

STUDY SERIES
(*Survey Methodology* #2007-1)

**Interagency Collaboration
among the Cognitive Laboratories:
Past Efforts and Future Opportunities**

Elizabeth Martin¹
Susan Schechter²
Clyde Tucker³

Director's Office¹
U.S. Census Bureau
Washington, DC 20233

National Center for Health Statistics²

Bureau of Labor Statistics³

Citation: Elizabeth Martin, Susan Schechter and Clyde Tucker. 1999. "Interagency Collaboration among the Cognitive Laboratories: Past Efforts and Future Opportunities." Pp. 359-387 in *Statistical Policy Working Paper 28: 1998 Seminar on Interagency Coordination and Cooperation*. Washington DC: Federal Committee on Statistical Methodology, Office of Management and Budget. April 1999.

Report Issued: January 26, 2007

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

INTERAGENCY COLLABORATION AMONG THE COGNITIVE LABORATORIES:
PAST EFFORTS AND FUTURE OPPORTUNITIES¹

by

Elizabeth Martin (Bureau of the Census),

Susan Schechter (National Center for Health Statistics), and

Clyde Tucker (Bureau of Labor Statistics)

This paper begins with a brief history of the federal agencies' cognitive laboratories², including a comparison of their common features and differences in scope, focus, and level of effort. We discuss efforts at the three established Federal labs, at National Center for Health Statistics (NCHS), Bureau of Labor Statistics (BLS), and the Bureau of the Census, but should not fail to mention cognitive research efforts at other agencies, such as National Agricultural Statistics Service (NASS) and Government Accounting Office (GAO).

In the second part of the paper, we consider the circumstances which have given rise to collaboration among the labs, namely:

- Collaborative research on jointly sponsored or conducted surveys
- Collaborative efforts in support of content areas of cross-cutting interest, and
- Collaboration to advance methodology or improve survey practice

¹This paper reports the results of research undertaken by staff of the Census Bureau, National Center for Health Statistics, and the Bureau of Labor Statistics. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion. The opinions expressed are the authors' and do not represent the positions of the Census Bureau, the NCHS, or the BLS. We thank Jennifer Rothgeb for helpful comments.

²As our discussant points out, the term "cognitive laboratory" is something of a misnomer, since the research activities conducted by the labs are not restricted to laboratory studies, nor are they exclusively cognitive in focus. However, since the label is commonly used and understood, we use it in this paper.

I. BRIEF HISTORY

A. National Center for Health Statistics.

The impetus behind the Federal government's first cognitive lab, established at Monroe Sirken's initiative at the National Center for Health Statistics in 1985, was a two-part seminar on the Cognitive Aspects of Survey Methodology (CASM) conducted in 1983 and 1984 by the Committee on National Statistics under a grant from the National Science Foundation. The participants were survey researchers and cognitive psychologists from academic institutions and government agencies. The seminar examined a number of developments in psychology that were expected to lead to improvements in survey questionnaires and interviewing procedures, using the National Health Interview Survey (NHIS) as a test case.

Subsequently, the National Science Foundation (NSF) funded a demonstration project applying cognitive knowledge and techniques to improve health surveys. Two distinct programs grew out of this project. The first was the establishment of a Questionnaire Design Research Laboratory (QDRL) in 1985. The second was the establishment of a Collaborative Research Program which sponsored extramural research by university scientists and researchers to study cognitive issues germane to improving the quality of health statistics. Up until 1993, NCHS also had a visiting scientist program which provided opportunities for psychologists, survey researchers, and statisticians to collaborate with NCHS staff in research on the cognitive aspects of surveys. Currently, due to staff shortages, emphasis is mostly on in-house laboratory testing and less on the extramural research program.

NCHS's first cognitive laboratory project was to develop and test the 1987 NHIS supplement on knowledge, attitudes, and practices regarding cancer risk factors (Royston and Bercini, 1987). The questions were difficult to formulate, and the laboratory proved an effective means for identifying conceptual problems which would have gone undetected using traditional questionnaire testing procedures.

