



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control and
Prevention

National Center for Health Statistics
3311 Toledo Road
Hyattsville, Maryland 20782

August 31, 2015

Margo Schwab, Ph.D.
Office of Management and Budget
725 17th Street, N.W.
Washington, DC 20503

Dear Dr. Schwab:

The staff of the NCHS' Center for Questionnaire Design and Evaluation Research (CQDER) (OMB No. 0920-0222, exp. 07/31/2018), in collaboration with the Division of Research and Methodology, plans to conduct a methodological study exploring if and how data from non-opt-in web panels may be linked to existing NCHS datasets and to quantify measurement error. We are requesting permission for two separate information collections: the evaluation of a sub-set of National Health Interview Survey (OMB Control # 0920-0214) questions, both in their original interviewer-administered mode and a new self-administered mode, using cognitive interviewing methodology; and the administration of a short web survey using a pre-existing, non-opt-in, commercial survey panel (in this case, the Gallup Panel). We propose to start both of these related sub-projects in September 2015.

Proposed Project: The NCHS Research and Development Survey (RANDS)

As the nation's principal health statistics agency, the National Center for Health Statistics (NCHS) is responsible for not only producing high-quality statistics on the health and wellbeing of the American public and the state of the country's health care system, but also for contributing to the development of survey methodologies that will allow the agency to continue producing these health statistics in the future.

To this effect, NCHS' Division of Research and Methodology proposes a methodological survey that will allow the agency to begin the long process of studying whether or not, and how, web panels may be integrated into NCHS' existing survey systems. The over-arching goal of this information collection is to discover new, and to improve existing, methods that will increase data quality in the midst of declining response rates and increased costs. It is well documented that the American public are becoming less likely to respond to surveys. While, on a whole, government surveys enjoy a higher response rate than ones conducted by private organizations, this overall trend is undeniable. Methodological research into new ways to capture respondents while minimizing their burden is necessary to ensure that the government is adapting to this changing survey environment in a proactive and efficient manner. The scope of this project involves both measurement error and sampling strategies. In particular, the project will research how measurement error can be examined in order to inform better-constructed questionnaires as well as laying the foundation for how recruited, web-based samples might be integrated alongside traditional sampling methods.

Linking Dissimilar Datasets. Non-probability sampling and the use of web-based panels are survey methods that have been extensively used in the business and marketing environments.

The American Association for Public Opinion Research (AAPOR), a professional association that focuses on the practical uses of surveying, has released two reports, a report on online panel surveys in 2010, and a report outlining the advantages and challenges of using non-probability samples in 2014¹. Methodological questions regarding the reliability and potential utility for the collection of official statistics remain—even for non-opt-in, recruited panels that claim national representativeness. In an effort to respond to these methodological issues, through a contract with the Gallup Organization, RANDS will examine how the results of a web-administered survey using current National Health Interview Survey (NHIS) questions on the Gallup Panel compare to the results of the current production NHIS. Furthermore, NCHS will attempt to adjust the Gallup Panel results to make them more comparable to the NHIS results, using a variety of modelling techniques. This analysis will move beyond the various attempts in the statistical literature that use demographic or socioeconomic variables to link datasets, and will instead test whether or not topically-relevant (in this case, health conditions and outcomes) variables are more successful.

Measurement Error. Construct validity, or the degree to which a question actually measures a concept that it purports to measure, is one of the key determinants of a survey's data quality and its potential response burden. If a questionnaire's items all have strong construct validity, it will tend to have lower respondent burden and higher data quality, whereas items with low construct validity will increase respondent burden and typically produce less reliable data. Construct validity is typically examined via cognitive interviewing methodology since this qualitative method can determine the phenomena that respondents consider when formulating responses. Thus, cognitive interviewing methodology, with its ability to reveal the substantive meaning behind a survey statistic, can provide critical insight into question performance. Nonetheless, cognitive interviewing methodology has its own limitations and cannot provide an entire picture of question performance. While cognitive interviewing can show that a particular interpretive pattern does indeed exist, it cannot determine the extent or magnitude to which that pattern would occur in a survey sample. Nor can cognitive interviewing studies reveal the extent to which variation of interpretive patterns would occur across various groups of respondents. Additionally, the method cannot fully determine the extent to which respondents experience difficulty when attempting to answer a question. In short, as a qualitative methodology, cognitive interviewing studies lack ability to provide quantitative assessment—a component particularly essential to the field of survey methodology. Likewise, strictly quantitative methods of question and questionnaire evaluation using metrics such as item non-response and missing rates can indicate, but not explain, source of response error. Recent works have attempted to bridge this issue by integrating qualitative and quantitative question evaluation methods through the use of targeted probing and follow-up questions.

