plan's several clinics between February and July 1968.

To reduce extraneous sources of variation in health reporting, a somewhat homogeneous population of respondents was selected. The sample was restricted to white females, aged 17-60, who were residents of the greater Detroit metropolitan area. In an attempt to increase the relative amount of reporting variance attributable to the experimental interviewing technique, a weighted selection of "sick" respondents was used. One-third of the persons in the original population had none of the special chronic conditions; two-thirds had one or more. But in the final sample chosen for the study, 88 percent of the women had at least one chronic condition and 12 percent had none of the conditions of interest. Most of the analyses which follow include only those respondents who had at least one of the chronic conditions.

Fifty percent of the respondents (half of each of the four experimental groups) selected for an interview were also scheduled for reinterviews. To facilitate data collecting, all respondents selected for original interviews during the first half of the data collection period were designated as the reinterview group. A small number of respondents whose scheduled original interviews came during the second half of this period were also reinterviewed in order to bring the reinterview sample sizes to approximately 50 persons for each of the four interviewing procedures.

Other controls.—In the sample design several variables were stratified or controlled. Respondents were assigned to 15 geographic area groups, four age groups, and four "sickness" groups. Sickness was defined by the number of chronic conditions checked "yes" on the Physician Summary Form (see appendix II).

Using these groupings, clusters with four respondents each were formed. The clusters were geographic, based on large areas corresponding roughly in size to one or more of the metropolitan census tracts. Within each of the clusters the four interviewing procedures were assigned at random, and within each interviewer's assignment an attempt was made to have a balanced variety of interviewing procedures to use as well as various age groups and sickness levels among the respondents interviewed. The design effect on the variances due to clustering is assumed to be zero and has not been calculated for this study.

While it would have been desirable to control interviewer variance by random or stratified assignments, this was not financially feasible as the interviewing area covered several hundred square miles. In this study interviewer variance might also be attributed to geographic location of the respondent.

All interviewers used all combinations of procedures (short questions, long questions, reinforcement, no reinforcement, no reinterview, and reinterview). Thus, interviewer differences were not seriously correlated with the effects of the interviewing procedures used. On the other hand, requiring interviewers to learn and use a variety of techniques probably introduced some error which might have reduced the contrast among the procedures.

With this compromise design and a less than optimal sample size it was expected that contrast between the experimental interviewing methods would be small. Specifically, the effects of reinforcement and question length might appear somewhat unstable; the reinterview effects would show up but be even more unreliable; and it would be difficult to detect and test for an interaction effect of the procedures.

Response Rates

The overall response rate, shown in table 4, was 90 percent for the original interview and 92 percent for the reinterview. The response rates for the four procedures ranged from 84 to 95 percent.

Interviewers and Training

Ten white female interviewers were employed in this research. Four had extensive experience with the Survey Research Center (SRC), two had been on the staff several months prior to this study, and four were new additions to the permanent staff. Of the last group one had extensive experience in market research interviewing. All 10 had training in basic SRC interviewing procedures prior to this study.

Interviewers received about 1 week of classroom training covering the basic definitions and procedures of health interviewing and were drilled in each of the four experimental interviewing methods. Considerable classroom work was directed toward acquiring a thorough understanding of the concepts underlying both the questionnaire and the techniques basic to all of these interviews. However, most of the classroom sessions were devoted to supervised role-playing interviews, after which immediate feedback was given by the research staff. The interviewers were not told that medical information was available about the respondents. Following classroom training, interviewers spent about 1½ weeks in practice interviewing and tape-recording of these practice interviews. Practice interview protocols and tapes were reviewed soon after completion, and the interviewer was given a critique of her performance shortly after the material was received. Comments were made, usually via telephone, and immediately afterward a written evaluation was sent to the interviewer along with her original protocols. Practice with reinterviews occurred 2 weeks after the start of interviewing and was handled in essentially the same way as the training for original interviews.

Table 4. Number of sample persons and response rates for original interviews and reinterviews, by interview procedure

Interview procedure	Original interview				Reinterview		
	Original sample	Eligible respondents	Respondents interviewed	Response rate	Eligible respondents	Respondents interviewed	Response rate
	Number of persons			Percent	Number of persons		Percent
All procedures--	511	448	404	90	224	205	92
Short questions:							
Without reinforcement----------	128	117	105	90	56	50	89
With reinforcement----------	127	107	99	93	53	49	92
Long questions:							
Without reinforcement----------	130	114	96	84	57	53	93
With reinforcement----------	126	110	104	95	58	53	91

Throughout the entire interviewing period for the study an attempt was made to observe each interviewer on at least two occasions each week. A member of the staff accompanied the interviewer during these two interviews, took systematic notes on strong and weak points of her interview performance, and transferred the notes to an Observation Form (see appendix II) for the interviewer's use. The form served as a basis for discussing the completed interview with the observer. After the interviewer finished studying the Observation Form, she was asked to return it to the central office for the records. Although observations were thought to be necessary and helpful, there was concern about their effect on the data, as about half the observations were made by males and both interviewer and respondent were female. Sex of observer was not recorded, but the data indicate no relationship between the presence or absence of an observer and the accuracy of reporting chronic conditions. This lack of relationship held true for all experimental procedure combinations within each respondent education group (see table 16).

Interview Techniques

A description of the interviewing techniques, reinforcement and question length, is given below. The basic features of each interviewing technique used were actually written into the questionnaires. Further detail regarding each procedure is available in the sample pages from the questionnaires in appendix I.

Reinforcement.—In those interviews where reinforcement was used, each time the respondent reported an instance of illness or health service utilization (up through question 14 of the original interview and through question 17 of the reinterview), the interviewer recorded the answer, looked up at the respondent, and used one of several reinforcing statements. After question 14 or 17, reinforcing statements were used but were worded and scheduled at the interviewer's discretion.

The reinforcing statements were printed in the following form at the top of each applicable page of the reinforcement questionnaires.

Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

The interviewer combined a statement in the first column with one in the second column to form a complete reinforcing statement. Infrequently, a large number of consecutive reinforcing statements were required, and the interviewer was given the option of occasionally using only a short statement from column one to avoid monotony.

A great deal of training and observation time was devoted to the problem of making the reinforcing statements sound natural. The interviewers eventually were able to do this satisfactorily. The natural effect would seem desirable for the interview situation, yet, there is some feeling in retrospect that it is not essential for successful application of the reinforcement procedure.

Question length.—The contrast in question length was basically between an interview using only short questions and an interview using a mixture of short and long questions, heavily weighted with the latter. The following examples

show how short questions about chronic conditions were lengthened:

Short questions	*Long questions*
a. Asthma?	a. Asthma is the first one we need information about. Have you had asthma?
b. Hay fever?	b. What about hay fever? Have you had that?
c. Thyroid trouble or goiter during the past 12 months?	c. Another area of interest to the Health Service is thyroid and goiter trouble. Have you had any thyroid trouble or goiter during the past 12 months?

For every lengthened question an attempt was made to add only words which neither changed the meaning of the question nor clarified what was wanted by the question.

For the long question interviews an attempt was made to intersperse short questions at different intervals, since pilot study data indicated this technique was effective. However, any probes that were used to follow up a main question were the same length for all interview procedures. Because of the nature of the probe questions, a respondent who volunteered information about illness in response to an open question was asked proportionately more of these short questions than one who either did not furnish information on the open questions or who reported illness only on the structured, checklist questions. This particular aspect of the design would tend to minimize the effect of question length on numbers of items reported—the more information reported, the more the long question interview became like the short question interview.

Reinterviews

Whenever possible, the reinterviews (with approximately half the respondents in the original sample) took place 2 weeks after the initial interview. The reinterview content was very similar to that of the original interview except for the addition of some "filler" questions about health insurance and a slight change in the order of the questions. An attempt was made to hold everything else constant.

Respondent Education

The educational level of the respondent was used as a mediating variable in the data analysis. The amount of education was coded from the respondent's answers to the following questions asked near the end of the original interview:

18. What was the highest grade you attended in school?

18a. Did you complete that grade?

Respondents who had not completed the 12th grade (had not graduated from high school) were classified as "less educated." Those who had completed the 12th grade or more were classified as "more educated."

Respondent education was not used as a stratification control in the sample design because this information was not available on the health records which served as a basis for sample selection. Therefore, respondent education was not distributed evenly among the experimental procedure groups. The analyses of variance employing respondent education as a control variable were calculated using an unweighted means solution in order to compensate for unequal cell sizes arising from the procedure x education sampling variation.

Dependent Variables

Several types of dependent variables were used for this study. The main emphasis was on the validity of reporting 13 selected chronic conditions. The other dependent variables were the average frequency of reporting a particular kind of health event and the correlations between the number of events reported on the original interview and on a reinterview.

Under- and overreporting of chronic conditions.—The basic dependent variables were indexes reflecting the degree of agreement be-

tween the respondent and a physician about the presence of each of 13 chronic conditions.

• Chronic condition items. The 13 chronic conditions were selected from those reported frequently in the Health Interview Survey. For this study they were included in a list of 19 conditions and placed approximately in the middle of the interview. The respondent was asked if she had any of the following during the preceding 12 months: asthma, chronic bronchitis, hay fever, chronic skin trouble, hemorrhoids, hernia, ulcer, and varicose veins. She was then asked if she had *ever* had arthritis, heart trouble, stroke, hypertension, or diabetes.

• Physician information. Information about the respondent's status on the 13 chronic conditions was obtained from the Physician Summary Form (see appendix II). For each of the 13 conditions, the physician checked one of these categories:

1. Definitely or probably present (scored as "yes" in data analysis)
2. Definitely or probably not present (scored as "no" in data analysis)
3. Don't know, no information

For each condition and for each respondent the data from the interview and from the Summary Form were combined and assigned to one of five match categories:

1. Respondent, yes - physician, yes
2. Respondent, no - physician, yes
3. Respondent, yes - physician, no
4. Respondent, no - physician, no
5. Missing data (a response other than "yes" or "no" from either source)

Comparing the respondent's answers to the physician data provided a measure of reporting validity. However, it should be pointed out that the physician data were assumed to be only approximations of the "truth" about respondent chronic conditions. There were several reasons for this: (1) Even though each physician filled out the Physician Summary Form immediately after the respondent left the clinic and the respondent's complete medical record was in the hands of the physician, time constraints undoubtedly prevented a thorough scrutiny of the record's contents; (2) some of the chronic condition diagnoses were tentative at the time the form was filled out; (3) physicians probably differed with respect to how certain a diagnosis must be before it was entered as a "probably present" on the Physician Summary Form; and (4) some time elapsed between the date the physician's form was filled out and the date of the household interview. In some cases the time interval was as great as 6 months. Both health and diagnostic precision could be expected to change within these time intervals.[g]

• Indexes of reporting error. For statistical purposes it was assumed that some proportion of respondent-physician disagreement was valid and that this kind of disagreement was not confounded with assignment to experimental interviewing procedure groups. It was also assumed that some of the disagreement represented respondent reporting error and that the extent of this error was reflected in an attenuated form

[g]Within the higher education group of respondents there was a slight negative correlation between the recency with which the physician filled out the Physician Summary Form for the respondent and reporting error. That is, for these respondents, the amount of reporting error was slightly less for those cases in which physician data were recent relative to the time of the interview. This relationship was extremely small, confined only to the more educated group of respondents, and was not confounded with any of the other experimental procedures. Because of the existence of this and other kinds of bias in the matching procedures, the reader should not attempt to interpret the error scores presented herein as exact or even close approximations of the true rates of misreporting of chronic conditions. On the other hand, the existence of this kind of bias, distributed evenly among procedures, does not preclude making meaningful comparisons among the reinforcement and question length experimental procedures for their relative effects on reporting error.

by the calculated mismatch rates described below.

Two basic mismatch rates were calculated based on the following scheme:

RESPONSE TO EACH CHRONIC CONDITION ITEM

Respondent says:	*Physician says:*	*Match category:*
Yes	Yes	A
No	Yes	B
Yes	No	C
No	No	D

(1) Underreporting, $\frac{B}{A+B}$; the number of conditions for which the respondent said "no" *and* the physician said "yes" divided by the total number of conditions for which the physician said "yes."

(2) Overreporting, $\frac{C}{C+D}$; the number of conditions for which the respondent said "yes" *and* the physician said "no" divided by the total number of conditions for which the physician said "no."

The underreporting and overreporting index scores were combined into two additional indexes of reporting error: an index of total error and an index of net bias. It was felt that these additional indexes gave the clearest picture of the underlying dynamics of recalling and reporting chronic condition information.

(3) Total error, $\frac{B}{A+B} + \frac{C}{C+D}$; the unweighted sum of the rate of underreporting and the rate of overreporting. In this study the opportunities to overreport were much greater than the opportunities to underreport. Therefore, the index of total error contained a "disproportionate" representation of observed underreporting errors. An index of total error which represents each type of error according to the actual frequency of occurrence, not computed here, would have the form $\frac{B+C}{A+B+C+D}$.

(4) Net bias, $\frac{B}{A+B} - \frac{C}{C+D}$; the rate of underreporting minus the rate of overreporting. Again, this index gave equal weight to both overreporting and underreporting errors, making it easier to talk about the underlying psychological processes involved in recall and reporting. However, it distorted the fact that opportunities for overreporting were much more frequent than for underreporting. A correctly weighted net bias score could be computed for this or any other single study as $\frac{A-C}{A+B+C+D}$.

These dependent variable index scores were proportionate scores. The underreporting and overreporting index scores could range from .00 to 1.00. The total error score had a possible range from 0.00 to 2.00. The net bias score had a theoretical range from -1.00 through 0.00 to 1.00. Practically speaking, however, the ranges of the total error and net bias scores were much less. In the absence of completely consistent, deliberate lying the underreporting and overreporting scores should not be expected to average above 0.50. Therefore, the practical range of the net bias score was between -0.50 and +0.50. Average scores much beyond these practical ranges would represent the phenomenon of respondents having available accurate information to report to the interviewer but reporting answers exactly opposite of the truth.

Signal detection theory,[9] developed in the field of sensory psychology, uses special statistical and experimental procedures which combine the total error score and the net bias score along with certain theoretical assumptions to make estimates of two psychological parameters involved in reporting: the accuracy or sensitivity of recall and the bias in the decision about how to report recalled information. By using certain experimental procedures and statistical techniques and by making certain assumptions about human sensation and memory, it is possible to estimate when an experimental interviewing procedure has actually changed the respondent's ability to recall information. This is opposed to the case in which an experimental technique has merely changed the respondent's decision to say "yes" or "no" when he is uncertain of the correct answer.

While it would have been desirable to make these two kinds of estimates for the experimental procedures in this study, the data necessary to make empirical estimates of the two parameters were not collected, and it was not feasible to make the necessary assumptions about human memory in order to derive the separate estimates from theory. Therefore, when changes in accuracy of recall were observed as a function of the independent variables in this research, it could not be determined if these changes had been mediated solely by a change in the bias (what to report when uncertain) parameter. Improvement in reporting accuracy may also (or solely) have been a function of improvement in the actual ability of respondents to recall chronic condition information. At present, it seems that the observed relationships between the independent and dependent variables provide useful knowledge despite the problem of pinpointing the exact mediating processes.

All cases for which the physician indicated that none of the 13 chronic conditions was present were deleted from the chronic condition (validity) analyses. Thus, the sample for whom the indexes of reporting accuracy and bias were computed consisted only of respondents who were potential over- or underreporters.

Quantity reported.—Another objective of this study was to ascertain the effect of reinforcement, question length, and reinterview upon the amount of illness and health service utilization reported. A description of the dependent variables used in these analyses is given here.