Based on these early demonstrations of the effectiveness of cognitive methods in addressing questionnaire problems, the QDRL has been involved in all stages of questionnaire design at NCHS, with the goal of augmenting traditional field pretesting procedures by evaluating questionnaires from the respondent's perspective (Bercini, 1992). The lab develops and tests questionnaires prior to field testing and investigates unresolved problems that remain after the field tests. Table 1 shows tests conducted over the past 4 years under NCHS's OMB clearance which provides for expedited clearance of lab activities. Testing is primarily conducted for NCHS, for other Center for Disease Control and Prevention (CDC) organizations outside of NCHS (most often the National Center for Chronic Disease Prevention and Health Promotion), and for other agencies outside of CDC but usually within Department of Health and Human Services, such as the National Institutes of Health. The surveys tested in 1995-98 were primarily health-related, such as the Behavioral Risk Factor Surveillance System Survey, National Survey of Family Growth, and the National Health Interview Survey, though testing was also conducted for the review of racial and ethnic standards sponsored by the Office of Management and Budget (OMB) and for the Federal Interagency Forum on Child and Family Statistics. NCHS staff also conduct their own methodological research projects to advance the state of the art, and have also worked to standardize and document their methods, for example in a training manual that is widely used at NCHS and by other agencies (Willis, 1994).

The research methods usually used in the NCHS lab are concurrent or retrospective protocol

analysis. In concurrent protocol analysis, a volunteer subject is asked to think aloud as he/she answers questions, and the interviewer probes to determine a subject's interpretation of the question and the recall process used to arrive at an answer. This method uncovers ambiguities in question wording, strategies for dealing with vague questions, or questions that ask for information which is not readily available. In retrospective protocol analysis, the interviewer first administers the entire draft questionnaire, and then probes for reactions to the questions. On rare occasions, focus groups are conducted to discuss general concepts that later are converted into specific survey questions.

NCHS staff also conduct what are termed "mini-pretests" for the National Health Interview Survey. Census Bureau interviewers are brought to the lab and trained, and lab staff accompany them on household interviews in the field. Interviews are taped and behavior coded, permitting more detailed examination of questionnaire and interview problems. This method, which is unique to the NCHS lab, provides a middle ground between lab testing and full scale field tests.

In addition to cognitive methods research, the Office of Research and Methodology conducts research on sampling and nonsampling measurement error and its impact on statistical estimation and analyses, nonresponse and coverage issues, and automated statistical and graphical technology for advanced survey and statistical methods.

B. Bureau of Labor Statistics

The Bureau of Labor Statistics (BLS) began expanding its capabilities for laboratory testing in 1987, when Cathy Dipppo took the lead in planning for a Survey Research Design Center (SRDC), intended as a state-of-the-art, multidisciplinary program of research to improve BLS survey methodology and reduce nonsampling error. A laboratory facility, the Collection Procedures Laboratory (CPRL), also was constructed. BLS also enlisted the services of contractors to assist with studies that required more resources than BLS could provide, and the University of Michigan and Westat, Inc. continue in this support role.

During the first year of its operation, research began for the early phases of the redesign of the Current Population Survey (CPS) (Bienias 1988), and laboratory experiments were conducted to address difficulties respondents had reporting in the Consumer Diary Survey (CED) (Tucker et al. 1989). Work was conducted on the Survey of Employer Practices (Palmisano, 1988), the Current Employment Survey (CES), and the Consumer Expenditure Interview Survey (CEI) (Lessler 1988). Over the years, the scope of research expanded to include more work on establishment surveys, research on confidentiality, new technology, nonresponse, as well as an expanded program of research on usability testing of human-computer interfaces.

The research undertaken at BLS over the past ten years also has been facilitated by a generic OMB clearance for methodological research. Table 2 shows the number of research projects cleared under the clearance from 1991 to 1997. These numbers show an increase in the amount of research conducted at BLS, with research projects split almost evenly between work on establishment surveys and household surveys.

Current research at BLS involves pretesting various questionnaires, including the Telephone Point-of-Purchase Survey, the Consumer Expenditure Quarterly Interview, and a time-use supplement for the Current Population Survey. Forms design for establishment surveys also is a common task. In particular, work is ongoing with the North American Industry Classification System (NAICS) implementation and the Occupational Safety and Health Administration (OSHA) log for reporting job-related injuries and illnesses. Recently, efforts have been

underway to conduct usability tests for new survey software and website design. Theoretical work remains a priority. Studies of the advantages and disadvantages of conversational interviewing have been done, and the development of micro-level indicators of measurement error is being explored.