Sub-Project 1: RANDS Cognitive Interviewing Study

Background Information about Cognitive Testing of Questionnaires: The methodological design of this proposed study is consistent with the design of typical cognitive testing research. As you know, the purpose of cognitive testing is to obtain information about the processes people use to answer survey questions as well as to identify any potential problems in the questions. The analysis will be qualitative.

Specific Plans for the Proposed Study: The staff of the NCHS Center for Questionnaire Design and Evaluation Research (CQDER) is requesting approval to conduct sixty total cognitive interviews, broken into three separate rounds. All interviews will take place at the CQDER or at a mutually agreed-upon location (such as a library or an office), and all

¹<http://isites.harvard.edu/fs/docs/icb.topic1331066.files/Baker%20et%20al%20Online%20Sampling.pdf>
https://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/NPS_TF_Report_Final_7_revised_FNL_6_22_13.pdf

respondents will be sampled from the Washington DC and Baltimore areas. These interviews will follow the procedures laid out in the CQDER generic ICR (0920-0222).

The first round of cognitive testing will be on the full proposed questionnaire for the first administration of RANDES/Gallup Panel Sub-Project (see the next section below for information on the various administrations of RANDES/Gallup Panel Sub-Project). This questionnaire will be comprised of current NHIS questions from the sample adult and sample adult functional disability questionnaires². This initial questionnaire can be seen in Appendix 1. This round will include up to 20 interviews, and will evaluate the questions in their original, interviewer-administered format.

The second round will evaluate both the usability and validity of the self-administered version of the questions tested in the first round. The self-report versions of the questions will be designed following the first round of cognitive testing. The Gallup Organization (henceforth, Gallup) will program the questions into their survey software and will provide NCHS with test codes. Respondents will then take the survey on a laptop computer while a CQDER interviewer observes and administers either concurrent or retrospective probes. This round will include up to 20 interviews. The questionnaire for this round of testing is not yet available, as Gallup will not program the questionnaire until both OMB and ERB approvals have been granted.

Following the first two rounds of cognitive interviewing, the CQDER will develop “structured probe questions” that will be added to the second administration of the RANDES/Gallup Panel Sub-Project. These structured probes will be designed to elicit the specific patterns of interpretation that a respondent used when answering a proceeding NHIS question on the RANDES³. The third round of cognitive testing will then test the full questionnaire of the second administration of the RANDES/Gallup Panel Sub-Project, which will include both a sub-set of the first questionnaire’s NHIS questions and these structured probe questions.

Recruiting and Informed Consent Procedures for the Cognitive Interviewing Study: We propose to recruit the up-to 60 adult respondents (Age 18 and older) through newspaper advertisements, flyers, special interests groups, word-of-mouth, CQDER Respondent Database, etc. The newspaper advertisement/flyer used to recruit respondents is shown in Appendix 2. Within these constraints, we hope to recruit participants with some demographic (particularly in terms of gender, education, and race/ethnicity) variety. The 5 minute screener used to determine eligibility of individuals responding to the newspaper advertisement/flyer, etc. is shown in Appendix 3a. The screener used to determine eligibility of individuals from the CQDER Respondent Database is shown in Appendix 3b. Note that the wording of the template has been approved and is contained within our umbrella package. Only project specific information has been added to the document. It is anticipated that as many as 72 individuals may need to be screened in order to recruit 60 participants.

Interviews averaging 60 minutes (including the completion of the Respondent Data Collection Sheet) will be conducted by CQDER staff members with English speaking respondents. All interviews conducted in the Center for Questionnaire Design and Evaluation Research will be video and audio recorded to allow researchers to review the behaviors and body language of the respondents. Interviews conducted offsite will only be audio recorded. These recordings will primarily allow researchers to ensure the quality of their interview notes.

² ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2015/english

³ The CQDER has already successfully developed and administered these structured probe questions in previous projects. For example, see Chapter 9 in Miller, Willson, Chepp, and Padilla 2014.