• Chronic conditions. In the original interview (questions 6 and 7) and reinterview (questions 9 and 10) there were two lists of chronic conditions. Together, they contained 19 chronic conditions, 13 of which were used in connection with the validity analysis described above. For analysis of quantity reported, the number of "yes" answers on both lists were combined to form a variable, regardless of whether answers were obtained to all 19 items.

• Total chronic and acute illness. All reports of injuries, present effects of old injuries, and other kinds of conditions, either chronic or acute and present within 2 weeks of the original interview or 4 weeks of the reinterview, were eligible for inclusion in this comprehensive variable. All instances of conditions which met the objectives of the question for which they were reported were included. Symptoms of such maladies were included only if the respondent could not report the underlying cause of the symptom. Otherwise, the cause of the symptom was included and the symptom itself was deleted. Multiple symptoms relating to the same cause, even if the cause was vague, were deleted and the cause retained. Also omitted were redundant reports of sickness, reports of normal menstruation and pregnancy, and symptoms reported only in question 1 (symptoms list). If the interviewer was in doubt about whether to retain an item, she asked a series of screening questions to determine eligibility. The screening questions on the originial interview are questions T1-T3 on the Original Interview Illness Table (see appendix II). They are questions T0-T3 on the Reinterview Illness Table (see appendix II). Answers to these questions were used to determine if the health item was retained.

• Symptoms. The first health question in both the initial interview and reinterview asked the respondent if she ever had had each of 19 health symptoms. The list may be found in appendix I. The respondent's "symptoms score" consisted of the number of "yes" answers given to items on the list. If one or more items for a respondent had answers which could not be classified either "yes" or "no," that respondent was excluded from data analysis for the entire symptom variable.

• Medicine and treatment taken in the past 2 weeks. In the original interview each respondent was asked if she had taken medicine or treatment for any condition during the past 14 days and, if so, what she took. At the end of the interview the respondent was again asked about medicines or treatments taken during the preceding 2 weeks. In coding, the total number of unduplicated medicines and treatments was recorded. Data on the numbers and names of medicines and treatments taken were not obtained in the reinterview.

• Physician visits. A complex set of procedures and definitions was used by interviewers to ascertain the number of times a respondent received the services of a physician during an approximate 3-month period.

For the initial interviews the respondent circled the dates on a calendar on which she

talked to a doctor or went to a doctor's office or clinic for herself during the 10-week period prior to the 2-week reference period of the interview. The interviewer cleared up any ambiguities or misunderstandings and confirmed the correct number of circles for the 10-week reference period. This number, subject to modification by the following series of probe questions, was the value the respondent received on the variable "physician visits, 10 weeks."

Because the definition of "use" in regard to physician services might not have been entirely clear to the respondent, the interviewer asked whether the respondent had done any of the following during the reference period: saw a doctor in an emergency room, saw a doctor at respondent's home, or talked to a doctor over the phone (excluding calls only to make appointments). If the respondent answered "yes" to any of these, she was asked for the date of the contact. The interviewer entered any new information on the calendar and corrected the numbers entered for the 10 weeks on the questionnaire.

The procedure was somewhat simplified for the reinterviews. The respondent was handed a calendar with the original 10 weeks outlined in blue and the 4 weeks following that in red. If the reinterview took place more than 2 weeks after the original, the 4 weeks outlined in red ended some time before the day of the reinterview. Thus, the recall reference period could differ for each respondent. Although data were obtained on physician contacts in the period between the original interview and the reinterview, these data were not used in the present study. The physician visit data are those for the 10-week reference period which ended 2 weeks prior to the original interview and at least 4 weeks prior to the reinterview.

• Dentist visits for 12 months. Original interview question 13 and reinterview question 17 were worded identically. In these questions the respondent was asked how many times she had gone to the dentist during the 12 months prior to the interview. Therefore, the days included in the 12-month reference periods differed between the original interview and the reinterview by about 2 weeks. For the reinterview score the respondent was asked to state how many dentist visits she made during the most recent 2 weeks. This number was subtracted from the total figure given for 12 months to obtain the reinterview dentist visit number. This resulted in the reinterview reference period being 11½ months and the original interview reference period 12 months. This difference is assumed to be inconsequential when original interview data are compared with reinterview data.

• Hospitalizations in 12 months. The respondent was asked how many times she had been in a hospital or nursing home overnight or longer during the 12 months preceding the interview. This information was obtained from question 12 in the original interview and question 13 for the reinterview. The equating of the 12-month reference periods was done in the same way as previously described for dentist visits.

EXPERIMENTAL RESULTS

Validity

Experimental treatment effects on the accuracy of chronic condition reporting in the original interview.—The data presented here indicate that reinforcement and question length had effects both on the error and on the bias involved in reporting of the 13 selected chronic conditions. The effects of the experimental variables were mediated by respondent education so that the variables which improved reporting for the less educated respondents interfered with good reporting for the more educated respondents. Conversely, the procedures which aided the reporting of the respondents with more education had a detrimental effect on the reporting of the respondents with less education.

The underreporting, overreporting, total error, and net bias index scores are given in tables 5-8 and are shown graphically in figures 2 and 3. The reader is cautioned that these dependent variable scores cannot be projected to any other population because a special sample was used in the study, and the physician "criterion" data are not to be considered as completely valid. These indexes were computed for the purpose of examining comparative effects of the different experimental interviewing procedures.

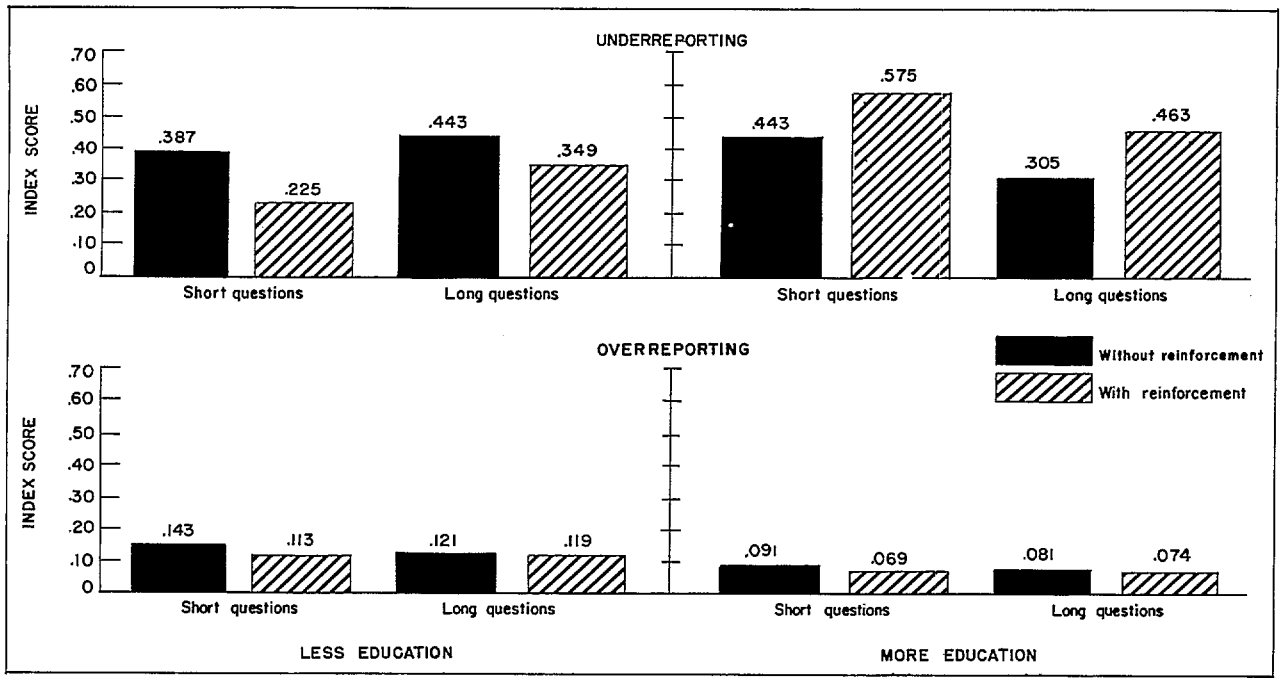

Figure 2. Index scores of underreporting and overreporting of chronic conditions in original interviews by reinforcement, question length, and respondent education.

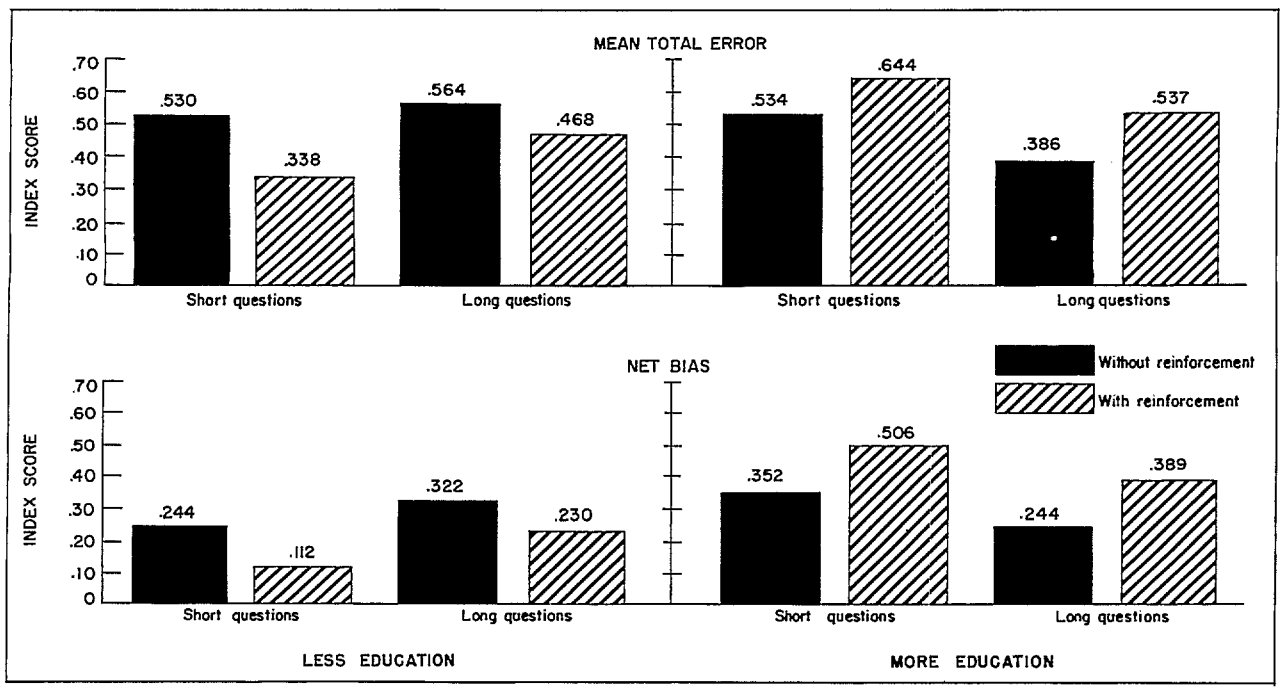

Figure 3. Mean total error and net bias index scores for the reporting of chronic conditions in original interviews by reinforcement, question length, and respondent education.

The underreporting index scores (table 5 and figure 2) indicate that, for respondents who had not completed high school, the use of reinforcement in interviews significantly reduced underreporting error. However, reinforcement significantly increased underreporting error for the more educated respondents. These data also indicate that question length affects underreporting. For less educated respondents long questions tended to increase underreporting (not statistically significant), while for more educated respondents the long questions significantly reduced underreporting. Although education interacted with reinforcement and question length to affect underreporting, within each education group the two experimental procedures had additive effects and did not interact to determine error.

The overreporting errors (table 6 and figure 2) were generally low (7 to 14 percent) and were much less affected by the experimental procedures. The small, nonsignificant trends in the data suggest that reinforcement reduced over-

Table 5. Number of respondents and index scores of underreporting of chronic conditions in original interviews, by reinforcement, question length, and respondent education

Interview procedure	Less education		More education	
	Number of respondents[1]	Index score	Number of respondents[1]	Index score
Short questions:				
Without reinforcement	42	.387	53	.443
With reinforcement	50	.225	39	.575
Long questions:				
Without reinforcement	44	.443	40	.305
With reinforcement	36	.349	46	.463

[1]Not included are three respondents for whom education was not ascertained and 51 persons for whom the physician indicated there were no chronic conditions present.

Summary of analyses of variance

Source of variation	Less education			More education		
	d.f.	M.S.	*F* value	d.f.	M.S.	*F* value
Reinforcement (R)	1	.696	[b]4.77	1	.921	[b]5.78
Question length (Q)	1	.342	2.34	1	.688	[b]4.32
R x Q	1	.049	0.33	1	.008	0.05
Error	168	.146		174	.159	

[b]$p \leq .05$.

NOTE: d.f. = degrees of freedom; M.S. = mean square.

reporting in both education groups. This trend runs counter to what might have been expected. One hypothesis was that reinforcement or reward on the part of the interviewer every time the respondent reports a sickness will result in the respondent "making up" sicknesses so that she will gain further approval of the interviewer. The data from this study indicate that the reinforcement procedure did not produce this effect. On the contrary, for both education groups, regardless of question length, the reinforcement interview tended to produce fewer false positive reports of chronic conditions than interviews which did not include reinforcement.

Table 7 and figure 3 present the total error index scores which represent the unweighted sum of the individual underreporting and overreporting index scores. Since underreporting scores had a higher mean and variance than overreporting scores, the total error scores tended to show the same effects of experimental procedures and respondent education

Table 6. Number of respondents and index scores of overreporting of chronic conditions in original interviews, by reinforcement, question length, and respondent education

Interview procedure	Less education		More education	
	Number of respondents[1]	Index score	Number of respondents[1]	Index score
Short questions:				
Without reinforcement	41	.143	53	.091
With reinforcement	50	.113	38	.069
Long questions:				
Without reinforcement	44	.121	40	.081
With reinforcement	36	.119	46	.074

[1]Not included are three respondents for whom education was not ascertained, 51 persons for whom the physician indicated there were no chronic conditions present, and two persons for whom physician data were incomplete.

Summary of analyses of variance

Source of variation	Less education			More education		
	d.f.	M.S.	F value[1]	d.f.	M.S.	F value[1]
Reinforcement (R)	1	.011	0.58	1	.009	1.19
Question length (Q)	1	.003	0.15	1	.000	0.04
R x Q	1	.008	0.41	1	.003	0.32
Error	167	.019		173	.008	

[1]None of the F values is statistically significant ($p \leq .05$).

NOTE: d.f. = degrees of freedom; M.S. = mean square.

Table 7. Number of respondents and mean total error index scores for the reporting of chronic conditions in original interviews, by reinforcement, question length, and respondent education

Interview procedure	Less education		More education	
	Number of respondents[1]	Index score	Number of respondents[1]	Index score
Short questions:				
Without reinforcement	41	.530	53	.534
With reinforcement	50	.338	38	.644
Long questions:				
Without reinforcement	44	.564	40	.386
With reinforcement	36	.468	46	.537

[1]Not included are three respondents for whom education was not ascertained, 51 persons for whom the physician indicated there were no chronic conditions present, and two persons for whom physician data were incomplete.

Summary of analyses of variance

Source of variation	Less education			More education		
	d.f.	M.S.	*F* value	d.f.	M.S.	*F* value
Reinforcement (R)	1	.934	[b]5.96	1	.749	[b]4.47
Question length (Q)	1	.205	1.59	1	.724	[b]4.32
R x Q	1	.115	0.73	1	.017	0.10
Error	167	.157		173	.168	

[b] $p \leq .05$.

NOTE: d.f. = degrees of freedom; M.S. = mean square.

observed in connection with underreporting. For less educated respondents, reinforcement significantly lowered total error while long questions increased it (the increase was not significant). For the more educated respondents, reinforcement significantly increased total error and long questions significantly reduced it. Within education groups, the experimental procedures did not interact but produced their effects as additive main effects.