BLS uses a number of methodological techniques. One-on-one cognitive interviews are commonly done when pretesting a questionnaire. They may involve either concurrent think-aloud or retrospective recall techniques. Focus groups often are used in the early stages of developing a new survey or to investigate the experiences of interviewers and survey respondents. Field tests also have been used, such as for developing a time-use survey. The field tests often include respondent debriefings. BLS also engages in a great deal of large-scale secondary analysis of survey results to uncover measurement problems.

C. The Census Bureau

The Census Bureau's questionnaire design and cognitive testing activities originated with Donnie Rothwell's program of laboratory research on self-administered census questionnaires prior to the 1970 census (Rothwell, 1983, 1985). Although they were inspired by marketing research rather than cognitive psychology, several research methods introduced informally then foreshadow methods that are applied more systematically now. The main methods were laboratory experiments, and observations. In the experiments, or "classroom tests," groups of people were assembled in connection with decennial census outreach activities, and each participant was randomly assigned one of several versions of a census questionnaire to fill out. Afterwards, they filled out a "What-do-you think" questionnaire, asking for opinions about the census form. This method resembles what today we call "respondent debriefing." Because these early experiments were true factorial designs, comparisons among treatments permitted inferences about which questionnaire versions or design features obtained more complete data and were more acceptable to respondents.

The second principal method was observations conducted in informal group settings. After a group discussion about mail questionnaires, an observer would pair up with one or two participants, and observe them as they each filled out a questionnaire. Observers encouraged participants to talk aloud as they filled out questionnaires--an early and less structured version of what we now call a think-aloud interview.

Since the early work on self-administered questionnaires, the scope and methods of the work carried out by the Census Bureau's Center for Survey Methods Research (which was created in 1981) have expanded. In 1988, its laboratory facility was established, and procedures and training began to be somewhat more standardized and formalized. In 1991, OMB approved a generic clearance process for questionnaire pretesting, which eliminated the clearance requirement for small tests involving cognitive or think-aloud interviews, behavior coding, respondent debriefing, focus groups, or split-sample experiments. This clearance covers small-scale tests for economic, demographic, and decennial areas of the bureau, and removed a major barrier to more thorough pretesting. Center staff developed a protocol for pretesting (DeMaio et al., 1993), and in 1995 the Census Bureau established a policy that all new or revised questions must be pretested using one of the more rigorous pretesting techniques just mentioned before being fielded in a demographic survey (Bureau of the Census, 1998).

Questionnaire design research and pretesting are a heavy focus of CSMR's research activities. Laboratory as well as field testing covers both interviewer-administered and self-

administered surveys and censuses. Research activities have also included ethnographic research on coverage and household rostering, record check studies and other investigations of measurement error, research on public attitudes related to privacy and census participation, and other methodological and social science topics. Currently, it is expanding its research to include usability testing of human-computer interfaces, and building a new lab facility to accommodate that need. It is also currently implementing its own test vehicle for conducting split-panel questionnaire experiments, the Questionnaire Design Experimental Research Survey (QDERS) (1998, SRD). Research has been conducted on a great diversity of surveys, including the Current Population Survey, the Survey of Income and Program Participation, Survey of Program Dynamics, the American Housing Survey, Schools and Staffing Survey, the American Community Survey, the Continuing Survey of Food Intakes for Individuals, and many others.

At present, 27 people are employed in an interdisciplinary staff that includes psychologists, sociologists, anthropologists, and survey methodologists. Work is conducted for internal sponsors and (under reimbursable agreement) for external sponsors of Census Bureau surveys, as well as other agencies and organizations who sponsor methodological work independent of any particular survey. The Center provides in-house training on cognitive interviewing and sponsors training on other methods.

Table 3 provides information about pretests conducted by the Census Bureau since the inception of the generic pretesting clearance, based on annual reports to OMB prepared by Terry DeMaio. Most pretesting has been conducted on decennial or demographic surveys, with only a few tests per year involving economic surveys. Most tests involved questionnaires, though a handful involved survey procedures. Most tests used cognitive interviews, alone or in combination with other methods. In 1997, 2,203 burden hours were used under the generic pretesting clearance.

Table 4 summarizes the current situation of the three cognitive labs, including their commonalities and their differences.

II. COLLABORATIVE ACTIVITIES

Collaboration among the cognitive laboratories has primarily occurred in 3 circumstances:

- ▶ First, as part of a program of research on jointly sponsored or conducted surveys,
- ▶ Second, collaborative development and testing in support of content areas of cross-cutting interest, and
- ▶ Third, coordinated efforts to advance methodology and improve survey practices.

