

# **INSTRUMENT REDESIGN: TRYING TO MAKE A MULTITUDE OF USERS HAPPY WHILE DEVELOPING A HIGH QUALITY SURVEY INSTRUMENT**

**John Tsapogas and Joseph P. Gannon**

## **1. Introduction**

The National Science Foundation began a redesign of the Scientific and Technical Personnel Data System (STPDS) in the late 1980s. The redesign effort included an investigation of all STPDS survey efforts including survey operations, questionnaire content and format, and sample design. Since the late 1980s research has been conducted to assist NSF in making the correct decisions on the redesign of the STPDS. This paper explores one part of the redesign effort: the questionnaires.

We will discuss the research that has been conducted on the questionnaires that were used in the 1980s, determine the impact this research has had on the development of new survey instruments for the 1990s, and show the measures taken to test the new survey instruments. The paper discusses the strategies that were used to come up with improvements and the methods used to implement the changes that were recommended.

## **2. Background**

### **Questionnaire content in previous years**

The various surveys of the STPDS were conducted by different contractors during different years. The U.S. Bureau of Census conducted the National Surveys of Natural and Social Scientists and Engineers (Postcensal survey) for the National Science Foundation in 1982, 1984, 1986, and 1989. The Institute for Survey Research of Temple University conducted the Survey of Natural and Social Science and Engineering Graduates (New Entrants Survey) in 1982, 1984, 1986, and 1988 and the National Research Council conducted the Survey of Doctoral Recipients (SDR) in 1981, 1983, 1985, 1987, and 1989.

The content areas of the STPDS surveys collected in the 1980s was influenced in large part by strategies and designs developed in the late 1970s. The Proceedings of the NSF Conference on the Scientific and Technical Personnel Data System: Strategy and Design detail some of the major topic areas that were instrumental in developing the content areas of the 1980s questionnaires.

The content areas covered in the questionnaires in each survey did not vary widely in size and scope during the 1980s. Except for some minor additions and deletions of questions, the data collected in the beginning of the decade was consistent with the data collected at the end of the decade, for each of the individual surveys. Consistency of content areas among the surveys was a different matter with the Postcensal and the New Entrants survey content areas being fairly high, while differences did prevail between the SDR and the other two surveys.

The record shows that while these surveys were set up as a system the questionnaires were far from comparable. There was consistent coverage for some questions during the 1980's, but for other questions there was less consistent and less frequent coverage.

### **Sequencing of major topic areas of the questionnaires in previous years**

The sequencing of major topic areas did not maintain any consistency across surveys by years nor across years by survey. The Postcensal/experienced S&E surveys and the New Entrants surveys maintained some consistency over time in that the education (and training) section always preceded the employment status section which always preceded the employment profile section and the other information section. The demographic characteristics section moved to various locations in the two surveys eventually finding a location in the beginning of the questionnaire by the end of the decade, the same order it held for both surveys at the beginning of the decade. The major topic areas of the Survey of Doctoral Recipients, employment profile and demographic characteristics shifted their order in 1983 and 1985, only to shift back to their original order in 1987 and 1989.

### **Formatting features of STPDS questionnaires in previous years**

There weren't any major formatting differences among the STPDS survey instruments of the 1980s. All the questionnaires were printed in standard size booklet form except for the New Entrants questionnaire which was printed in undersized booklet form. All questionnaires were printed on white paper but used different ink color, dark blue ink for the Postcensal/Experienced Sample Survey, blue ink for the New Entrants Survey and brown ink for the Survey of Doctorate Recipients. The formatting of the front cover varied among the STPDS surveys, with the Postcensal containing a respondent letter with instructions but no survey questions, the New Entrants containing no letter, no instructions, and no survey questions, and the SDR front page including no respondent letter or instructions but including survey questions.

### **3. Goals of the questionnaire redesign effort**

Building on previous research, NSF staff and its consultant (Mathematica Policy Research, Incorporated) evaluated the questionnaire redesign effort and developed several goals in the development of the questionnaires.

**Understanding the unique needs of each survey's population:** Since the redesign effort involved three questionnaires, the redesign of the questionnaires required an understanding of the population of each survey (college graduates, recent bachelor's and master's degree recipients, and doctoral degree holders). For example the education and training questions were much more important questions for recent bachelor's, master's, and doctoral degree recipients than they were for the experienced sample population and as a result should be the first set of questions in the questionnaires for those surveys.