The net bias index scores are shown in table 8 and figure 3. They indicate that higher educated respondents tended to exhibit more bias in the interview than respondents who had not graduated from high school. The effects of reinforcement and question length on net bias were also mediated by respondent education and were somewhat unstable for the less educated group. For the more educated group, long questions reduced bias and reinforcement increased

Table 8. Number of respondents and mean net bias index scores for the reporting of chronic conditions in original interviews, by reinforcement, question length, and respondent education

Interview procedure	Less education		More education	
	Number of respondents[1]	Index score	Number of respondents[1]	Index score
Short questions:				
Without reinforcement	41	.244	53	.352
With reinforcement	50	.112	38	.506
Long questions:				
Without reinforcement	44	.322	40	.244
With reinforcement	36	.230	46	.389

[1]Not included are three respondents for whom education was not ascertained, 51 persons for whom the physician indicated there were no chronic conditions present, and two persons for whom physician data were incomplete.

Summary of analyses of variance

Source of variation	Less education			More education		
	d.f.	M.S.	*F* value	d.f.	M.S.	*F* value
Reinforcement (R)	1	.574	3.32	1	1.123	[b]6.67
Question length (Q)	1	.366	2.12	1	0.664	[b]3.94
R x Q	1	.026	0.15	1	0.001	0.01
Error	167	.173		173	0.169	

[b]$p \leq .05$.

NOTE: d.f. = degrees of freedom; M.S. = mean square.

it. For the less educated group, the trends were in the opposite direction; long questions increased bias and reinforcement reduced it.

To summarize, the more educated group tended to underreport at a higher rate and exhibit a higher level of total error and more net bias than did the less educated group. It should be noted, however, that when conventional interviewing procedures were used (short questions without reinforcement), the rounded total error scores were the same for both education groups—.53. It was only when the experimental procedures were introduced that the groups showed different levels of response error and bias. It can also be seen that the two groups achieved the same low levels of error under interviewing conditions most appropriate (on the basis of findings in this study) to their education levels: the lowest average total error score was .34 for the less educated group,

achieved by using short questions and reinforcement; the lowest average total error achieved by the more educated group was .39, a result of the long question, no reinforcement interview procedure.

Experimental treatment effects on the accuracy of chronic condition reporting in the reinterview.—A second interview covering the same topics as the original interview was conducted for about half (205) of the respondents. The reinforcement and question length procedures to which the respondent was exposed originally were repeated during the second interview. The data presented here show the cumulative effects of the reinforcement and question length procedures. Because of the small sample sizes, the effects were not statistically significant.

The trends in the reinterview data were, in general, similar to those observed in connection with the original interview: for less educated respondents, reporting accuracy was generally improved by reinforcement; for more educated respondents, accuracy was generally increased by long questions. The best reporting performance for the less educated group was obtained by using short questions with reinforcement. For the more educated group, the best reporting was obtained by the interviewer using long questions without reinforcement. The average index scores on the four measures of variation for the eight experimental groups are given in table 9. Reinforcement reduced underreporting for less educated respondents and increased it for more educated respondents. Long questions tended to increase underreporting among the less educated group but decrease it among the more educated group, especially in the case when reinforcement was not used. In general, average underreporting was relatively higher for the more educated respondents.

The overreporting data were quite unreliable but suggested a somewhat different picture from that observed in connection with the original interview. Because of the extremely low F values, these conclusions should be considered highly tentative. As in the original interview, for less educated respondents overreporting in the reinterview was highest in the short question, no reinforcement group. The other combinations of procedures obtained approximately the same level of overreporting (lower than the short question without reinforcement procedure in the less educated group but higher than any of the combinations for more educated respondents). However, in contrast to the original interview, the reinterview data suggest a possible interaction between reinforcement and question length on overreporting among both the more educated and less educated respondents. In the reinterview, reinforcement tended to lower overreporting when short questions were asked and to raise it when long questions were used.

The total error data suggest that the previously observed independent main effects of reinforcement and question length were similarly operative here for the less educated respondents. Reinforcement lowered total error for this group and long questions increased it. The two variables were additive and did not interact. For the more educated group, reinforcement and question length appeared to interact in determining total error. The long question without reinforcement procedure produced an extremely low level of total error in comparison with all other combinations, which produced total error at approximately the same level (between .603 and .615).

The effects of the reinforcement, question length, and respondent education variables on net bias were similar to those observed previously in connection with underreporting and total error. For the less educated respondents, the net bias index score declined when reinforcement was used and increased when long questions were used. There appeared to be no interaction between these two experimental interviewing procedures within the less educated group. Net bias scores were uniformly high for the more educated group in all interview procedure combinations except long questions without reinforcement. In this procedure the net bias score was about half of what it was in the other procedures (table 9).

Thus, the cumulative effects of the experimental interviewing procedures of reinforcement and long questions were essentially the same in the reinterview as in the original interview. The one possible exception to this trend was that in reinterviews of the more educated group of

Table 9. Number of respondents and underreporting, overreporting, total error, and net bias index scores for the reporting of chronic conditions in reinterviews, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Index score for chronic condition reporting			
		Under-reporting	Over-reporting	Total error[2]	Net bias[2]
Less education					
Short questions:					
Without reinforcement	20	.400	.166	.566	.234
With reinforcement	25	.230	.131	.361	.099
Long questions:					
Without reinforcement	22	.504	.121	.625	.383
With reinforcement	21	.380	.140	.521	.240
More education					
Short questions:					
Without reinforcement	27	.495	.120	.615	.375
With reinforcement	20	.521	.082	.603	.438
Long questions:					
Without reinforcement	26	.286	.086	.373	.200
With reinforcement	24	.509	.099	.608	.410

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.
[2]May not add or subtract exactly due to rounding error.

Summary of F values from analyses of variance[1]

Dependent variable	Source of variation for respondents with less education[2]			Source of variation for respondents with more education[3]		
	Reinforcement (R)	Question length (Q)	R x Q	Reinforcement (R)	Question length (Q)	R x Q
Underreporting	3.04	2.27	0.08	2.10	1.66	1.32
Overreporting	0.07	0.40	0.92	0.54	0.25	2.18
Total error	3.33	1.66	0.36	1.66	1.87	2.02
Net bias	2.24	2.43	0.00	2.42	1.34	0.70

[1]None of the F values is statistically significant ($p \leq .05$).
[2]Degrees of freedom = 1,84.
[3]Degrees of freedom = 1,93.

respondents, the beneficial effects of long questions were obtained only when reinforcement was not used. When reinforcement was combined with long questions in the reinterview, underreporting, total error, and net bias index scores were at approximately the same high levels as when long questions were not used.

Validity of reinterview data compared with original interview data.— In order to compare the various indexes of reporting obtained in the original interviews and reinterviews, data in tables 10 and 11 are presented for only those respondents who were included in both interviews. One of the main conclusions from these data was that reinterviews tended to reduce the amount of underreporting error made by the more educated respondents. Reinterviews did not reduce underreporting error (in fact, tended to increase it) for less educated respondents. There were some exceptions to these trends: the short question with reinforcement reinterview did not produce an increase in underreporting with less educated respondents; and both the short question without reinforcement reinterview and the long question with reinforcement reinterview did not reduce underreporting error for the more educated respondents. Finally, although reinterviews did help the more educated respondents reduce underreporting error, this type of error was still generally higher in their reinterviews than in those of the less educated group.

The effects of the reinterview on overreporting error were less clear. The data indicated that overreporting index scores tended to increase in reinterviews for both education groups when the standard type of interviewing procedure (short questions without reinforcement) was used. The trend was in the same direction for most of the other types of reinterviews. From these data it seems reasonable to conclude that overreporting error was not reduced in an important way and showed a tendency to increase when reinterviews were employed.

As mentioned above with respect to underreporting, reinterviews brought about a slight improvement for higher educated respondents who were interviewed using either one of two of the experimental procedures. However, when interviewing was conducted by the conventional procedure (short questions without reinforcement), the reinterview did not produce better data than did the original interview regardless of the respondent's educational level. Since these findings are contrary to what might be expected, the next section of this report explores the possibility that original interview and reinterview data may be combined in some way to produce lower individual and total error index scores than would be obtained from original interview data only.

The comparative total error and net bias index data for original interviews and reinterviews are presented in table 11. Total error was not reduced in any of the four experimental groups in reinterviews with the less educated respondents. Between the two interviews the total error increased most when the short question without reinforcement procedure was used. There were intermediate levels of increase when long questions were used either with or without reinforcement. Almost no change occurred when the short question with reinforcement procedure was used. The lower educated group showed the largest increase in net bias scores in reinterviews using long questions. However, as for the total error scores, there was almost no change in net bias scores when the short question with reinforcement procedure was used in the reinterview.

For the group of respondents with more education, the data showed that total error was reduced in reinterviews for two of the four experimental procedures. When short questions and reinforcement were used, total error in the reinterview was .60 compared to a score of .70 for this same group in the original interview. The most interesting result, however, concerns the reduction in total error for the long question without reinforcement procedure. This procedure yielded the lowest rate of total error for these respondents in the original interview (.46). When these respondents were reinterviewed using long questions without reinforcement, the total error dropped to .37. This is the lowest average rate of total error for any of the experimental procedures for the higher educated group. It is almost identical to the lowest rate

Table 10. Number of respondents and underreporting and overreporting index scores for the reporting of chronic conditions in original interviews and reinterviews, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Underreporting index score			Overreporting index score		
		Original interview	Reinterview	Difference	Original interview	Reinterview	Difference
Less education							
Short questions:							
Without reinforcement	20	.346	.400	-.054	.134	.166	-.032
With reinforcement	25	.230	.230	.000	.126	.131	-.005
Long questions:							
Without reinforcement	22	.452	.504	-.052	.131	.121	+.010
With reinforcement	21	.331	.380	-.049	.138	.140	-.002
More education							
Short questions:							
Without reinforcement	27	.488	.495	-.008	.097	.120	-.023
With reinforcement	20	.638	.521	+.117	.065	.082	-.017
Long questions:							
Without reinforcement	26	.364	.286	+.078	.094	.086	+.008
With reinforcement	24	.454	.509	-.055	.090	.099	-.009

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

Summary of F values from analyses of variance[1]

Variable	Source of variation for respondents with less education[2]			Source of variation for respondents with more education[3]		
	Reinforcement (R)	Question length (Q)	R x Q	Reinforcement (R)	Question length (Q)	R x Q
Underreporting:						
Original interview	2.03	1.55	0.00	1.90	3.15	0.12
Reinterview	3.04	2.27	0.08	2.10	1.66	1.32
Overreporting:						
Original interview	0.00	0.02	0.07	1.04	0.37	0.64
Reinterview	0.07	0.40	0.92	0.54	0.25	2.18

[1]None of the F values is statistically significant ($p \leq .05$).
[2]Degrees of freedom = 1,84.
[3]Degrees of freedom = 1,93.

Table 11. Number of respondents and total error and net bias index scores for the reporting of chronic conditions in original interviews and reinterviews, by reinforcement, question length, and respondent education

Respondent education and interview prodedure	Number of respondents[1]	Total error index score			Net bias index score		
		Original interview	Reinterview	Difference	Original interview	Reinterview	Difference
Less education							
Short questions:							
Without reinforcement--	20	.480	.566	-.086	.212	.234	-.022
With reinforcement-----	25	.356	.361	-.005	.104	.099	+.005
Long questions:							
Without reinforcement--	22	.584	.625	-.041	.322	.383	-.061
With reinforcement-----	21	.470	.521	-.051	.192	.240	-.048
More education							
Short questions:							
Without reinforcement--	27	.585	.615	-.030	.390	.375	+.015
With reinforcement-----	20	.703	.603	+.100	.572	.438	+.134
Long questions:							
Without reinforcement--	26	.457	.373	+.084	.270	.200	+.070
With reinforcement-----	24	.543	.608	-.065	.364	.410	-.046

[1] Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

Summary of F values from analyses of variance[1]

Variable	Source of variation for respondents with less education[2]			Source of variation for respondents with more education[3]		
	Reinforcement (R)	Question length (Q)	R x Q	Reinforcement (R)	Question length (Q)	R x Q
Total error:						
Original interview-	1.83	1.52	0.00	1.41	2.78	0.03
Reinterview--------	3.33	1.66	0.36	1.66	1.87	2.02
Net bias:						
Original interview-	1.79	1.26	0.01	2.28	3.26	0.23
Reinterview--------	2.24	2.43	0.00	2.42	1.34	0.70

[1] None of the F values is statistically significant ($p \leq .05$).
[2] Degrees of freedom = 1,84.
[3] Degrees of freedom = 1,93.

of total error achieved by the "best" group of less educated respondents in the reinterview (short questions with reinforcement—.36). Reinterviews of the more educated group decreased the net bias score for three of the interviewing procedures but increased it for the long question with reinforcement procedure. Net bias was lowest in the long question without reinforcement reinterview, the procedure which also produced the lowest rate of total error.

The data suggest that meaningful improvement in reporting accuracy can be obtained by reinterviewing more educated respondents if long questions are asked and reinforcement is not used. Reinterviews with less educated respondents did not improve reporting accuracy.

Special combinations of original interview and reinterview data.— Analyses were made of various ways of combining data from original interview and reinterview sources to produce error and bias rates which were lower than those from the best of either source alone. Results of two ways of combining data from both interviews are shown in this section: a procedure which attempts to minimize underreporting or false negative errors and a procedure designed to minimize overreporting or false positive errors.

- Combination to minimize false negative errors. As long as the respondent did not say "no" when asked about the presence of a selected chronic condition in *both* the original interview and the reinterview, the response to that item was scored "yes." A "yes" was scored if the respondent said "no" in the original interview and "yes" in the reinterview, "yes" in the first interview and "no" in the reinterview, "yes" in both interviews, or "yes" in one interview and "don't know" in the other interview. However, a respondent who said "don't know" in both interviews was not scored "yes."
- Combination to minimize false positive errors. Data were also combined so that a respondent received a score of "no" on any chronic condition item unless she reported the condition in *both* interviews.

Data using these scoring methods were matched against the information provided by the physician, and the four dependent variable index scores were recalculated. The underreporting and overreporting data using each of these two combination scoring procedures are shown in table 12. Table 13 contains the total error and net bias scores for the two combination approaches. In tables 14 and 15 these total error and net bias scores are compared with those from the original interviews and reinterviews separately.

Data in tables 12 and 13 indicate that the recalculation of the dependent variable scores did not alter the previously observed effects of respondent education and interviewing procedure on reporting errors. As expected, the statistical procedure designed to minimize underreporting or false negative error produced lower underreporting index scores than those calculated by a procedure intended to minimize overreporting or false positive errors. The same kind of phenomenon occurred with respect to the overreporting index data. The calculation procedure designed to minimize overreporting errors yielded lower average overreporting index scores than did the procedure designed to minimize underreporting errors.

In order to understand the net effects of these apparent tradeoffs between overreport and underreport data, it is necessary to look at the total error index scores and the net bias index scores (table 13). With one minor exception, the calculation which minimized false negative reports produced lower total error index scores and lower net bias scores than did the calculation which reduced false positive errors. While the latter procedure reduced overreporting error in the data, it did so at the expense of introducing more than a compensating amount of false negative errors. Regardless of which of the two calculation procedures were used, the relative effects of reinforcement, question length, and respondent education on total error and net bias remained as observed in previous analyses: reinforcement improved the reporting of less educated respondents and hindered that of the more educated respondents; the long question interview interfered with optimal performance in the less educated group and promoted it in the more educated group.