EXAMPLE 1: CPS REDESIGN

A good example of interlaboratory collaboration of the first type was focussed on the 1994 redesign of the Current Population Survey. In 1986, senior executives and researchers from the Bureau of Labor Statistics and Bureau of the Census met to discuss the future of the CPS and plan a comprehensive redesign of all aspects of the survey, including its sample, automated interviewing, a redesigned instrument, and a modernized processing system. Six areas of research were identified to move the CPS toward the leading edge, including questionnaire

design research and application of cognitive psychology research methods to improve the CPS questionnaire. The high priority given this effort reflected several considerations. First, over the long history of the CPS, its conceptual foundation had been well-elaborated in terms of the technical definitions of key labor force constructs, such as unemployment. However, despite refinements of the concepts, the questionnaire had remained almost unchanged since a redesign in 1967. Second, several presidential commissions had recommended revisions and additions to CPS which had not been incorporated. Third, an interagency task force reviewed the CPS instrument, and identified cognitive and other problems in the instrument, including problems with terminology or concepts in the labor force questions, reliance on volunteered information, subjectivity, and difficult tasks of recall and estimation (BLS and Bureau of the Census, 1986). BLS and Census developed and embarked on a research agenda which included both laboratory and non-laboratory research as well as a multi-year program of split-sample field experimentation and evaluation of questionnaire alternatives.

The CPS Redesign was much of the impetus behind establishment of BLS's cognitive lab, and BLS's earliest cognitive research focussed on CPS (e.g., Beinias, 1988). The Census Bureau conducted debriefing surveys of respondents and interviewers which confirmed some of the conceptual ambiguities which had been suspected (Campanelli, Rothgeb, and Martin, 1989). A thorough questionnaire revision was undertaken to improve labor force measures and take advantage of the capabilities of automation, and redesigned questionnaires were experimentally tested in two successive large-scale RDD field tests in 1990 and 1991 (Rothgeb et al., 1992). After each field experiment, results were used to determine which of several alternative versions performed better, and to further refine and improve question wordings. Evaluation of the performance of alternative versions was based upon traditional measures of data quality, such as item nonresponse and response distributions, but also used respondent debriefing, vignettes, and behavior coding data to evaluate whether revised questions achieved better reporting and a closer match between respondents' understandings of questions and the intended technical definitions. The final, redesigned questionnaire was evaluated in an overlap sample in 1992-3 before being fully implemented in 1994.

The level of effort and degree of collaboration between the two agencies in the research involved in redesigning and testing the CPS instrument was extraordinary. Most papers and reports included authors from both agencies. Regular meetings of an interagency working group chaired by associate commissioner Jack Breggar were held to review and discuss virtually every questionnaire revision and testing result over a period of several years, and in fact the same group, with a different name but many of the same participants, still meets to this day, to discuss and plan research relevant to keeping CPS a leading edge survey. Finally, this collaborative effort was unusual in the extent to which senior staff from BLS were not only supportive of but directly involved in the questionnaire revision and evaluation process. It is an indication of the seriousness with which BLS officials took and still take the measures on which labor force statistics are based. One consequence is that the conceptual underpinnings of the Current Population Survey are unusually well-articulated and thought through, and revisions to the survey are thoroughly researched and evaluated.

EXAMPLE 2: PRINT DISABILITY RESEARCH

A second circumstance in which collaborative work between the cognitive labs has been

common is on content areas of mutual, often government-wide, interest. The flexibility permitted by the labs' generic OMB clearances has facilitated broad, interagency efforts, and has made it possible for the statistical agencies to respond quickly to demands for development and testing of new content areas. For instance, the cognitive research supporting the OMB research program to revise standards for collecting race and ethnicity data was highly collaborative, involving more than 8 agencies and all three labs; the Census Bureau alone conducted 8 separate tests in connection with this effort. A second example is research on disability, which is described in the other paper in this session. A third example, also related to disability, is less well-known, but interesting in several ways, and may point to areas where collaborative efforts need to be intensified.