**Emphasis on the STPDS as a system:** The STPDS is a system of surveys that develops information on each of the populations surveyed and uses survey results to develop national estimates. Therefore, it is important that core questions be included in each survey so that data developed for each population can be compared and national estimates can be developed.

**Enhanced respondent participation:** A major criticism of the STPDS surveys of the 1980s is the low response rates in those surveys. In 1986 the response rates as a percent of the original sample were 44% for the Postcensal/Experienced Sample Survey, 48% for the New Entrants Survey, and 58% for the Survey of Doctorate Recipients. A major goal of the questionnaire redesign effort was to develop questionnaires that enhanced respondent participation.

**Minimum respondent burden:** The questionnaire had to be designed in such a way as to minimize respondent burden. Features such as questionnaire size, overall format, color and weight

of the paper, cover letter, front page cover, page format, and questionnaire length had to be appealing to the respondent and contribute towards easy completion of the questionnaire.

**Clear and unambiguous questions:** Questions needed to be stated clearly and simply in order to enhance data accuracy, avoid response bias, and encourage respondents to complete the survey.

**Uniformity of questions among all the STPDS surveys:** Since composite estimates were developed from the surveys, consistency of questionnaire wording had to exist among the various STPDS surveys.

**Historical comparability:** Since a redesign involves some changes from the past, the redesign effort requires some decision making on the questionnaire items that need to be maintained so that comparability can be maintained with previous surveys. In the 1990s the STPDS will need to maintain a link with the surveys of the 1980s.

**Consistency with other surveys:** In order to enhance data comparability and utility to the extent practicable and to reduce duplication of efforts and costs, NSF should make an effort to design surveys that are consistent with other surveys. Specifically some comparability should exist, where feasible, between the questions in the STPDS surveys and those of the Bureau of Census' Current Population Survey, the Decennial Census, and the National Center of Education Statistics' Recent College Graduate Survey.

**Minimize S&E bias in both response rate and in questionnaire design:** The orientation of the Postcensal/Experienced Sample survey and the New Entrants Survey reflected in the survey titles printed on the questionnaires, the cover letter, and other features of the questionnaire packages contributes to an upward bias to the NSF estimates of the S&E population. The questionnaire packets must be designed in such a way as to limit or eliminate this S&E bias.

#### **4. Recommendations from studies of the STPDS and their impact on STPDS instruments**

##### **Data Users Conference**

The National Science Foundation held a data users conference during May 9-10, 1985 in Washington, D.C. to explore current and emerging issues with respect to the scientific and technical personnel. The objective of the conference was to identify the type of information that is required to address current and emerging policy issues, establish priorities among competing demands for data and to identify data that is presently being collected that may no longer be needed. The conference generated many recommendations about proposed data collection activities.

Some of the most important suggestions included the need for data on the educational and post-degree activities of foreign students in the United States and data on new emerging technologies and information on personnel and their work activities that support these new emerging activities. Many users also expressed an interest in data on the career paths of scientists and engineers with detail on crossfield transfers, retraining needs, geographic mobility, and dual-career families. Great emphasis was placed on coming up with methods of measuring the quality and output of scientists and engineers, salary information, geographic information such as location of employment, and finer disaggregation within fields of science. Many of the recommendations were ultimately included in our final survey instrument.

## **Engineering Workshop on Data Needs for the 1990s**

The National Science Foundation funded a workshop in which the topic was “Data Needs in the 1990s for Monitoring Labor Market Conditions for Engineers”. The workshop which was held on May 28, 1988 addressed four major issues: (1) occupational mobility and flow dynamics, (2) international flows of engineers, (3) technical currency and (4) the role of underrepresented groups in engineering.

Many of the recommendations of the workshop were similar to those of the May 9-10 Data Users Conference. Additional recommendations that related to questionnaire design and content issues included the need to develop questions and categories that are more engineering specific in nature. Information on both the formal (education) and informal (coursework and training) mechanisms by which individuals maintain their technical currency was requested. Data on factors related to the success and failure of women and minorities in the engineering labor force is very important and should be developed. During the workshop, an interest also emerged on expanding the taxonomies on highest degree field and occupation to provide more detailed occupational disaggregation.

## **CNSTAT Report**

The Committee on National Statistics of the National Academy of Sciences was given the responsibility to evaluate the current STPDS and to propose a system for the 1990s. In 1989 the Committee completed its work and issued a final report, Surveying the Nation's Scientists and Engineers, A Data System for the 1990s (CNSTAT Report). The various recommendations made in this report have influenced the questionnaire currently being developed for the surveys of the 1990s. These recommendations include improvements in the following proposed content areas: (1) Data on kinds of work performed by scientists and engineers, (2) Data on career paths, and (3) Data on career outcomes.