The fact that net bias scores were lowest when data from the original interviews and reinterviews were combined to minimize false negative reporting seems to reflect the possi-

Table 12. Number of respondents and index scores of underreporting and overreporting of chronic conditions in original interviews and reinterviews (combined to minimize false negative and false positive errors), by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Underreporting index score		Overreporting index score	
		Combined to minimize false negative errors	Combined to minimize false positive errors	Combined to minimize false negative errors	Combined to minimize false positive errors
Less education					
Short questions:					
Without reinforcement	20	.333	.412	.172	.129
With reinforcement	25	.217	.243	.146	.112
Long questions:					
Without reinforcement	22	.417	.532	.147	.107
With reinforcement	21	.283	.428	.152	.129
More education					
Short questions:					
Without reinforcement	27	.432	.551	.128	.092
With reinforcement	20	.521	.638	.083	.065
Long questions:					
Without reinforcement	26	.286	.364	.105	.075
With reinforcement	24	.454	.509	.109	.080

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

Summary of F values from analyses of variance[1]

Variable	Source of variation for respondents with less education[2]			Source of variation for respondents with more education[3]		
	Reinforcement (R)	Question length (Q)	R x Q	Reinforcement (R)	Question length (Q)	R x Q
Underreporting:						
Combined to minimize false negative errors	2.34	0.85	0.01	2.29	1.59	0.21
Combined to minimize false positive errors	2.52	3.11	0.14	1.76	3.25	0.11
Overreporting:						
Combined to minimize false negative errors	0.11	0.08	0.25	1.17	0.01	1.74
Combined to minimize false positive errors	0.01	0.01	0.52	0.46	0.00	1.01

[1]None of the F values is statistically significant ($p \leq .05$).
[2]Degrees of freedom = 1,84.
[3]Degrees of freedom = 1,93.

Table 13. Number of respondents and total error and net bias index scores for the reporting of chronic conditions in original interviews and reinterviews (combined to minimize false negative and false positive errors), by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Total error index score		Net bias index score	
		Combined to minimize false negative errors	Combined to minimize false positive errors	Combined to minimize false negative errors	Combined to minimize false positive errors
Less education					
Short questions:					
Without reinforcement	20	.505	.542	.162	.283
With reinforcement	25	.362	.356	.071	.131
Long questions:					
Without reinforcement	22	.564	.639	.270	.425
With reinforcement	21	.436	.556	.131	.299
More education					
Short questions:					
Without reinforcement	27	.560	.642	.304	.459
With reinforcement	20	.604	.703	.438	.572
Long questions:					
Without reinforcement	26	.392	.439	.182	.288
With reinforcement	24	.563	.589	.344	.429

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

Summary of *F* values from analyses of variance[1]

Variable	Source of variation for respondents with less education[2]			Source of variation for respondents with more education[3]		
	Reinforcement (R)	Question length (Q)	R x Q	Reinforcement (R)	Question length (Q)	R x Q
Total error:						
Combined to minimize false negative errors	2.56	0.80	0.01	1.74	1.59	0.58
Combined to minimize false positive errors	2.36	3.33	0.35	1.48	3.20	0.26
Net bias:						
Combined to minimize false negative errors	1.61	1.02	0.07	2.78	1.31	0.03
Combined to minimize false positive errors	2.27	3.21	0.02	2.04	2.91	0.02

[1]None of the *F* values is statistically significant ($p \leq .05$).
[2]Degrees of freedom = 1,84.
[3]Degrees of freedom = 1,93.

bility that a "no"-saying response bias—the tendency to deny to the interviewer the presence of existing health conditions—was a fairly predominant characteristic in chronic condition reporting for all respondents. Procedures which reduced the effects of this tendency to underreport, whether they were experimental interviewing strategies or a special statistical strategy, yielded lower net bias scores. Inspection of these data also revealed that the lowest net bias scores within education groups were obtained by the same experimental combination of interview procedures which produced the lowest total error scores.

Tables 14 and 15 present, for summary purposes, the main response error and response bias findings of this research. Here the relative effects of the major analysis variables (reinforcement, question length, respondent education, reinterviews, and special statistical procedures for minimizing error) can be seen.

Table 14. Number of respondents and total error index scores for the reporting of chronic conditions in original interviews, reinterviews, and in original interview-reinterview data combined to minimize false negative and false positive errors, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Total error index score			
		Original interview	Reinterview	Combined to minimize false negative errors	Combined to minimize false positive errors
Less education					
Short questions:					
Without reinforcement	20	.480	.566	.505	.542
With reinforcement	25	.356	.361	.362	.356
Long questions:					
Without reinforcement	22	.584	.625	.564	.639
With reinforcement	21	.470	.521	.436	.556
More education					
Short questions:					
Without reinforcement	27	.585	.615	.560	.642
With reinforcement	20	.703	.603	.604	.703
Long questions:					
Without reinforcement	26	.457	.373	.392	.439
With reinforcement	24	.543	.608	.563	.589

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

NOTE: Lowest scores for each procedure are underscored.

Within each education group, the range of total error and net bias index scores was quite wide, indicating that the way in which the interview was conducted and the method of data analysis had particularly large effects on the quality of survey results. The highest total error index score was almost twice as high as the lowest total error index score. Within education groups, the highest net bias index score was about six times greater than the smallest net bias index score.

The data also indicate that the experimental procedures interacted or combined in different ways to influence reporting quality. The only main effect observed was that combining the data from first and second interviews to minimize false negative reporting (underreporting) always resulted in the lowest average net bias scores. This was true for all interviewing procedures, regardless of respondent education.

Table 15. Number of respondents and net bias index scores for the reporting of chronic conditions in original interviews, reinterviews, and in original interview-reinterview data combined to minimize false negative and false positive errors, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents[1]	Net bias index score			
		Original interview	Reinterview	Combined to minimize false negative errors	Combined to minimize false positive errors
Less education					
Short questions:					
Without reinforcement	20	.212	.234	.162	.283
With reinforcement	25	.104	.099	.071	.131
Long questions:					
Without reinforcement	22	.322	.383	.270	.425
With reinforcement	21	.192	.240	.131	.299
More education					
Short questions:					
Without reinforcement	27	.390	.375	.304	.459
With reinforcement	20	.572	.438	.438	.572
Long questions:					
Without reinforcement	26	.270	.200	.182	.288
With reinforcement	24	.364	.410	.344	.429

[1]Includes only respondents who were reinterviewed, but excludes two respondents for whom education was not ascertained and 18 persons for whom the physician indicated there were no chronic conditions present.

NOTE: Lowest scores in each procedure are underscored.

This method of combining original interview and reinterview data also resulted in relatively low total error index scores within experimental groups, but only when a question length procedure was used which was found to be least effective for the education group. Thus, the analysis minimizing false negative error yielded lowest total error index scores among less educated respondents asked long questions and among more educated respondents asked short questions.

If one wished to minimize the total error index scores for chronic condition reporting, he would interview less educated respondents only once, using reinforcement and short questions. He would interview more educated respondents twice, using long questions without reinforcement and disregarding data from the original interview. Were these procedures to be followed, the total error rates for each education group would be about the same—.36 for less educated respondents and .37 for more educated respondents (table 11). It is interesting to note, however, that if less educated respondents were interviewed using the procedure optimal for more educated respondents, and vice versa, maximum total error rates (almost twice as high) would result.

One additional conclusion from these optimal procedure data may be that better data would result from redesigning procedures of interviewing rather than trying to introduce statistical corrections into survey data already collected.

Other correlates of accuracy and bias of chronic condition reporting.—In addition to the effects of education, reinforcement, question length, and reinterviews, the data from this study may be inspected for other correlates of reporting error and bias. The relationship of reliability and of number of items reported to validity will be discussed in the next section. In this section, the effects of personal characteristics of respondents and of various administrative considerations on four of the dependent variables are explored.

Of the five personal characteristics shown in table 16, only "health rating" revealed a consistent pattern of effects and then only for the less educated respondents. Respondents rated themselves on a four-point scale describing their general state of health as either excellent, good, fair, or poor. The data indicated that those less educated respondents who rated themselves toward the "poor" end of the scale (essentially those who did not claim excellent health): (1) underreported less, (2) overreported more, (3) had lower total error scores, and (4) lower net bias ("no"-saying bias) scores. The direction of these effects was similar for the more educated respondents, but only the effect on net bias was statistically significant.

The second part of table 16 shows the correlations of the dependent variables with various administrative variables. One would hope that none of the correlations would be significant, and almost none were. As mentioned earlier, there was a negative relationship between the length of time since the physician record was made and total error for the more educated group. This indicates that slightly less reporting error was detected when the physician data were recent. The effect was small and was not confounded with interview procedure effects. The only other linear relationship which met the criterion of statistical significance was the positive relationship between length of interview and overreporting in the less educated group. Longer interviews apparently had a slightly greater tendency to contain chronic condition overreports than did shorter interviews if the respondent had not completed high school.

It is interesting to note the lack of linear influence of the other variables on error and bias. It apparently made no important difference if the interview was observed, if the interviewer had to make a number of callbacks, or if the interview occurred in a particular week. Respondent reporting error was not influenced by the number of conditions the person had according to the physician. This lack of effect increases the possibility of generalizing the present findings to a more representative cross section of the population in which the incidence of occurrence of the 13 selected chronic conditions is considerably lower than that in this research sample.

Table 16. Product-moment correlation coefficients between characteristics of respondents in original interviews and the underreporting, overreporting, total error, and net bias index scores by respondent education

Variable	Less education				More education			
	Under-reporting	Over-reporting	Total error	Net bias	Under-reporting	Over-reporting	Total error	Net bias
Personal characteristics:								
Age	-.05	.07	-.02	-.06	-.03	.10	-.01	-.05
Education	-.01	.01	.01	.00	.00	-.05	-.01	.01
Income	.01	.10	.04	-.02	-.05	-.01	-.06	-.05
Number of children	.04	.07	.06	.01	.00	-.08	.01	.02
Health rating	[a]-.29	[a].21	[a]-.20	[a]-.34	-.16	.18	-.12	[a]-.20
Administrative variables:								
Interval since Physician Summary Form was filled out	.04	.14	.09	-.01	.16	[a]-.21	[a]-.20	-.11
Number of conditions on Physician Summary Form	.05	-.02	.04	.05	.13	-.08	.11	.14
Length of interview	-.12	[a].22	-.04	-.18	-.09	.17	-.06	-.13
Week of interview	.00	.00	.01	.01	-.05	.10	-.05	-.05
Number of calls to obtain interview	.10	-.03	.09	.11	.12	-.04	.11	.12
Observer present at interview	-.07	-.03	-.08	-.06	.02	.05	.03	.01

[a] $p \leq .01$.

Reliability

Effects of experimental interviewing procedures on reliability of interview data.— Since the same questions were asked in the original interview and the reinterview, it is possible to examine the effects of reinforcement and question length on reliability.

The following data approach reliability from two different points of view: (1) for chronic conditions (the 19-item list and the 13-item subset which was validated) and symptoms, the calculations are in terms of the consistency of reporting the individual items on the lists; and (2) for the other health variables (doctor contacts, dentist visits, hospital episodes, total chronic and acute conditions, and health rating) and secondary variables (height, weight, education, and income), a Pearson product-moment correlation coefficient is reported. The coefficient is based on the total number of items reported in each interview, regardless of the consistency of reporting each item making up the total. However, the latter approach may mask the existence of compensating errors. For example, three physician visits might be reported in both interviews even though they might not be the "same" three visits in both interviews.

Tables 17 and 18 present the average number of responses in the original interview and reinterview match categories in each education group for the interviewing procedures for reporting of items on the symptom list and on the chronic condition list. In the far right column of the tables is the overall original interview-reinterview agreement rate; the other agreement rate columns contain two other coefficients called "yes" match rate and "no" match

Table 17. Number of respondents, number of reported symptoms by original interview-reinterview match category, and rates of original interview-reinterview agreement, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents	Total number of symptom items[1]	Average number of symptom items reported					Original interview-reinterview agreement rates[3]		
			Missing data	Match category A[2]	Match category B[2]	Match category C[2]	Match category D[2]	"Yes" match rate, $\frac{A}{A+B+C}$	"No" match rate, $\frac{D}{B+C+D}$	Overall agreement rate, $\frac{A+D}{A+B+C+D}$
Less education										
Short questions:										
Without reinforcement	21	19.01	0.05	6.57	1.29	1.24	9.86	.722	.796	.867
With reinforcement	26	19.02	0.08	5.85	1.35	1.12	10.62	.703	.811	.870
Long questions:										
Without reinforcement	23	19.00	0.04	6.09	1.65	1.61	9.61	.651	.747	.828
With reinforcement	25	19.00	0.28	5.56	1.44	1.20	10.52	.678	.799	.859
More education										
Short questions:										
Without reinforcement	29	18.99	0.03	5.55	1.21	1.41	10.79	.679	.805	.862
With reinforcement	23	19.00	0.17	3.78	0.83	1.87	12.35	.583	.821	.857
Long questions:										
Without reinforcement	28	19.00	0.00	4.79	0.71	1.86	11.64	.651	.819	.865
With reinforcement	28	19.01	0.61	4.04	0.96	1.36	12.04	.635	.838	.875

[1]Symptom items may not add to 19 due to rounding error.

[2]Codes for original interview-reinterview match categories are as follows:

Match category	Respondent answer in original interview	Respondent answer in reinterview
A	Yes	Yes
B	No	Yes
C	Yes	No
D	No	No

[3]For further discussion of agreement rates, see p. 33.

Table 18. Number of respondents, number of reported chronic conditions by original interview-reinterview match category, and rates of original interview-reinterview agreement, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of respondents	Total number of chronic condition items[1]	Average number of chronic condition items reported					Original interview-reinterview agreement rates[3]		
			Missing data	Match category A[2]	Match category B[2]	Match category C[2]	Match category D[2]	"Yes" match rate, $\frac{A}{A+B+C}$	"No" match rate, $\frac{D}{B+C+D}$	Overall agreement rate, $\frac{A+D}{A+B+C+D}$
Less education										
Short questions:										
Without reinforcement	21	13.01	0.10	2.38	0.48	0.14	9.91	.793	.941	.952
With reinforcement	26	13.00	0.04	2.65	0.23	0.19	9.89	.863	.959	.968
Long questions:										
Without reinforcement	23	13.00	0.09	2.39	0.26	0.52	9.74	.754	.926	.940
With reinforcement	25	13.00	0.04	2.28	0.16	0.32	10.20	.826	.955	.963
More education										
Short questions:										
Without reinforcement	29	13.01	0.07	1.76	0.35	0.17	10.66	.772	.954	.960
With reinforcement	23	13.00	0.13	1.04	0.35	0.00	11.48	.748	.970	.973
Long questions:										
Without reinforcement	28	13.00	0.00	1.86	0.25	0.18	10.71	.812	.961	.967
With reinforcement	28	13.01	0.11	1.75	0.18	0.18	10.79	.829	.968	.972

[1]Chronic conditions included are the 13 validated items. Items may not add to 13 due to rounding error.

[2]Codes for original interview-reinterview match categories are as follows:

Match category	Respondent answer in original interview	Respondent answer in reinterview
A	Yes	Yes
B	No	Yes
C	Yes	No
D	No	No

[3]For further discussion of agreement rates, see p. 33.

rate. The "yes" match rate is the average number of items for which the respondent said "yes" in both interviews divided by the number of items for which the respondent said "yes" in either interview. The "no" match rate was computed in a similar manner for the "no" answers to items on the symptom and chronic condition lists.

The tables indicate only small differences among the interview procedures on the consistency of respondent reports. The rates of agreement were all very high—higher for chronic condition reporting than for symptom reporting and higher for "no" answers than for "yes" answers. The data for chronic condition reporting show that differences among interviewing procedures were not very large, indicating that reinforcement and question length did not have especially strong effects on consistency of illness reporting.