A small laboratory project was conducted jointly by NCHS and the Census Bureau to evaluate measures of print disability, which is operationalized as the inability to see well enough to read ordinary newspaper print. Different measures of print disability are used in the Census Bureau's Survey of Income and Program Participation, and in the National Health Interview Survey, sponsored by NCHS. The different measures resulted in over a two-fold difference in estimated prevalence of print disability among people 15 and older--9.7 million people have difficulty reading newspaper print according to SIPP, while 4.2 million people cannot read newspaper print according to NHIS (Beatty and Davis, 1998). A private foundation, the American Federation for the Blind, sponsored a methodological study to investigate possible cognitive and questionnaire sources of the difference in the survey estimates. Based on a small-scale experiment using cognitive interviews, Beatty and Davis (1998), found that one likely reason the NHIS and SIPP obtain different estimates is because slightly different questions are asked, and the questions measure different degrees of disability. Additionally, in the NHIS, the question is preceded by other questions on vision which may influence respondents to give more sanguine reports about their ability to see newspaper print. Finally, the SIPP clearly establishes that the question is about respondents' ability to see with glasses (if they ordinarily wear them), whereas NHIS is not so clear. Although the cognitive research study was small--only 30 subjects--the results are interesting and suggestive, and it would be desirable to replicate the study in a larger field experiment, to compare and evaluate the quality of the data produced by the different questions. If compatible with NHIS and SIPP objectives, it also would be desirable to use more comparable measures in both surveys in order to produce more consistent estimates. The current use of noncomparable measures appears to be a source of some confusion about the prevalence of print reading disability.

Several features of this collaborative effort deserve comment. First, this research once again highlights the fact that rather minor differences in question wording and context can have potentially large effects on survey estimates. Despite this fact, most government surveys do not invest the substantial time and effort in the research needed to develop and test valid and reliable measures; it is probably fair to say that most questions currently fielded in government surveys were never rigorously pretested. One result is that errors and problems slip through which may affect the quality of government statistics. For example, in the NHIS the print disability question has been asked for persons of all ages; one wonders how respondents answer a question about ability to "see well enough to read" for children who aren't old enough to read. We don't know, because the research has not been conducted. Similarly, one wonders how either SIPP or NHIS respondents answer these questions for persons who are not literate, but again, we don't know because the issue hasn't been addressed in research.

A second notable feature of this research is that it was paid for by the American Foundation for the Blind, through The Library of Congress's National Library Service for the Blind and Physically Handicapped. It is unusual that a private foundation funds Government research, and it indicates that the discrepancy between the two estimates was bothersome enough to some data users and advocates that they were moved to fund methodological research to understand the problem. One might have supposed that the sponsoring agencies would have taken it on themselves to document and investigate and perhaps resolve the discrepancies between their measures, but this was not the case. This would not be an easy job, since there are many differences among the major surveys--interviewing mode, sample design, and so on--which, in addition to the questions themselves, may contribute to different estimates. Special studies would be required to tease apart the causes of noncomparable estimates. Nonetheless, it is uncommon for statistical agencies to attempt to achieve comparability across surveys or, sometimes, to even notice that measures and estimates are not comparable. The case just described, which is not unique, suggests there is an unmet need for more methodological research to address substantive measurement problems. Research to develop and scientifically evaluate important survey measures should be conducted more systematically than is now the case. This kind of research provides opportunities for collaboration among the sponsoring agencies, and requires collaboration between methods experts and subject matter specialists as well. The different results obtained by NHIS and SIPP print disability measures are troublesome, but also provide an opportunity for comparative research that can shed light on the measurement properties of the different measures, perhaps leading to improvements.

EXAMPLE 3: CASM II

A third type of collaboration among the labs is coordinated efforts to advance methodology and improve standards of practice in government surveys.

One way this occurs is that the labs benefit from what might be termed "a useful competition." We have learned from each other, and if one lab introduces an innovation or sets a new standard, this spurs the other labs to upgrade their capabilities and intensify their efforts. We believe that the quality of our work, and our receptiveness to new ideas, have been enhanced by this informal competition as well as by our cooperative efforts. Also, sometimes our home agencies have been moved to invest in our facilities and raise standards because otherwise they might appear to be falling behind the other statistical agencies.