After respondents have been briefed on the purpose of the study and the procedures that CQDER routinely takes to protect human subjects, respondents will be asked to read and sign an Informed Consent document (Appendix 4). Only project specific information has been added to the document. Respondents will also be asked to fill in their demographic characteristics on the Respondent Data Collection Sheet. This document is contained in our umbrella package and is included in this submission, however the burden for completion of this form is captured in the interview.

The interviewer will then ask the respondent to confirm that he/she understands the information in the Informed Consent, and then state that we would like to record the interview. The recorder will be turned on once it is clear that the procedures are understood and agreed upon. The interviewer will then orient the respondent to the cognitive interview with the following introduction:

[Fill name] may have told you that we will be working on some questions that may be on, or will eventually be added to, national and international health surveys. Before that happens, we like to test them out on a variety of people. Most of the questions in this study are about your health history, behaviors, and opinions. We are interested in your answers, but also how you go about making them. I may also ask you questions about the questions—whether they make sense, what you think about when you hear certain words, and so on.

I will read each question to you, and I'd like you to answer as best you can. Please try to tell me what you are thinking as you figure out how to answer. Also, please tell me if:

- *there are words you don't understand,*
- *the question doesn't make sense to you,*
- *you could interpret it more than one way,*
- *it seems out of order,*
- *the answer you are looking for is not provided.*

The more you can tell us, the more useful it will be to us as we try to develop better questions. Okay? Do you have any questions before we start?

The probing in this sub-study will follow the normal semi-structured procedures laid out in the CQDER generic package. Each interview in this sub-study will be followed by a semi-structured debriefing session with the respondent which will allow the interviewer to discuss any inconsistencies that arose as well as additional questions that may have arisen as part of the scripted portion of the interview.

After the interview, respondents will be given the thank-you letter (document contained in umbrella package) signed by Charles J. Rothwell, Director of NCHS, a copy of the informed consent document, and \$40. After the cognitive interview is over, respondents will be asked to read and sign the Special Consent for Expanded Use of Video and Audio Recordings (Appendix 5). There will be no coercion and the respondents will be told that they can call and reverse the decision at any time if they change their minds. If respondents do sign the special consent form they will be given a copy of that as well. Extreme care will be taken with all recordings and paperwork from the interviews conducted off-site. Recordings and identifying paperwork will be stored in a secured travel case until returned to NCHS, at which point they will be transferred to the usual secured locked storage cabinets.

We propose paying participants \$40, which is our standard payment. In total, for this sub-project, the maximum respondent burden will be 66 hours of interviewing.

Sub-Project 2: RANDES/Gallup Panel

The RANDES itself will be conducted for NCHS by Gallup, using their standing, proprietary, recruited panel (known as the Gallup Panel). The survey will be administered twice, and will capture a total of 4000 complete cases. The first administration of the survey will use a questionnaire that is comprised only of current, production NHIS questions, while the second administration will repeat some of these questions while adding an interspersed set of CQDER-developed structured probe questions.

Background Information about the Gallup Panel: The sample for completing the 4000 web surveys will be pulled from the Gallup Panel. Currently, the panel includes about 60,000 members, approximately 50,000 of whom can be contacted via email. Gallup Panel selects potential members using random-digit-dialing (RDD) of landline telephones and cellphones or address-based sampling (ABS) to contact U.S. households at random. During the recruitment call, respondents take a short survey about current event topics, and are asked if they would be interested in participating in additional surveys as a member of the Gallup Panel.

Unlike opt-in panels, the recruitment process for Gallup panel starts with a random sample of telephone numbers or addresses and, as a result, it is possible to derive the selection probability and hence the sampling weight for each respondent on the panel.

There is no time commitment to membership in the Gallup Panel. Rather, households and individuals are encouraged to remain members as long as they are willing and interested. Surveys are administered by an interviewer (over the phone) or are self-administered (either by mail or Web, depending on the Internet accessibility of the respondent). There are no financial incentives for participating in the Gallup Panel, and Gallup's panel performs very well in terms of eliminating "professional" respondents who tend to volunteer for most of the opt-in panels. As with any longitudinal design, Gallup's Panel is also affected by attrition. Significant efforts are taken to retain panelists for as long as possible.