For the less educated group of respondents, reinforcement tended to be associated with higher coefficients of agreement. Long questions tended to produce lower rates of consistency than did short questions. Thus, within the less educated group of respondents, the effects of the experimental procedures on consistency of reporting followed the same pattern as the effects on the validity of reporting.

For more educated respondents the effects of the interviewing procedures on consistency were not identical to the effects on validity. While the differences in consistency among interviewing procedures were small, the general pattern for the "no"-match rate and the overall agreement rate was the production of higher levels of consistency between interviewing procedures in responding to specific items when reinforcement was used. Thus, it appears that reinforcement produced slightly more reliable symptom and chronic condition reporting in both education groups.

In table 19 product-moment correlation coefficients are presented showing the consistency of reporting in the original interview and the reinterview of the number and/or quantity reported of various types of information. The table organizes the reporting into three general types: illness, utilization of health services, and miscellaneous items.

All three types of reporting followed similar patterns. First, the coefficients of agreement were almost uniformly high, providing very little opportunity for the interviewing procedure to show any effects. Second, for less educated respondents there was a slight tendency for the coefficients of agreement to be higher for interviews in which short questions and reinforcement were used than for other kinds of interviews. Thus, within this education group, the interviewing procedures which produced the most valid reporting of chronic conditions also showed a slight tendency to produce more reliable data when original interview and reinterview responses were compared. Among more educated respondents interviews in which long questions and no reinforcement were used did show a slight tendency to produce more reliable reporting for illness items. However, this pattern did not carry over to any great extent when health service reporting and miscellaneous item reporting were examined. For the more educated group of respondents the most reliable reporting of health service use and miscellaneous items was obtained, on the average, by the long question and reinforcement procedures.

To summarize, several different reliability rates were calculated for the reporting of health-related and nonhealth-related variables. The overall level of agreement between reporting in the original interview and the reinterview was high, leaving little variance upon which the experimental interviewing procedures could act. No large differences between procedures were found, although there was a very slight tendency for reliability to be highest among less educated respondents when short questions and reinforcement were used. While these data are far from definitive, it would appear that the relationship between reliability and validity of reporting health information in survey studies is fairly small and that persons or agencies who employ reliability of response as an index of the validity of collected data might benefit from further experimentation concerning this relationship.

Physician Summary Form reliability.—Respondent reports of chronic conditions were compared with data provided by physicians on the Physician Summary Form (PSF) in order to calculate the error and bias rates discussed previously.

Table 19. Product-moment correlation coefficients of agreement between quantities of various health items reported in original interviews and reinterviews, by reporting item, reinforcement, question length, and respondent education

Reporting item	Short questions without reinforcement		Short questions with reinforcement		Long questions without reinforcement		Long questions with reinforcement	
	Paired number	Correlation coefficient	Paired number	Correlation coefficient	Paired number	Correlation coefficient	Paired number	Correlation coefficient
	Less education							
Illness:								
Validated chronic conditions	21	.88	26	.96	23	.89	25	.92
All 19 chronic conditions	21	.89	26	.95	23	.90	25	.89
Total chronic and acute conditions	21	.85	26	.94	23	.87	25	.76
Symptoms	20	.85	24	.93	22	.78	22	.83
Health services:								
Physician visits, 10 weeks	18	.56	26	.88	21	.81	25	.28
Dentist visits, 12 months	21	.78	26	1.00	23	.90	25	1.00
Hospitalizations, 12 months	21	.93	26	.90	23	.99	25	.69
Miscellaneous:								
Health rating	21	.49	25	.89	23	.61	25	.75
Height	20	1.00	26	.98	22	.99	23	.98
Weight	20	1.00	26	.99	23	.99	24	1.00
Education	20	1.00	26	.96	23	.99	24	.98
Income	20	.96	23	.96	22	.95	24	.92
	More education							
Illness:								
Validated chronic conditions	29	.84	23	.72	28	.89	28	.93
All 19 chronic conditions	29	.88	23	.81	28	.89	28	.88
Total chronic and acute conditions	29	.82	23	.73	28	.81	28	.68
Symptoms	29	.76	20	.72	28	.88	25	.78
Health services:								
Physician visits, 10 weeks	25	.43	19	.91	26	.74	25	.74
Dentist visits, 12 months	29	.87	23	.94	28	.94	28	.96
Hospitalizations, 12 months	29	1.00	23	.69	28	.94	27	1.00
Miscellaneous:								
Health rating	29	.83	22	.86	28	.81	28	.77
Height	29	.98	23	.99	26	.92	27	.95
Weight	29	.99	23	.95	28	.94	28	.99
Education	29	.99	22	.96	28	.84	27	.99
Income	28	.86	22	.61	26	.87	26	.98

Also, as discussed previously, it was anticipated that the physician data could not adequately represent the entire "truth" about respondent chronic condition status as it existed at the time of the interview with the respondent. However, it was assumed that record "error" would be distributed randomly across interviewing procedures and thus would have no systematic effect on the comparisons between procedures. The effect of random error in the records, therefore, would be to attenuate the coefficients of respondent-physician agreement so that they would be lower than those expected if the physician records were truly valid.

An effort was made to obtain a rough estimate of the degrees of attenuation in the match rate coefficients due to record instability. Physicians filled out the chronic condition summary form when a patient was seen at a clinic. There were 88 persons for whom at least two physician records of chronic condition status were available. These records were made at two different times, but it is not known whether they were filled out by different physicians. Inconsistencies in the records may represent a refinement of diagnosis over time, a new appearance or a cure of a chronic condition, or a variance introduced in the interpretation of the medical records of the patient. When more than two PSF's were received for a person, comparisons were made between pairs of forms contiguous in time (namely, first with second, second with third, third with fourth, fourth with fifth). No more than five forms were received for any one patient. By adding these paired forms to the original 88 instances in which at least two forms were available, a total of 97 pairs of forms could be analyzed. The distributions of physician responses in terms of average numbers of conditions in each response category were as follows:

Average number of chronic conditions reported by response category				Missing data	Total chronic conditions
Form 1: Yes Form 2: Yes	No Yes or Yes No		No No		
1.21	.78		10.70	.31	13.00

The overall agreement rate between the two forms was $\frac{(1.21 + 10.70)}{(13.00 - .31)}$, or approximately 94 percent. The average consistency of checking "yes" for a particular condition was $\frac{1.21}{(1.21 + .78)}$, or about 61 percent.[h] The corresponding figure for checking "no" was 93 percent.

One cannot attach a great deal of meaning to these figures. The 88 persons or 97 duplicate cases were not altogether representative of the sample actually used. (The actual sample contained a higher proportion of persons with at least one condition checked "yes" on the PSF.) If the figures were representative and if respondents were completely truthful, then a lower limit which the mismatch coefficients could reach is expressed by 1.00 minus $\sqrt{\text{PSF reliability}}$. This figure would be .22 for "yes" answers and .04 for "no" answers. It is interesting to note that the underreporting scores achieved by less educated respondents interviewed with short questions and reinforcement approached this lower limit quite closely.

Number of Health Events Reported

In this section the effects of the experimental interviewing procedures and respondent education on amount of illness and use of health services reported are examined. In addition, the results of this study are compared with those of a previous study on reinforcement, and the possibility of using the number of chronic conditions reported as an index of accuracy of reporting is investigated.

Effects of interviewing procedures on the amount of health items reported.—In table 20 the average number of health items reported in original interviews and reinterviews is shown for each of the experimental groups. The F values greater than 2.00, indicating the statistical significance of the various trends, are also shown in this table. Generally speaking, for original in-

[h]The Pearson product-moment correlation coefficient for the total number of items checked "yes" on both forms is .77. Comparing the total number of items checked "yes" is not a sensitive measure to the existence of "compensating errors" on individual items. Therefore, the Pearson product-moment coefficient is not preferred as a measure of reliability here.

Table 20. Average number of health items reported in original interviews and reinterviews, by reinforcement, question length, and respondent education

Health item	Less education				More education			
	Short questions		Long questions		Short questions		Long questions	
	Without reinforcement	With reinforcement	Without reinforcement	With reinforcement	Without reinforcement	With reinforcement	Without reinforcement	With reinforcement
	Average number of items reported							
Original interview:								
Validated chronic conditions	2.40	2.80	2.66	2.69	2.08	1.59	2.08	1.98
Total chronic and acute conditions	5.02	6.02	5.57	5.56	4.34	3.97	4.58	4.39
Symptoms	7.00	7.54	7.55	7.92	6.45	5.90	6.50	5.33
Physician visits, 10 weeks	4.10	2.58	2.48	2.44	2.04	2.68	2.50	2.59
Dentist visits, 12 months	0.98	0.86	1.50	2.58	1.74	1.79	3.30	1.80
Hospitalizations, 12 months	0.24	0.26	0.16	0.19	0.09	0.08	0.18	0.24
Reinterview:								
Validated chronic conditions	2.95	3.00	2.64	2.81	2.22	1.55	2.19	2.04
Total chronic and acute conditions	5.90	5.96	5.68	5.24	4.56	3.60	4.27	4.33
Symptoms	8.16	7.22	7.52	7.39	6.89	4.06	5.77	4.95
Physician visits, 10 weeks	2.00	2.16	6.90	1.90	2.30	2.06	2.50	3.10
Dentist visits, 12 months	1.20	1.24	1.23	3.05	1.89	2.05	3.58	2.13
Hospitalizations, 12 months	0.36	0.32	0.23	0.00	0.07	0.00	0.19	0.35

Summary of F values greater than 2.00 from analyses of variance

Variable and health item	Source of variation		
	Reinforcement (R)	Question length (Q)	R x Q
Less education			
Reinterview:			
Hospitalizations, 12 months		3.48	
More education			
Original interview:			
Validated chronic conditions	2.24		
Symptoms	[b]3.96		
Dentist visits, 12 months		2.69	2.61
Hospitalizations, 12 months		3.17	
Reinterview:			
Validated chronic conditions	2.72		
Symptoms	[b]9.57		2.97
Dentist visits, 12 months		2.32	

[b] $p \leq .05$.

terviews the effects of interviewing procedures on numbers of items reported were unstable and did not show many consistent patterns. Further analysis, not shown here, has suggested that the effects of the interviewing procedures differ for some types of reporting depending upon whether the respondent was scheduled for a reinterview. As described earlier, the selection of respondents for reinterviews was not random. Original interviews with respondents scheduled for reinterviews were generally taken in the first few weeks of the 6-to-8-week interviewing period, and some clinics from which records were sampled were represented less frequently in the reinterview group. Rather than offer tentative conclusions about the effects of procedures on amount of health reporting here, it is suggested that further research is needed, employing designs which do not contain the confounding mentioned above. Fortunately, the procedural effects on the amount of reporting of the 13 validated chronic conditions were not mediated by the reinterview selection to any meaningful extent.

Comparison of reinforcement effects on amounts reported with previous research.—In a previous study by Marquis and Cannell[3] a reinforcement procedure of interviewing was found to produce about 25 percent more symptoms and chronic and acute conditions reported than were obtained without reinforcement. The previous research confounded reinforcement and question length because respondents receiving reinforcement were also asked questions that were longer (contained extra words) than were questions asked of nonreinforced respondents. In addition, the sample consisted of a cross section of adult white women living within the city limits of Detroit which differs somewhat from the sample of women in the present research. A reanalysis of the data from the previous research, controlling for respondent education (shown in the first section of table 21), shows that respondent education had little effect on amount of reporting when reinforcement was not used. When reinforcement was used, less educated respondents reported more health items, in general, than did more educated respondents.

Data from the present study are correspondingly rearranged and are shown by question length in the lower section of table 21. A similar education-reinforcement interaction can be seen when short questions were used, but only for symptom reporting when long questions were used. The lack of a reinforcement-education interaction for the long question procedure in the present study probably reflects differences between the two studies in the actual length of questions. In the present study, "long" questions were considerably longer and occurred more frequently than in the previous study.

By examining these two sets of data, it can also be seen that in the first study, the reinforcement effects on amount of reporting were generally confined to the less educated group. Moreover, reinforcement did not decrease amounts reported by the more educated respondents to any important extent, whereas in the present study, a definite reinforcement suppressing effect on illness reporting was often obtained for the more educated group.

Relationship between amount reported and accuracy of reporting.—The amount of health items reported is of interest because it has been used as an indicator of accuracy of reporting when validity data were not available (for example, in the study by Cannell, Fowler, and Marquis[2] and in the pilot studies for the present research).

The correspondence between the number of selected chronic conditions reported and the accuracy of reporting them is shown for the various experimental groups in table 22. When the experimental group was the unit of analysis, the rank order correlation between amount and accuracy of reporting was reasonably high (Spearman's rho = .64; $p \leq .05$). When the long question without reinforcement procedures for both education groups were removed, the rank order correlation was 1.00. This suggested that, for this kind of chronic condition reporting, differences among groups in amounts reported (especially those not asked long questions) tended to reflect, albeit imperfectly, group differences in validity. On the other hand, it would probably not be safe to conclude that the increased level of reporting produced by long questions reflected increases in accuracy of reporting. Why increases in number of conditions reported brought about by reinforcement reflected increases in accuracy, while increases in reporting produced by long questions did not, is a question which deserves further research.

Table 21. Average number of conditions and symptoms reported, with percent difference by education in the present study, by reinforcement, question length, and respondent education, compared with data by respondent status derived from an earlier study

Variable and health item	Without reinforcement			With reinforcement		
	Less education	More education	Percent difference[1]	Less education	More education	Percent difference[1]
	Average number of items reported					
Data from earlier study						
Chronic conditions:						
For self	1.46	1.09	34	1.59	1.44	10
By proxy	1.01	0.91	11	1.84	0.81	127
Total chronic and acute conditions:						
For self	2.14	2.19	-2	2.81	2.68	5
By proxy	1.46	1.39	5	2.41	1.36	76
Symptoms	5.23	5.07	3	7.42	5.51	35
Data from present study						
Short questions:						
Validated chronic conditions	2.40	2.08	15	2.80	1.59	76
Total chronic and acute conditions	5.02	4.34	16	6.02	3.97	52
Symptoms	7.00	6.45	9	7.54	5.90	28
Long questions:						
Validated chronic conditions	2.66	2.08	28	2.69	1.98	36
Total chronic and acute conditions	5.57	4.58	22	5.56	4.39	27
Symptoms	7.55	6.50	16	7.92	5.33	49

[1] Computed as: $\frac{\bar{X}_L - \bar{X}_M}{\bar{X}_M}$

Where $\bar{X}_L$ = mean number reported by less educated respondents

and $\bar{X}_M$ = mean number reported by more educated respondents.

Table 22. Average number of chronic conditions reported and average accuracy score, by reinforcement, question length, and respondent education

Interview procedure[1]	Average number of chronic conditions reported	Average accuracy score (1.00= total error)
Less education		
With reinforcement:		
Short questions-	2.80	.66
Long questions--	2.69	.53
Without reinforcement:		
Long questions--	2.66	.44
Short questions-	2.40	.47
More education		
Without reinforcement:		
Short questions-	2.08	.47
Long questions--	2.08	.61
With reinforcement:		
Long questions--	1.98	.46
Short questions-	1.59	.36

[1]Listed by rank order of column 1, average number of chronic conditions reported.

NOTE: Spearman's rho = .64; $p \leq .05$.

Finally, data in table 23 show the relationships of number of chronic conditions reported (out of 13 validated items) and the four indexes of reporting error and bias when the individual respondent is used as a unit of analysis. A marked tradeoff between overreporting and underreporting can be seen. Correlations of amount reported with underreporting were large and negative. Correlations of amount reported with overreporting were also large but in a positive direction. The net effect of this pattern on the total error score was quite small, indicating that decreases in underreporting errors were potentially canceled by increases in overreporting errors. It should be pointed out that the total error index gave a disproportionate weight to overreporting errors, so that for practical purposes the reduction in total error can be assumed to be somewhat greater than that indicated by the correlation coefficients.