A second form of collaboration is coordinated, interlaboratory efforts to document, evaluate, and validate the methods used in the cognitive laboratories. Interlaboratory replication is adopted in the physical sciences as a way of evaluating the vulnerability of basic physical and chemical measurements to different conditions of measurement, and to provide quality control for the laboratories. This sort of effort has begun in an informal way. For example, over the past 2 years, NCHS has been involved in a study initiated by Gordon Willis to compare expert reviews of questionnaires and cognitive tests conducted independently by the 3 labs, in order to assess interlaboratory variation in how methods are applied and what results they yield. This sort of collaboration should be expanded.

The laboratories can also cooperate in expanding into new areas and adopting new methods, especially when fresh perspectives and scientific ideas are introduced into the mix. We mention just one example of such a collaborative effort. Two years ago, NCHS sponsored a Second

Advanced Seminar on Cognitive Aspects of Survey Methodology, or CASM II, held in June of 1997 in Charlottesville, VA. The conference was organized by our three agencies, as well as survey methodologists and cognitive researchers from outside the government. It was a reprise of the conference held in St. Michael's 14 years earlier, and its purpose was to assess progress and chart next steps in integrating cognitive and other sciences and survey methodology. The conference explored a diverse set of new areas for research, for example applying computational models to the design of questionnaires, applying cognitive methods to statistical graphs and data analysis, and using conversational analysis to examine the interview process. A number of papers were prepared for discussion at the conference, and conference attendees participated in working groups to develop research agendas for new applications of cognitive science to survey methodology--for example, in establishment surveys, and to improve measures of income and disability (see Sirken, Jabine, Willis, Martin, and Tucker, in press; and Sirken, Herrmann, Schechter, Schwarz, Tanur, and Tourangeau, in press).

III. LESSONS LEARNED

Greater investment by the statistical agencies in testing of survey measures and research to improve them is needed to sustain and improve data quality. The cognitive labs are in an especially advantageous position to identify and take advantage of opportunities for cross-cutting research, since they do not specialize in particular surveys. Their central position in their home agencies also permits them to take a more detached, objective view of measurement problems and inconsistencies. Since they have no stake in any particular survey's approach, the labs can and often do play the role of honest broker. In addition to laboratory testing, more field testing and experiments are needed, and the lab staffs are experienced in these as well as laboratory methods.

Despite their critical role and their potential to contribute substantially more to survey methods research, the labs are not staffed sufficiently to meet many of the basic agency needs for their services. Even with the necessary support provided by outside contractors, the combined level of effort in the statistical agencies is insufficient to provide cursory pretesting for all government surveys, much less more intensive research on measurement issues. There are only 56 staff in all three agencies working full or part-time on research activities in the cognitive labs, and this includes management and support staff. Though the Census Bureau has the largest staff, the Census Bureau's pretesting policy is not always followed, in part because the staff resources are not sufficient to fully support the effort. It is particularly unfortunate to see the dwindling of the research staff of the cognitive laboratory at NCHS. This lab is the only government lab involved in testing health-related measures, yet only 3 people are working there as of this writing, and the agency does not have a policy or a practice of cognitively testing all questions in its surveys.

We see three areas where there is the greatest potential gain from more intensive collaborative efforts among the cognitive laboratories:

- ▶ **coordinated programs of cognitive research and field experiments on measures of common interest**

Here, research on race and ethnicity might serve as a model for other areas in which a coordinated program of research might yield important improvements in the quality of data,

as well as basic understanding of the sources of error affecting the data.

▶ **interagency research to evaluate testing protocols and methods and to replicate testing results**

As Willis, DeMaio, and Harris-Kojetin (in press) note, more research is needed to evaluate and validate the methods currently being used to design and evaluate questionnaires. Much of this research could be fruitfully be conducted collaboratively, and interlaboratory replication could provide quality control for routine tests conducted by the labs.

▶ **collaboration to develop and set government-wide standards for surveys--for example, standards for pretesting questionnaires**

Currently, there is no agreement among statistical agencies about basic standards for conducting surveys. For example, the practice and policy for pretesting questionnaires prior to a field survey are variable both within and across agencies. The quality of survey data collected by the Government would be improved if all questionnaires were pretested before they were fielded, in order to identify and correct problems of interpretation and understanding, lack of knowledge, recall problems, sensitivity, and other problems which make it difficult for respondents to provide accurate answers to survey questions. Addressing such problems, even minimally, would also reduce the burden of participating in government surveys.