As mentioned above, the Gallup Panel will be used as the sample source for completing the surveys. All Panel participants have been fully screened and a substantial amount of background data have already been collected (e.g., health and well-being, socio-economic and occupational status, media usage, political views, age, gender, race, ethnicity, etc.), which will be attached to the final files delivered by Gallup to NCHS, allowing for extensive non-response bias analysis. Following the delivery of the second dataset and the final methodological report, Gallup will remove all RANDES data from its servers, including backups. This will include not only the responses to the survey itself, but the metadata associated with the RANDES (including response, non-response, participation, and sampling flags). Gallup has extensive cyber and physical security in place in order to protect both the security of the front-end survey interface and the back-end storage of the survey's data. Gallup's security plans are detailed in Appendix 7.

Specific Plans for the Proposed Study: The RANDES itself will be administered twice, with each fielding using a slightly different questionnaire.

The first phase of the RANDES will administer only NHIS questions (that have been adapted for a self-report web mode by NCHS, with advice from Gallup). The non-adapted questionnaire can be seen in Appendix 1. The questions, in the form that they will actually be used on the RANDES, will be pre-tested by NCHS after Gallup has programmed the questionnaire into its Computerized Self-Administered Questionnaire (CSAQ) software, as explained above in Sub-Project 1. The analysis of this phase of the RANDES will focus on whether or not it is possible to statistically manipulate the Gallup Panel data so that it

resembles the data from the production NHIS. The first phase of the RANDS will capture 2000 complete cases, which will be sampled using stratified sampling from the full Gallup Panel. Strata will include age, income, race, ethnicity, and education-based strata.

The second phase of the RANDS will follow the initial analysis of the first phase (as well as the second round of cognitive testing, as explained above in Sub-Project 1). The analytic intent of this phase of data collection is twofold: First, NCHS will use data from this administration of the survey to continue investigating the possible ways online panel data can be matched to data from the production NHIS. Additionally, “structured probe questions” will be added to the RANDS questionnaire in this phase in an attempt to better understand and quantify measurement error for certain RANDS questions. Given this dual-focus purpose, the questionnaire for the second phase of the RANDS will differ from the first phase: questions that do not contribute to the analytic model attempting to link the Gallup and NHIS data will be dropped, and a series of structured probe questions will be added.

While the final questionnaires for the two rounds are obviously not yet finalized, they will be constructed so that its average burden is 20 minutes. Gallup’s own experience with their panel suggests that a questionnaire of 50-60 questions takes 20 minutes on average to complete. The questionnaires will begin with an introduction screen similar to what is seen in Appendix 6, explaining the survey and providing the usual confidentiality and Paperwork Reduction Act language. No written consent will be asked for: consent will be assumed when the respondent clicks from the introduction screen to the first survey screen. We have requested a waiver of signed consent from the NCHS ERB.

Following both phases of the RANDS, Gallup will process the survey data and prepare data files. The data files will not include the respondents’ names, addresses, or any other primary Personal Identifiable Information (PII), including any ISP data Gallup has about the computer from which the respondent replied to the survey. As stated above, all metadata tying the respondents to their inclusion in the RANDS sample will be eliminated from the Gallup servers, including the backups, following final delivery. Again, both the cyber and physical security plan Gallup uses to protect its survey data can be found in Appendix 7. These will be transferred to NCHS via either a secure File Transfer Protocol (FTP) web portal or by loading them directly on an encrypted memory stick. Following confirmation that the transfer is complete and successful, Gallup will delete the data file from their secured servers.

Respondents will not be provided with an incentive to participate in either the first or second phase of the RANDS. Additionally, Gallup does not provide incentives to panel members. The estimated average burden of each phases’ questionnaire is 20 minutes, therefore in total the maximum respondent burden for this sub-study will be 1,333 hours.

Overall Project Burden

A burden table for the entire project consisting of 1399 total burden hours is shown below:

| Form Name | Number of Participants | Number of Responses per Participant | Average hours per response (in minutes) | Response Burden (in hours) |
|-------------------------------------|------------------------|-------------------------------------|---|----------------------------|
| Screener | 72 | 1 | 5/60 | 6 |
| Cognitive Interview Questionnaire & | 60 | 1 | 60/60 | 60 |

| | | | | |
|---------------------------|------|---|-------|------|
| Demographics | | | | |
| RANDS/Gallup Panel | 4000 | 1 | 20/60 | 1333 |

Attachments (7)

cc:

Verita Buie

Tony Richardson

DHHS RCO