For the entire sample, the correlation between the number of chronic conditions (of the 13 validated items) reported and total error is -.24 ($p < .01$). This indicates that the previously assumed positive relationship of quantity and accuracy of health reporting is confirmed but that only a small amount of the accuracy variance is explained by reporting quantity ($r^2 = .0576$). Thus, for future studies which are designed to explore for variables influencing accuracy of reporting (that is, studies that are not planned for comparisons among groups defined on an a priori basis) and which use quantity of reporting as an index of accuracy, large sample sizes are needed. The relative costs of increasing sample sizes and collecting actual validity data should be considered in the study plan.

Table 23. Product-moment correlation coefficients between number of chronic conditions reported and index scores of underreporting, overreporting, total error, and net bias, by reinforcement, question length, and respondent education

Respondent education and interview procedure	Number of chronic conditions reported	Correlation coefficient			
		Under-reporting	Over-reporting	Total error	Net bias
Less education					
Short questions:					
Without reinforcement	41	[a]-.48	[a].72	-.16	[a]-.69
With reinforcement	50	[a]-.46	[a].83	-.15	[a]-.68
Long questions:					
Without reinforcement	44	[a][b]-.40	[a][a].76	-.15	[a][a]-.63
With reinforcement	36	-.37	[a].77	-.15	[a]-.57
More education					
Short questions:					
Without reinforcement	53	[a][b]-.57	[a][a].82	[a]-.39	[a][a]-.70
With reinforcement	38	[b]-.39	[a].46	-.30	[a]-.50
Long questions:					
Without reinforcement	40	[b]-.34	[a][a].82	[a]-.13	[a][a]-.53
With reinforcement	46	[a]-.59	[a].63	[a]-.44	[a]-.72
Combined education groups					
Long and short questions:					
Without and with reinforcement	348	[a]-.47	[a].74	[a]-.24	[a]-.64

[a] $p \leq .01$.

[b] $p \leq .05$.

DISCUSSION

Education Differences in Ability and Bias

Earlier research in the NCHS-SRC series demonstrated that there is variation in the accuracy with which health information is reported by respondents. The initial studies suggested that some of the response error could be reduced by asking only for health information that was easy to recall accurately (for example, relatively recent and of high impact).

Since changing the nature of the events asked about is not always feasible, a search for other correlates of response error in health reporting was undertaken. This search indicated that the most immediate influences on response bias were to be found in the conduct of the interview itself rather than in the attitudes or demographic characteristics of respondents.

The results of the present research confirmed the earlier conclusion that response error can be affected greatly by even minor changes in the conduct of the interview. Making only two changes at the major points of leverage in the interview—the wording of the questions and the way in which the interviewer reacts to answers—had significant effects on information accuracy and response bias. A major surprise, however, was to find that these proximal influences were mediated by respondent education. These highly reliable effects suggest again that efforts to improve the data collected by personal interviews should take into account the fact that the interview is an instance of the two-person social interaction, potentially governed by the variables important in such interactions. Any changes in the behavior of one participant, the interviewer or the respondent, will have potentially complex effects on the behavior of the other. As the following discussion infers, the thinking in regard to variables affecting response accuracy in health interviews has come almost full circle.

Earlier research failed to confirm the hypothesized importance of the main effects of respondent cognitive and demographic (for example, education) variables on reporting accuracy. Current results and interpretations indicate that these variables are indeed important, but only as mediators of the effects of different interviewing practices, rather than as prime causal variables in the "main effect" sense.

An attempt is made in the material presented below to construct a social-psychological frame of reference which will provide a setting for the current results and which may be used as a basis for future research. In the discussion certain assumptions derived from signal detection theory are made—changes in total error scores reflect changes in ability to recall and changes in net bias scores reflect changes in reporting bias tendencies.

Ability to recall.—It is generally assumed that a significant correlation exists between education level and general ability. It is assumed, too, that more able people tend also to be more educated. A corresponding hypothesis is that less educated respondents are less able to recall their chronic conditions than are more educated persons.

The data from the current study, however, do not support this hypothesis. When conventional interviewing procedures (short questions without reinforcement) were used, the total error scores of the two education groups were approximately the same. In addition, the minimum total error index scores of each group were similar, although achieved with different interviewing conditions.

Interview procedure	Total error index score	
	Less education	More education
Conventional interviewing procedure (short questions without reinforcement):		
Original interview----------	.53	.53
Reinterview----------	.57	.62
"Best" procedure (minimum total error index score):		
Original interview, short questions with reinforcement----------	.36	
Reinterview, long questions without reinforcement------		.37

If reinforcement improved the ability of less educated respondents to recall, reinterview total error index scores would be lower than those obtained in an initial interview. The data show opposite trends for the less educated group. Thus, assuming total error scores to be an index of ability to recall, the different experimental procedure effects on reporting could not be attributed to different ability levels of the education groups. Furthermore, the differential reinforcement and question length effects could not be attributed to their effect on different ability levels.

Reporting bias.—The current data do suggest that the two education groups differed somewhat in their tendency to deny to the interviewer the presence of existing health conditions—the "no"-saying response bias. Under conventional procedures (short questions without reinforcement) the net bias scores were lower for the less educated group than for the more educated group. Moreover, ignoring the statistical correction analysis (attempts to minimize over- or underreporting statistically), the lowest average net bias index score obtained by the less educated group was about half of the lowest average net bias index score achieved by the high school graduate group:

Interview procedure	Net bias index score	
	Less education	More education
Conventional interviewing procedure (short questions without reinforcement):		
Original interview	.24	.35
Reinterview	.23	.38
"Best" procedure (minimum net bias index score):		
Original interview, short questions with reinforcement	.10	
Reinterview, long questions without reinforcement		.20

One hypothesis relevant to understanding the bias problem is that the education groups differ in the extent to which confidence or certainty is required before sickness or health service use is reported. In other words, the threshold of certainty for responding "yes" may differ between groups. Thus, when a more educated respondent is unsure of an answer, he may be inclined to search his memory for further confirmation of his tentatively recalled facts and respond "yes" only when he is fairly sure this is the correct answer. Procedures, therefore, which stimulate and allow extra time for this confirmation memory search (such as long questions and reinterviews, both of which present the question or search stimulus a second time) may increase the accuracy of reporting of persons who are "conservative" in the sense of not admitting sickness, or other facts, when there is some uncertainty about the answer.

Reinforcement, coming as it did immediately after an answer, may "cut off" additional memory search or answer-confirmation activity and thereby maintain the existing conservatism bias of higher educated respondents. Future research might be directed toward testing the hypothesis that the education groups differ in the degree of certainty required before reporting health information and toward exploring the idea that increased memory-search or answer-confirmation time will aid accuracy in reporting of more educated respondents. If the conservatism bias hypothesis is confirmed, other interviewing procedures might be developed to reduce it directly.

There is some independent evidence that higher educated respondents might be more reluctant to report sickness to an interviewer. Phillips and Clancy[10] found that higher status respondents (status defined in terms of education and income) rated more items on a 22-item mental illness symptoms list as "less desirable to have" than did lower status respondents. If the social undesirability of reporting symptoms of physical illness showed the same interaction effect with education as did the reporting of mental illness symptoms, then the conservatism or high level of "no"-saying bias observed for the higher education group in this research may be a product of a special sensitivity to the undesirability of mentioning illnesses in public.

However, if the observed bias of the higher educated group is a social desirability bias, it is not at all clear how it might be reduced. Clinical interviewers, such as psycholgists, might claim that this type of bias is reduced by creating more rapport: a warm, supportive, friendly atmosphere of communication. As will be discussed below, however, one hypothesis suggested by the data from this study is that there is too much rapport in interviews with higher educated respondents and that it is this rapport which leads to increases in the "no"-saying bias tendencies.

Social Status, Rapport, and Task Orientation

Hyman and associates[11] have pointed to the distinction between task involvement and total involvement in characterizing the relationship between the interviewer and respondent during an interview. They suggested that valid data may be more a function of task involvement. Too much rapport or involvement in the nontask aspects of the relationship may detract from accuracy of reporting when a "hen party" atmosphere is maintained.

Back and Gergen[12] made a somewhat similar distinction between the "information game" and the "ingratiation game," suggesting that the interviewer's problem is to maximize information giving and minimize ingratiation. These writers theorized that as social distance between the interviewer and respondent decreases (that is, they are more alike in demographic and other status characteristics), the tendency increases for the respondent to give ingratiating rather than accurate, task-oriented answers.

One series of studies (Dohrenwend, Williams, and Weiss[13]) focusing on the possible interaction effects of rapport and social distance on accuracy and bias in the personal interview differed on a number of dimensions and often produced different conclusions. These studies did seem to suggest, however, that response bias is a complex function of the difference (or similarity) of status between the interviewer and respondent and the amount of rapport created in the interview. For example, Weiss' data[14] indicated that response inaccuracy is inversely related to social status similarity, and that when status differences are small, rapport decreases accuracy of reporting.

To clarify the relevance of the above discussion, several things about the variables in the current study should be mentioned. The interviewers were all "high status" in terms of education. All had completed high school and many had some college training. Hence, when we talk about variation in respondent education we are also talking about variation in the similarity of interviewer-respondent educational status: social distance was relatively large when the respondent had not graduated from high school and relatively small when the respondent had graduated. Second, the reinforcement procedure might be regarded as one which was somewhat high in rapport, since it was generally friendly and accepting of reports of illness.

If the above reinterpretations of education and reinforcement are made, the results of this study tend to agree with those of the other researchers in that, when social differences are small (high educated respondent and high educated interviewer here), an increase in rapport (here, using reinforcement) also increases bias.

Fowler[15] made an extensive reanalysis of data from a study of health reporting by Cannell, Fowler, and Marquis[2] and has presented perhaps the most elaborated theory to date concerning the relationship between rapport and respondent education, and between cognitive approaches and respondent education.

Fowler theorized that because highly educated and less educated respondents differ on a number of dimensions, the personal interview should be conducted to emphasize different things for the two education groups. Less educated respondents are seen as having less skill or ability to report accurately, although Fowler's analysis suggested that interviewers can and do compensate for skill differences in various ways, such as probing or clarifying inadequate answers. The reporting performance of more highly educated respondents is largely under their own cognitive control. Reporting accuracy, therefore, is mainly a function of how well the highly ed-

ucated respondent understands what he is supposed to do. The reporting performance of less educated respondents is not internally controlled but much more dependent upon guidance and help from the interviewer. Explaining to the less educated respondent about the reporting requirements of the research (for example, complete accuracy and coverage) is hypothesized to have no effect on reporting accuracy since this performance is not strongly governed by the degree of "understanding" achieved. Finally, the nature of the social interaction is hypothesized to differ according to the level of respondent education. Because survey research interviewers tend to be high school graduates or even college graduates, a respondent who has not completed high school is at a clear disadvantage vis-à-vis the interviewer in the interview social interaction. Fowler hypothesized that this status discrepancy presents a problem to the lower status respondent and that his reporting performance is a function of how this problem is resolved.

In some unpublished pilot studies carried out at the Survey Research Center, Wood[16] was able to experimentally control three independent variables (ability to perform a recall task, reinforcement, and cognitive control over performance) and to measure the effects of combinations of these variables on recall accuracy. Her results (which were not always statistically significant due to the small sample sizes used in the pilot research) were somewhat parallel to those obtained in the present study. For the low recall ability group, both reinforcement and understanding of the task requirements had a beneficial effect on recall accuracy. But for those with high recall ability, reinforcement and task understanding worked against each other. Respondents who understood the recall task and had the ability to perform it reasonably well actually did worse when the interviewer reinforced correct answers than when the interviewer did not use reinforcement.

Most of the studies cited above produced the common finding that the status relationship between the interviewer and respondent has an effect on reporting accuracy and bias. Some writers go further to state that social status interacts with the performance of the interviewer to influence the data.

The implication made by most of these writers is that the natural relationship between a lower status respondent and a higher status interviewer is not a personal one. Back and Gergen[12] infer that this is good because the information gain can be maximized. Fowler,[15] Weiss,[14] and others also see a potential benefit in a large social distance because, under these conditions, if the interviewer is of higher status than the respondent, her behavior has a great influence on the respondent's performance. Fowler[15] and deKoning[17] concluded that probing and question clarification are necessary and beneficial. Fowler went even further to suggest that abstract explanations will have no effect. The present study showed that the redundant long questions and reinterviews (all task-oriented experimental phenomena) had negative effects on the less educated respondents. All of these findings suggest to the present writers that interview situations where there is a wide social discrepancy call for task-oriented interview behavior which (1) is contingent on respondent behavior and (2) has a friendly, supportive component.

The reinforcement procedure seems to meet both criteria. It should be noted, however, that reinforcement is not the only way to achieve these two goals. As Fowler theorized, more traditional interviewing styles employing effective probing strategies and occasional, noncontingent interpersonal behavior (for example, a pleasant personal comment, a joke, and some laughter) should achieve the same effects. The question now becomes whether effective probing (essentially negative reinforcement because probes are used when the respondent fails to perform his role properly) plus some positive but irrelevant comments can serve the same function as a schedule of positive reinforcement for the less educated respondent group.

Caution is advised against accepting the idea that noncontingent supportiveness really

has any effect on accuracy of reporting. In a study by Marquis *et al.*[18] which was designed specifically to test the effects of a supportive versus challenging atmosphere on reporting accuracy the data clearly indicated that respondents enjoyed the supportive atmosphere more and thought they did a better job of reporting. However, accuracy of recall was not even slightly influenced by the contrast in interrogation atmospheres.

The researchers and theorists referred to above point out that a different problem exists when the status of the interviewer and respondent are similar. The natural relationship in this case is not a task-oriented one, but an interpersonal, friendship one. If the interviewer does anything to support this perception, the resulting atmosphere will encourage minimum attention to the reporting task as well as answers appropriate to a hen party (socially acceptable, ingratiating answers). The problem with the equal status relationship appears to be one of too much rapport (total involvement). Apparently, in this situation, an interviewer must be attentive to establishing a task-oriented attitude for the respondent. Fowler suggests that this can be done cognitively by the interviewer presenting a clear explanation of the repondent role. The current results show that long, redundant questions and reinterviews also seem to achieve the desired effects. Neither of these procedures is contingent on the behavior of the respondent.

The above findings imply that in situations where the interviewer and respondent are of the same status (in this case high education women), a reinforcement procedure will accentuate a natural tendency to establish an interpersonal rather than task-oriented relationship. The more friendly the interview situation becomes, the worse the data become.

A final and alternative hypothesis (based on the work of Wood mentioned above) concerns the negative effects of reinforcement in the high status relationship. The problem becomes one of not creating too much rapport with reinforcement but requiring the interviewer to engage in behavior which is inappropriate or unwanted in an equal status situation. Reinforcement may be perceived as unnecessary and awkward after every report of illness by the highly educated respondent. Such a respondent undoubtedly feels she understands what the interviewer wants and that she is perfectly capable of performing the apparently easy reporting task without the interviewer's constant reminders that performance is adequate.[i] Possibly it is the confident, highly educated respondent who might benefit from negative reinforcement for inadequate performance. The major benefits for this respondent would seem to be derived from changing her perception that minimal effort is needed to report accurately and that the reporting task is only a minor part of her interactive relationship with the interviewer.

Summary of Discussion

It is hypothesized that the effects of reinforcement, question length, and reinterviews differ among education groups for the following reasons:

Less educated respondents rely on interviewer cues to direct their reporting performance. Thus, appropriate use of reinforcement, probes, and other feedback by the interviewer can aid recall and reporting accuracy a great deal. Because of the social status discrepancy, feedback is perceived as appropriate and actually welcomed. Abstract explanations are probably irrelevant to performance for this group, and long questions appear only to serve as a source of confusion. Performance is about as good as it can be in an appropriately conducted initial interview; there appears to be no additional benefit from a reinterview.