Table 1. Projects conducted under NCHS Generic OMB Clearance for 1995-98

Sponsor	Number of projects ¹	Number of subjects ²
NCHS Requests National Health Interview Survey National Survey of Family Growth National Health and Nutrition Examination Survey Second Longitudinal Study on Aging	24	269
CDC requests Behavioral Risk Factor Surveillance System Survey National Immunization Survey Survey on Intimate Partner Violence Survey on Environmental Factors	12	71
Non-NCHS or CDC requests National Cancer Institute projects OMB Race & Ethnicity project Forum on Child and Family	9	61
Total	45	401

¹The count of projects refers to each request made to test either a new instrument or part of a new instrument. Only pretests conducted in the QDRL are included; extramural research is not included.

²Each subject represents one burden hour.

Table 2. Number of BLS Research Projects Conducted under OMB Generic Clearance 1991-97,
by Area and Year

Type	FY 1991	FY 1992	FY 1993	FY 1994	FY 1995	FY 1996	FY 1997	TOTAL
Household surveys	3	3	3	8	5	11	12	45
Establishment surveys	0	2	2	7	7	3	8	29
Total	3	5	5	15	12	14	20	74

Table 3. Number of Census Bureau Projects conducted under the Generic OMB Pretesting Clearance 1991-97, by Area and Year

	9/91-9/92	10/92-12/93	1994	1995	1996	1997	Total
Decennial	1	6	5	9	11	4	36
Demographic	1	12	11	8	6	13	51
Economic	3	2	5	1	5	1	17
Decennial/demographic (OMB race-ethnicity project)		3	1	4			8
Total	5	23	22	22	22	18	112
Burden hours	263	1252	3024	1778	2013	2203	

Note: Census Bureau projects involved one or more of the following methods: cognitive interviews, focus groups, respondent debriefings, split-ballot experiments, behavior coding of field interviews, or ethnographic interviews.

Recent (1997) Census Bureau projects covered by generic pretesting clearance:

- ▶ Special places facility questionnaire for Census 2000
- ▶ Decennial questions on disability
- ▶ Decennial questions on race and ethnicity
- ▶ Decennial questions of grandparents as caregivers
- ▶ Communication channels and messages for encouraging census response
- ▶ Self-administered roster research
- ▶ Focus groups on barriers to enumeration of American Indians
- ▶ Service-based Enumeration
- ▶ Survey of Income and Program Participation
- ▶ Current Population Survey March Income Supplement
- ▶ Survey of Program Dynamics
- ▶ American Housing Survey
- ▶ Print disability measures on NHIS and SIPP
- ▶ Survey of Minority Owned Business Enterprises
- ▶ Survey of Employee Fringe Benefits
- ▶ Schools and Staffing Survey
- ▶ Census of Juvenile Detention, Correctional, and Shelter Facilities
- ▶ Juvenile Facility Questionnaire
- ▶ Integrated Postsecondary education data system

Table 4. Features of Cognitive Labs at Three Agencies

	NCHS	BLS	CENSUS
<i>Common Features</i>			
In-house laboratory facility	✓	✓	✓
Respondents paid	✓	✓	✓
Similar recruitment strategies: flyers, leaflets, ads, by word of mouth, through community organizations	✓	✓	✓
Interviews are audio/video-taped	✓	✓	✓
Interdisciplinary staff (social science, psychology, survey methodology)	✓	✓	✓
Small tests covered by OMB generic clearance	✓	✓	✓
Methods used:			
Cognitive interviews	✓	✓	✓
Focus groups	✓	✓	✓
Respondent debriefing	✓	✓	✓
Split-panel experiments	✓	✓	✓
Behavior coding	✓	✓	✓
Unstructured/ethnographic interviewing		✓	✓

	NCHS	BLS	CENSUS
<i>Differences</i>			
Scope	Primarily health topics	Primarily economic topics	Broader: housing, income, labor force, nutrition, etc.
Type of survey	Household surveys	Household and establishment surveys	Household + some establishment
In-house training/manual of cognitive methods	Yes	No	Yes
Current staff size	4	15	27
1997 burden hours conducted under generic OMB lab clearance ¹	119	2,903	2,203
1997 tests conducted under generic lab clearance ²	11	17	18
Does agency have an official pretesting policy?	No	No	Yes for demographic surveys; de facto, for decennial; no for economic area

¹The burden hours given in the table are not fully comparable across the three agencies. The NCHS burden hours represent the number of laboratory interviews and do not include field activities. At BLS, the generic clearance covers many types of research activities, including field tests of new instruments. At the Census Bureau, the generic clearance covers testing activities that involve focus groups, cognitive interviews, respondent debriefings in a field setting, behavior coding, or split-ballot experiments involving a few hundred cases. All the labs engage in research activities that are not covered by their generic clearances, and hence not represented in these tables, either because they do not involve any collection of data, or because they were larger scale efforts covered under separate clearances.

²Some tests were contracted out: 0 for NCHS, 2 for Census, and 2 for BLS.

REFERENCES

- Beatty, P., and W. Davis, (1998), Evaluating Discrepancies in Print Reading Disability Statistics through Cognitive Interviews, Working Paper Series No. 25, Hyattsville, MD: National Center for Health Statistics, and Bureau of the Census Working Paper in Survey Methodology No. 98/1.
- Bienias, J.L. (1988), "Preliminary Laboratory Research on the Current Population Survey Questionnaire with Employed Persons," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 185-190.
- Bercini, D., (1992), "Pretesting Questionnaires in the Laboratory: An Alternative Approach," Journal of Exposure Analysis and Environmental Epidemiology (2):241-248.
- Bureau of the Census, (1995; reissued 1998), Pretesting Policy and Options: Demographic Surveys at the Census Bureau. Washington DC: Bureau of the Census.
- Bureau of Labor Statistics and Bureau of the Census, (1986), Report of the BLS-Census Bureau Questionnaire Design Task Force.
- Campanelli, P.C., Rothgeb, J.M., and Martin, E.A. (1989), "The Role of Respondent Comprehension and Interviewer Knowledge in CPS Labor Force Classification." Proceedings of the Section on Survey Research Methods, American Statistical Association.
- DeMaio, T. J., N. Mathiowetz, J. Rothgeb, M. E. Beach, and S. Durant, (1993), Protocol for Pretesting Demographic Surveys at the Census Bureau. Bureau of the Census.
- Lessler, J.T. (1988), Consumer Expenditure Interview Survey, Preliminary Cognitive Laboratory Studies, Research Triangle Institute, Research Triangle Park.
- Palmisano, M. (1988), "The Application of Cognitive Survey Methodology to an Establishment Survey Field Test," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 179-184.
- Rothgeb, J. M., Polivka, A. E., Creighton, K. P., and Cohany, S. R. (1992), "Development of the Proposed Revised Current Population Survey," Proceedings of the American Statistical Association (Survey Research Methods Section).
- Rothwell, N. D., (1983), New Ways of Learning How to Improve Self-Enumerative Questionnaires: A Demonstration Project, unpublished manuscript, Bureau of the Census.
- _____, (1985), "Laboratory and Field Response Research Studies for the 1980 Census of Population in the United States," (1985), Journal of Official Statistics 1(2), pp. 137-157.

Royston, P. and D. Bercini, (August, 1987), "Questionnaire Design Research in a Laboratory Setting: Results of Testing Cancer Risk Factor Questions," paper presented at the annual meeting of the American Statistical Association (Survey Research Methods Section)

Sirken, M., Jabine, T., Willis, G., Martin, E., and Tucker, C., (Eds.), (In press), A New Agenda for Interdisciplinary Survey Research Methods: Proceedings of the CASM II Seminar, National Center for Health Statistics.

Sirken, M., Herrmann, D., Schechter, S., Schwarz, N., Tanur, J., and Tourangeau, R., (Eds.), (In press), Cognition and Survey Research. New York: Wiley.

Statistical Research Division, (1998), Steps Towards a Five Year Plan. Bureau of the Census: Statistical Research Division.

Tucker, C., Vitrano, F., Miller, L. A., and Doddy, J. (1989), "Cognitive Issues and Research on the Consumer Expenditure Diary Survey," paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL.

Willis, G., (1994), Cognitive Interviewing and Questionnaire Design: A Training Manual. Working Paper Series No. 7, Hyattsville MD: National Center for Health Statistics.

Willis, G., DeMaio, T., and Harris-Kojetin, B., (in press), "Is the Bandwagon Headed to the Methodological Promised Land? Evaluating the Validity of Cognitive Interviewing Techniques," in Sirken, M., Herrmann, D., Schechter, S., Schwarz, N., Tanur, J., and Tourangeau, R., (Eds.), (In press), Cognition and Survey Research. New York: Wiley.