More educated respondents carry out the reporting task largely according to their own understanding of it; they do not rely on cues from the interviewer. Reinforcement under these circumstances may encourage an interpersonal (versus task) orientation; may be perceived as

[i] It should be noted that respondents' perceptions of reporting requirements can be wrong. In one study[2] 45 percent of the respondents had incorrect ideas about accuracy of reporting requirements; in another study[3] 41 to 69 percent of respondents misunderstood the requirements to varying extents.

inappropriate, unnecessary, or even condescending; and may, at least for open-ended questions, "cut off" the additional search and confirmation time needed to avoid underreporting. The more educated respondent appears to have a stronger tendency to underreport chronic conditions, possibly due to a stronger conservatism or social desirability bias. This type of respondent may feel the need to rethink answers and be very sure of their accuracy before admitting the existence of a chronic condition. Long questions and reinterviews apparently provide the additional cues necessary to search memory to confirm ambiguous answers and, thereby, provide the confidence level required to report a chronic condition.

These highly speculative hypotheses are offered as a guide to further research on survey interview reporting accuracy. Hopefully, they highlight some of the important areas to consider when survey data validity is of concern. These "human" characteristics involving memory, recall, cognition, social status, and interpersonal interaction are often overlooked in methodological studies, yet appear to be very crucial in any meaningful understanding of the personal interview process.

REFERENCES

[1]National Center for Health Statistics: Interview data on chronic conditions compared with information derived from medical records. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 23. Public Health Service. Washington. U S. Government Printing Office, May 1967.

[2]National Center for Health Statistics: The influence of interviewer and respondent psychological and behavioral variables on the reporting in household interviews. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 26. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1968.

[3]National Center for Health Statistics: Effect of some experimental interviewing techniques on reporting in the health interview survey. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 41. Public Health Service. Washington. U.S. Government Printing Office, May 1971.

[4]Matarazzo, J. D., Saslow, G., and Wiens, A. N.: Studies in interview speech behavior, in L. Krasner and L. P. Ullman, eds., *Research in Behavior Modification.* New York. Holt, Rinehart and Winston, 1965.

[5]Ferber, R.: *Collecting Financial Data by Consumer Panel Techniques.* Urbana. Bureau of Economic and Business Research, University of Illinois, 1959.

[6]Neter, J., and Waksberg, J.: *Response Errors in Collection of Expenditures Done by Household Interviews: An Experimental Study.* Washington. U.S. Department of Commerce, Bureau of Census Technical Paper No. 11, 1965.

[7]National Center for Health Statistics: Chronic conditions and activity limitation, United States, July 1961-June 1963. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 10-No. 17. Public Health Service. Washington. U.S. Government Printing Office, May 1965.

[8]U.S. National Health Survey: Health interview responses compared with medical records. *Health Statistics.* PHS Pub. No. 584-D5. Public Health Service. Washington. U.S. Government Printing Office, June 1961.

[9]Green, D. M., and Swets, J. A.: *Signal Detection Theory and Psychophysics.* New York. John Wiley and Sons, 1966.

[10]Phillips, D. L., and Clancy, K. J.: Response biases in field studies of mental illness. *Am. Sociol. Rev.* 35 (3): 503-515, June 1970.

[11]Hyman, H. H., Cobb, W. J., Feldman, J. J., Hart, C. W., and Stember, C. H.: *Interviewing in Social Research.* Chicago. University of Chicago Press, 1954.

[12]Back, K., and Gergen, K.: Idea orientation and ingratiation in the interview: a dynamic model of response bias. *Proceedings of the Social Statistics Section of the American Statistical Association.* Technical Report No. 13, 1963, pp. 284-288.

[13]Dohrenwend, B. S., Williams, J. A., and Weiss, C. H.: Interviewer biasing effects: toward a reconciliation of findings. *Pub. Op. Quart.* XXXIII (1): 121-129, 1969.

[14]Weiss, C. H.: Validity of welfare mothers' interview responses. *Pub. Op. Quart.* XXXII (4): 622-633, 1968.

[15]Fowler, F. J., Jr.: *Education, Interaction, and Interview Performance.* Doctoral dissertation, The University of Michigan, 1965.

[16]Wood, L.: The role of task difficulty in operant verbal conditioning. Personal communication.

[17]deKoning, T. L.: *Interviewer and Respondent Interaction in the Household Interview.* Doctoral dissertation, The University of Michigan, 1969.

[18]Marquis, K. H., Marshall, J., and Oskamp, S.: Accuracy and completeness of testimony as a function of kind of question, interrogation atmosphere, and item content. *J. Appl. Soc. Psychol.* 1972, in press.

———O O O———

APPENDIX I

SAMPLE PAGES FROM QUESTIONNAIRES

INTERVIEW PROCEDURE A—Short Questions Without Reinforcement

Symptoms

P45970B
1-A

1. Now I'm going to ask you some questions about your health. Have you EVER had-------?

(DO NOT ENTER THIS INFORMATION ON TABLES)

	Yes	No
a. Bad headaches?		
b. Coughed up blood?		
c. Had fainting or blackout spells?		
d. Bad sore throats?		
e. Shortness of breath?		
f. Serious backaches?		
g. Ever felt your heart beating hard or acting funny?		
h. Pain in or around your heart or chest?		
i. Gas in your stomach?		
j. Bad stomach cramps?		
k. Loose bowels?		
l. Pain or soreness in the female organs?		
m. Pain or burning when you go to the bathroom?		
n. Ever had painful or swollen joints?		
o. Any broken bones?		
p. Itching skin?		
q. Mental illness?		
r. Ever had trouble sleeping?		
s. Ever had any venereal disease?		

45970B
6-A

6. Please tell me if you have had any of these conditions during the past 12 months?:

	Yes	No
a. Asthma?		
(IF YES) What kind of asthma is it? _____		
b. Hay fever?		
c. Thyroid trouble or goiter during the past 12 months?		
(IF YES) What kind of thyroid trouble is it? _____		
d. Repeated bronchitis?		
e. Repeated skin trouble?		
(IF YES) What kind of skin trouble is it? _____		
f. Paralysis of any kind?		
g. Hemorrhoids or piles during the past 12 months?		
h. Hernia or rupture?		
i. Repeated gall bladder or liver trouble?		
(IF YES) What kind of trouble is it? _____		
j. Peptic or stomach ulcer?		
k. Varicose veins?		
l. Have you had repeated trouble with your back or spine?		
(IF YES) What kind of back trouble is it? _____		

7. Please tell me if you have ever had any of these conditions:

	Yes	No
m. Arthritis or rheumatism?		
n. Rheumatic fever?		
o. Heart disease or any heart trouble?		
(IF YES) What kind was it? _____		
p. Ever had a stroke?		
(IF YES) How does it affect you now? _____ _____		
q. Hypertension or high blood pressure?		
r. Hardening of the arteries?		
s. Ever had diabetes?		

9 During the past 14 days, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q10)

9b. Please take this pencil and circle the date of each visit on the calendar.

9c. According to what you have circled, you saw a doctor _____ time(s) during the last 14 days. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OF VISIT IF EXACT DAY UNKNOWN)

10. During the three months outlined in blue on the calendar, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q11)

10b. Please circle the date of each visit on the calendar.

10c. According to what you have circled, you saw a doctor _____ time(s) during the three months outlined in blue. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OR MONTH OF VISIT IF EXACT DAY UNKNOWN)

11. (Not counting the visits you have already mentioned) During the times outlined in red and blue, did you:

	Yes	No	When was this? *
See a doctor in an emergency room?			
At your home?			
Talk to a doctor over the phone? (DO NOT COUNT CALLS ONLY TO MAKE APPOINTMENTS)			

*INTERVIEWER: CIRCLE ANY ADDITIONAL DATES OF CONTACTS MENTIONED

Symptoms

P45970
2-B

R STATEMENTS - USE SHORT AND LONG	
Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

1. Now I'm going to ask you some questions about your health. Have you EVER had-------?

(DO NOT ENTER THIS INFORMATION ON TABLES)

	Yes	No
a. Bad headaches?	R	
b. Coughed up blood?	R	
c. Had fainting or blackout spells?	R	
d. Bad sore throats?	R	
e. Shortness of breath?	R	
f. Serious backaches?	R	
g. Ever felt your heart beating hard or acting funny?	R	
h. Pain in or around your heart or chest?	R	
i. Gas in your stomach?	R	
j. Bad stomach cramps?	R	
k. Loose bowels?	R	
l. Pain or soreness in the female organs?	R	
m. Pain or burning when you go to the bathroom?	R	
n. Ever had painful or swollen joints?	R	
o. Any broken bones?	R	
p. Itching skin?	R	
q. Mental illness?	R	
r. Ever had any trouble sleeping?	R	
s. Ever had any venereal disease?	R	

R STATEMENTS - USE SHORT AND LONG	
Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

45970B
6-B

6. Please tell me if you have had any of these conditions during the past 12 months?:

	Yes	No
a. Asthma?	R	
(IF YES) What kind of asthma is it? _____	R	
b. Hay fever?	R	
c. Thyroid trouble or goiter during the past 12 months?	R	
(IF YES) What kind of thyroid trouble is it? _____	R	
d. Repeated bronchitis?	R	
e. Repeated skin trouble?	R	
(IF YES) What kind of skin trouble is it? _____	R	
f. Paralysis of any kind?	R	
g. Hemorrhoids or piles during the past 12 months?	R	
h. Hernia or rupture?	R	
i. Repeated gall bladder or liver trouble?	R	
(IF YES) What kind of bladder trouble is it? _____	R	
j. Peptic or stomach ulcer?	R	
k. Varicose veins?	R	
l. Have you had repeated trouble with your back or spine?	R	
(IF YES) What kind of back trouble is it? _____	R	

R STATEMENTS - USE SHORT AND LONG	
Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

45970B
7-B

7. Please tell me if you have ever had any of these conditions:

	Yes	No
m. Arthritis or rheumatism?	R	
n. Rheumatic fever?	R	
o. Heart disease or any heart trouble?	R	
(IF YES) What kind was it? _____	R	
p. Ever had a stroke?	R	
(IF YES) How does it affect you now? _____ _____	R	
q. Hypertension or high blood pressure?	R	
r. Hardening of the arteries?	R	
s. Ever had diabetes?	R	

45970B
9-B

R STATEMENTS - USE LONG ONLY

Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

9. During the past 14 days, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q10)

9b. Please take this pencil and circle the date of each visit on the calendar. [R]

9c. According to what you have circled, you saw a doctor _____time(s) during the last 14 days. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OF VISIT IF EXACT DAY UNKNOWN)

10. During the three months outlined in blue on the calendar, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q11)

10b. Please circle the date of each visit on the calendar. [R]

10c. According to what you have circled, you saw a doctor _____ time(s) during the three months outlined in blue. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OR MONTH OF VISIT IF EXACT DAY UNKNOWN)

11. (Not counting the visits you have already mentioned) During the times outlined in red and blue, did you:

	Yes	No	When was this? *
See a doctor in an emergency room?	[R]		[R]
At your home?	[R]		[R]
Talk to a doctor over the phone? (DO NOT COUNT CALLS ONLY TO MAKE APPOINTMENTS)	[R]		[R]

*INTERVIEWER: CIRCLE ANY ADDITIONAL DATES OF CONTACTS MENTIONED.

Symptoms

P45970
2-C

1. Now I'm going to ask you some questions about your health. By asking these questions, The Public Health Service can get a good picture of the nation's health. Have you EVER had-----?

(DO NOT ENTER THIS INFORMATION ON TABLES)

	Yes	No
a. Bad headaches?		
b. Coughing up blood is the next one. Have you coughed up blood?		
c. How about fainting spells or blackout spells. Have you had these?		
d. Now a question about bad sore throats. We're looking for information about these. Have you had bad sore throats?		
e. What about shortness of breath?		
f. The next item is serious backaches? Have you had serious backaches?		
g. Ever felt your heart beating hard or acting funny?		
h. What about pain in or around your heart or chest? Have you had that kind of pain?		
i. Gas in the stomach is the next item. This is another health problem we're interested in. Have you had gas in your stomach?		
j. We'd like to know about bad stomach cramps. Have you had them?		
k. The next question is about loose bowels? Have you had loose bowels?		
l. Have you had pain or soreness in the female organs?		
m. What about pain or burning when you go to the bathroom?		
n. How about painful or swollen joints. These are other items we would like to know about. Have you ever had painful or swollen joints?		
o. Broken bones is the next item. Have you had any broken bones?		
p. Now a question about itching skin. Have you had that?		
q. What about mental illness. This is another kind of condition we need information about. Have you had any mental illness?		
r. Have you ever had trouble sleeping?		
s. Venereal disease is the last item of this list. We'd like to get an estimate of this condition in the population. Have you ever had any venereal disease?		

6. Next, I'm going to read a list of some health conditions of special interest to the Health Service. Please tell me if you have had any of these conditions during the past twelve months.

	Yes	No
a. Asthma is the first one we need information about. Have you had asthma?		
(IF YES) What kind of asthma is it? _____		
b. What about hay fever? Have you had that?		
c. Another area of interest to the Health Service is thyroid and goiter trouble. Have you had any thyroid trouble or goiter during the past 12 months?		
(IF YES) What kind of thyroid trouble is it? _____		
d. The next item is repeated bronchitis. Have you had that?		
e. We are also interested in trouble with your skin. Have you had repeated skin trouble?		
(IF YES) What kind of skin trouble is it? _____		
f. Have you had paralysis of any kind?		
g. Hemorrhoids and piles is the next item we need information about. Have you had hemorrhoids or piles during the past 12 months?		
h. We are looking for information about hernias and ruptures. Have you had a hernia or rupture?		
i. Have you had any repeated gall bladder or liver trouble?		
(IF YES) What kind of trouble is it? _____		
j. We are also interested in how many people had peptic or stomach ulcers. Have you had a peptic or stomach ulcer?		
k. What about varicose veins? Have you had varicose veins?		
l. The last item on this list asks about repeated back trouble. Have you had repeated trouble with your back or spine?		
(IF YES) What kind of back trouble is it? _____		

7. Here are some other conditions we want to ask you about. This time we want to know if you have ever had any of them. Please tell me if you have ever had any of these conditions.

	Yes	No
m. The first item is arthritis or rheumatism. We're interested in how many people have this condition. Have you had arthritis or rheumatism?		
n. The next one is rheumatic fever. Have you had that?		
o. Heart disease or heart trouble is the next item. This is another area of special interest to the Health Service. Have you had heart disease or any heart trouble?		
(IF YES) What kind was it? _____		
p. The next item is stroke. Have you ever, at any time, had a stroke?		
(IF YES) How does it affect you now? _____ _____		
q. We also need information on hypertension or high blood pressure. Have you had hypertension or high blood pressure?		
r. What about hardening of the arteries. Have you had that?		
s. The last item is diabetes. Have you ever had diabetes?		

9. The next questions are about doctors, hospitals, and other places people use in connection with their health. The health service is very interested in how people use the services of doctors and clinics. During the past 14 days, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q10)

9b. Please take this pencil and circle the date of each visit on the calendar.

9c. According to what you have circled, you saw a doctor _____ time(s) during the last 14 days. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OF VISIT IF EXACT DAY UNKNOWN)

10. Now we are interested in the last three months. We ask people about the times they have talked to a doctor during this period. During the three months outlined in blue on the calendar, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q11)

10b. Please circle the date of each visit on the calendar.

10c. According to what you have circled, you saw a doctor _____ time(s) during the three months outlined in blue. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OR MONTH OF VISIT IF EXACT DAY UNKNOWN)

11. In addition to seeing a doctor in his office, there are other places where people see doctors about their health. During the times outlined in red and blue (but not counting what you've already mentioned) did you:

	Yes	No	When was this?*
See a doctor in an emergency room?			
How about a telephone call. Did you talk with a doctor about your health over the telephone? (DO NOT COUNT CALLS ONLY TO MAKE APPOINTMENTS)			
We are also interested in the times the doctor came to your home in connection with your health. Did you talk to a doctor at your home during the times outlined on the calendar?			

*INTERVIEWER: CIRCLE ANY ADDITIONAL DATES OF CONTACTS MENTIONED

Symptoms

P45970
2-D

R STATEMENTS - USE SHORT AND LONG

Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

1. Now I'm going to ask you some questions about your health. By asking these questions, The Public Health Service can get a good picture of the nation's health. Have you EVER had-----?

(DO NOT ENTER THIS INFORMATION ON TABLES)

	Yes	No
a. Bad headaches?	R	
b. Coughing up blood is the next one. Have you coughed up blood?	R	
c. How about fainting spells or blackout spells. Have you had these?	R	
d. Now a question about bad sore throats. We're looking for information about these. Have you had bad sore throats?	R	
e. What about shortness of breath?	R	
f. The next item is serious backaches? Have you had serious backaches?	R	
g. Ever felt your heart beating hard or acting funny?	R	
h. What about pain in or around your heart or chest? Have you had that kind of pain?	R	
i. Gas in the stomach is the next item. This is another health problem we're interested in. Have you had gas in your stomach?	R	
j. We'd like to know about bad stomach cramps. Have you had them?	R	
k. The next question is about loose bowels? Have you had loose bowels?	R	
l. Have you had pain or soreness in the female organs?	R	
m. What about pain or burning when you go to the bathroom?	R	
n. How about painful or swollen joints. These are other items we would like to know about. Have you ever had painful or swollen joints?	R	
o. Broken bones is the next item. Have you had any broken bones?	R	
p. Now a question about itching skin. Have you had that?	R	
q. What about mental illness. This is another kind of condition we need information about. Have you had any mental illness?	R	
r. Have you ever had trouble sleeping?	R	
s. Venereal disease is the last item of this list. We'd like to get an estimate of this condition in the population. Have you ever had any venereal disease?	R	

R STATEMENTS - USE SHORT AND LONG	
Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

P45970B
7-D

6. Next, I'm going to read a list of some health conditions of special interest to the Health Service. Please tell me if you have had any of these conditions during the past twelve months.

	Yes	No
a. Asthma is the first one we need information about. Have you had asthma?	R	
(IF YES) What kind of asthma is it? _____	R	
b. What about hay fever? Have you had that?	R	
c. Another area of interest to the Health Service is thyroid and goiter trouble. Have you had any thyroid trouble or goiter during the past 12 months?	R	
(IF YES) What kind of thyroid trouble is it? _____	R	
d. The next item is repeated bronchitis. Have you had that?	R	
e. We are also interested in trouble with your skin. Have you had repeated skin trouble?	R	
(IF YES) What kind of skin trouble is it? _____	R	
f. Have you had paralysis of any kind?	R	
g. Hemorrhoids and piles is the next item we need information about. Have you had hemorrhoids or piles during the past 12 months?	R	
h. We are looking for information about hernias and ruptures. Have you had a hernia or rupture?	R	
i. Have you had any repeated gall bladder or liver trouble?	R	
(IF YES) What kind of trouble is it? _____	R	
j. We are also interested in how many people had peptic or stomach ulcers. Have you had a peptic or stomach ulcer?	R	
k. What about varicose veins? Have you had varicose veins?	R	
l. The last item on this list asks about repeated back trouble. Have you had repeated trouble with your back or spine?	R	
(IF YES) What kind of back trouble is it? _____	R	

P45970 B
8-D

R STATEMENTS - USE SHORT AND LONG

Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

7. Here are some other conditions we want to ask you about. This time we want to know if you have ever had any of them. Please tell me if you have ever had any of these conditions.

	Yes	No
m. The first item is arthritis or rheumatism. We're interested in how many people have this condition. Have you had arthritis or rheumatism?	R	
n. The next one is rheumatic fever. Have you had that?	R	
o. Heart disease or heart trouble is the next item. This is another area of special interest to the Health Service. Have you had heart disease or any heart trouble?	R	
(IF YES) What kind was it? _____	R	
p. The next item is stroke. Have you ever, at any time, had a stroke?	R	
(IF YES) How does it affect you now? _____ _____	R	
q. We also need information on hypertension or high blood pressure. Have you had hypertension or high blood pressure?	R	
r. What about hardening of the arteries. Have you had that?	R	
s. The last item is diabetes. Have you ever had diabetes?	R	

P45970 B
10-D

R STATEMENTS - USE LONG ONLY

Thank you	We're interested in that.
Mm-hmm	That's the kind of information we need.
I see	(REPEAT ANSWER JUST GIVEN)
Yes	That's important.
O.K.	We need to know that.

9. The next questions are about doctors, hospitals, and other places people use in connection with their health. The health service is very interested in how people use the services of doctors and clinics. During the past 14 days, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q10)

9b. Please take this pencil and circle the date of each visit on the calendar. [R]

9c. According to what you have circled, you saw a doctor _____ time(s) during the last 14 days. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OF VISIT IF EXACT DAY UNKNOWN)

10. Now we are interested in the last three months. We ask people about the times they have talked to a doctor during this period. During the three months outlined in blue on the calendar, did you talk to a doctor or go to a doctor's office or clinic for yourself?

☐ Yes ☐ No (SKIP TO Q11)

10b. Please circle the date of each visit on the calendar. [R]

10c. According to what you have circled, you saw a doctor _____ time(s) during the three months outlined in blue. Is that correct?

(NOTE IN CALENDAR MARGIN: MORE THAN ONE VISIT ON A SINGLE DAY)
WEEK OR MONTH OF VISIT IF EXACT DAY UNKNOWN)

11. In addition to seeing a doctor in his office, there are other places where people see doctors about their health. During the times outlined in red and blue (but not counting what you've already mentioned) did you:

	Yes	No	When was this?*
See a doctor in an emergency room?	[R]		[R]
How about a telephone call. Did you talk with a doctor about your health over the telephone? (DO NOT COUNT CALLS ONLY TO MAKE APPOINTMENTS)	[R]		[R]
We are also interested in the times the doctor came to your home in connection with your health. Did you talk to a doctor at your home during the times outlined on the calendar?	[R]		[R]

*INTERVIEWER: CIRCLE ANY ADDITIONAL DATES OF CONTACTS MENTIONED

APPENDIX II

FORMS USED IN STUDY

INTERVIEWER'S FORM

P45970B

RESPONDENT'S NAME: _____

ADDRESS: _____

PHONE NO: _____

SERIAL NUMBER
OF INITIAL INTERVIEW: _____

FORM LETTER OF
INITIAL INTERVIEW: _____
(A,B,C, or D)

DATE OF INITIAL INTERVIEW: _____

DATE AND TIME OF
REINTERVIEW APPOINTMENT: _____

REMARKS:

OBSERVATION FORM

P. 45970B

Observed Interviewer Name:

Observer Initials:

Interview Number:

Interview Form: | A | B | C | D |

Specific Observation

Question or item origin

- in Forms A-C:	Gives erroneous feedback						
- in Forms B-D:	Misses reinf. statement						
	Gives inadequate reinf.						
- in all Forms:	Should have probed more						
	Uses inadequate probe						

General Observation

Introduction at door

Asking Questions

___ Looks up in middle of question
___ Doesn't look up at end of question
___ Rushes in reading questions
___ Inconsistent pace for short and long

Using Reinforcing Statements

___ Doesn't look up for R. stat.
___ Doesn't pause after R. stat.
___ Uses too many single R. stat.
___ Lack of variety in selecting R. stat.
___ Lack of naturalness in using R. stat.
___ Lack of R. after Q. 14

Eligible Conditions

Other Comments

PHYSICIAN SUMMARY FORM

Survey Research Center
The University of Michigan
Project 45970B
March 1968

Address: _____

Date: _____

Condition	Definitely or probably present	Definitely or probably not present	Don't know, no information
Chronic skin trouble	☐	☐	☐
Hernia, rupture	☐	☐	☐
Any heart disease	☐	☐	☐
Hypertension	☐	☐	☐
Arthritis, rheumatism	☐	☐	☐
Hay fever	☐	☐	☐
Hemorrhoids	☐	☐	☐
Peptic ulcer	☐	☐	☐
Diabetes	☐	☐	☐
Varicose veins	☐	☐	☐
Asthma	☐	☐	☐
Chronic bronchitis	☐	☐	☐
Stroke	☐	☐	☐

NOTE: The terminology used is that of the questionnaire as asked of respondents.

Please fill out for patients who are: Female
White
Age 17-60, inclusive

ORIGINAL INTERVIEW ILLNESS TABLE

Interviewer Initials ___
Serial Number _____
Page ____ of _____

45970B

NAME OF CONDITION OR SYMPTOM:			
Q FIRST MENTIONED #			
CHECK ALL APPLICABLE	CHECK Q. SOURCE BEGIN Q'S ☐ Q.2-7.....T4 ☐ Q.8.....T1 ☐ Other.....T1	CHECK Q. SOURCE BEGIN Q'S ☐ Q.2-7.....T4 ☐ Q.8.....T1 ☐ Other.....T1	CHECK Q. SOURCE BEGIN Q'S ☐ Q.2-7.....T4 ☐ Q.8.....T1 ☐ Other.....T1
T1. Did you have...(was your ...present) during the past 14 days?	/YES/ ASK T4 /NO/ ASK T2	/YES/ ASK T4 /NO/ ASK T2	/YES/ ASK T4 /NO/ ASK T2
T2. Did you have it during the past 12 months?	/YES/ ASK T3 /NO/STOP QUESTIONS	/YES/ ASK T3 /NO/STOP QUESTIONS	/YES/ ASK T3 /NO/STOP QUESTIONS
T3. For how long did ... last? (GET DURATION OF CONDITION, NOT SINGLE EPISODES)	☐ 3 months or more ASK T4 Less than 3 months ☐ On CClist-ASK T4 ☐ Not on CC list STOP QUESTIONS	☐ 3 months or more ASK T4 Less than 3 months ☐ On CClist-ASK T4 ☐ Not on CC list STOP QUESTIONS	☐ 3 months or more ASK T4 Less than 3 months ☐ On CClist-ASK T4 ☐ Not on CC list STOP QUESTIONS
T4. Did you ever talk to a doctor about ...?	/YES/ ASK T4b /NO/ ASK T5	/YES/ ASK T4b /NO/ ASK T5	/YES/ ASK T4b /NO/ ASK T5
T4b. What did the doctor say it was--did he give it a medical name?	/NO/	/NO/	/NO/
T5. During the past 12 months did you have to cut down on your usual activities because of ...?	/YES/ ASK T5b /NO/ ASK T6	/YES/ ASK T5b /NO/ ASK T6	/YES/ ASK T5b /NO/ ASK T6
T5b. How many days?	_____days	_____days	_____ days
T6. During the past 14 days, at its worst how much pain or discomfort has it caused you?	☐ Very much ☐ Much ☐ Some ☐ Almost none	☐ Very much ☐ Much ☐ Some ☐ Almost none	☐ Very much ☐ Much ☐ Some ☐ Almost none

REINTERVIEW ILLNESS TABLE

Interviewer Initials____
Serial Number_____
Page____of____

45970B

NAME OF CONDITION OR SYMPTOM:			
# Q FIRST MENTIONED			
SOURCE OF ILLNESS (CHECK ALL APPLICABLE)	SOURCE / BEGIN Q's ☐ Q 2-4-7-8-9-10.T4 ☐ Q 11.....T1 ☐ Other.....T0	SOURCE / BEGIN Q's ☐ Q 2-4-7-8-9-10.T4 ☐ Q 11.....T1 ☐ Other.....T0	SOURCE / BEGIN Q's ☐ Q 2-4-7-8-9-10..T4 ☐ Q 11.....T1 ☐ Other.....T0
LEAVE BLANK	/R2/ /2B/ /B/ /S/ /CC/ /Ot/	/R2/ /2B/ /B/ /S/ /CC/ /Ot/	/R2/ /2B/ /B/ /S/ /CC/ /Ot/
T0. Did you first notice . . . during the past 2 weeks or before that time?	/2 weeks/ ASK T4 /before/ ASK T1	/2 weeks/ ASK T4 /before/ ASK T1	/2 weeks/ ASK T4 /before/ ASK T1
T1. Did you have . . . (was your . . . present) during the past 4 weeks?	/YES/ ASK T4 /NO/ ASK T2	/YES/ ASK T4 /NO/ ASK T2	/YES/ ASK T4 /NO/ ASK T2
T2. Did you have it during the past 12 months?	/YES/ ASK T3 /NO/ STOP QUESTIONS	/YES/ ASK T3 /NO/ STOP QUESTIONS	/YES/ ASK T3 /NO/ STOP QUESTIONS
T3. For how long did . . . last? (GET DURATION OF CONDITION, NOT SINGLE EPISODES)	☐ 3 months or more ASK T4 Less than 3 months ☐ On CC list ASK T4 ☐ Not on CC list... STOP QUESTIONS	☐ 3 months or more ASK T4 Less than 3 months ☐ On CC list ASK T4 ☐ Not on CC list... STOP QUESTIONS	☐ 3 months or more ASK T4 Less than 3 months ☐ On CC list ASK T4 ☐ Not on CC list... STOP QUESTIONS
T4. Did you ever talk to a doctor about . . . .?	/YES/ ASK T4b /NO/ ASK T5	/YES/ ASK T4b /NO/ ASK T5	/YES/ ASK T4b /NO/ ASK T5
T4b. What did the doctor say it was -- did he give it a medical name?	/NO/	/NO/	/NO/
T5. During the past 12 months did you have to cut down on your usual activities because of . . . .?	/YES/ ASK T5b /NO/ ASK T6	/YES/ ASK T5b /NO/ ASK T6	/YES/ ASK T5b /NO/ ASK T6
T5b. How many days?	_____ days	_____ days	_____ days
T6. During the past 4 weeks, at its worst how much pain or discomfort has it caused you?	☐ Very much ☐ Much ☐ Some ☐ Almost none	☐ Very much ☐ Much ☐ Some ☐ Almost none	☐ Very much ☐ Much ☐ Some ☐ Almost none

———○○○———

☆ U.S. GOVERNMENT PRINTING OFFICE : 1972 482-005/13

VITAL AND HEALTH STATISTICS PUBLICATION SERIES

Originally Public Health Service Publication No. 1000

Series 1. Programs and collection procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

Series 2. Data evaluation and methods research.—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

Series 3. Analytical studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and committee reports.—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data from the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

Series 11. Data from the Health Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian, noninstitutional population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics; and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data from the Institutional Population Surveys.—Statistics relating to the health characteristics of persons in institutions, and their medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

Series 13. Data from the Hospital Discharge Survey.—Statistics relating to discharged patients in short-stay hospitals, based on a sample of patient records in a national sample of hospitals.

Series 14. Data on health resources: manpower and facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on mortality.—Various statistics on mortality other than as included in regular annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

Series 21. Data on natality, marriage, and divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

Series 22. Data from the National Natality and Mortality Surveys.—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, hospital experience in the last year of life, medical care during pregnancy, health insurance coverage, etc.

For a list of titles of reports published in these series, write to:

Office of Information
National Center for Health Statistics
Public Health Service, HRA
Rockville, Maryland 20852

DHEW Publication No. (HRA) 74-1028
Series 2-No. 45

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
PUBLIC HEALTH SERVICE
Health Resources Administration
5600 Fishers Lane
Rockville, Maryland 20852

OFFICIAL BUSINESS
Penalty for Private Use $300