## **Research Users Group Meeting**

In June, 1989, NSF convened a meeting of researchers that use the STPDS in their research. The meeting was primarily oriented towards the research needs of the data users and the adequacy of the STPDS in filling those needs. The meeting provided information that was similar to data generated in the reports mentioned above.

## **5. Input to Redesigned Questionnaires**

### **Input of Mathematica Policy Research Corp.**

Mathematica's role was to assist NSF in developing and implementing procedures for the questionnaire redesign. The tasks completed by Mathematica for NSF included reviewing data on the existing survey instruments, advising on the number and composition of technical working groups, assisting with the selection and recruitment of individual members of these groups, and recommending to NSF, with technical working group input, specific recommendations on redesign activities. Mathematica was responsible for scheduling and arranging meetings of groups and in developing the questionnaire format and content in conjunction with NSF staff.

### **Input of technical working groups (Non-Federal)**

Two technical working groups (composed of a technical design group and a survey content group) convened to discuss the STPDS. Selection of members for these groups was carefully made to represent users who had a good working knowledge of the data such as researchers and policy analysts and

representatives from various organizational sectors such as academia, private industry, and trade associations.

### **Technical design working group**

The technical design working group had some general comments on questionnaire design that built upon the work in the CNSTAT Report, namely that the questionnaire is not oriented to non-S&E employed persons who have S&E training much less to persons who have no background in S&E. They suggested using a more inclusive survey title, including the name of a contact person and a telephone number in the cover letter, changing the flow of questions to reflect the stated purpose of the survey. The technical design working group was also concerned with occupational coding and the comparability of this data element with other national data sets and the potential for inflated estimates of persons in S&E occupations over time. They recommended two options for minimizing spurious reports for occupational change. One was to provide the respondent with his/her previous wave response before completing the present wave questionnaire. The second one was to assess change by asking the respondent whether a previous occupation at a certain point in time is the same or different from the respondent's current occupation.

### **Survey content working group**

Aside from specific comments on individual questions which space does not permit us to include here, the survey content group mentioned the following items as important issues that need to be addressed by the STPDS data system. The evaluation of factors contributing to relative attractiveness of S&E versus other professions. The role of junior colleges in the educational career of S&E personnel, especially minorities, was mentioned as an area of great interest. Other areas included the identification of significant points of reentry into the educational/training system and "substitutability issues" such as the identification of S&E personnel who could cross over between closely related fields to more immediately respond to changes in demand. The impact of technology on the mobility between subfields and the availability, qualifications, and utilization of technical (non Ph.D) level S&E personnel were also mentioned as topic areas that need further exploration.

## **6. How the Decisions on Questionnaire Content Were Made**

NSF survey content staff evaluated the various reports on questionnaire design and content including the Data Users Conference report, the Engineering Workshop report, the CNSTAT report, technical working group reports, survey content analyses on the 1980s surveys, questionnaires of the Decennial Census, the Current Population Survey, and the National Center for Education Statistics (NCES) Recent College Graduate Survey. A draft questionnaire was developed that was 45 pages long that included most areas of interest to our users while adhering to the goals of the redesign effort. Through a process of eliminating questions, the draft questionnaire was eventually reduced to 17 pages.

The questions that were dropped fell into several categories. One category included questions that were not easily formulated because content areas were not amenable to clear and unambiguous phrasing (e. g., questions on productivity or quality of employee output). Others were eliminated because they could not be collected from a survey of individuals (e.g. SIC classification of a respondent's employer). Questions that were considered too burdensome on the respondent were immediately dropped from our questionnaire (a request for data on professional income, by income source was considered too burdensome because respondents would have to refer to secondary sources for this data--income tax returns). There were some questions that we decided need not be asked in each survey cycle, some of these were postponed to future years. Examples of these included questions on the number of subscriptions to journals, number of patents held, methods used to maintain professional skills. Some

questions were considered too intrusive (e.g. social security numbers). Finally, some questions were already collected in the Decennial Census and since the Postcensal/experienced sample was being drawn from the Decennial Census the data could be obtained during the sample selection. These data included race, ethnicity, and sex.

## **7. Testing the questionnaires**

### **Cognitive laboratory**

In September, 1991, four participants were invited to individual cognitive lab sessions. Each laboratory session lasted between 1 and 1/2 to two hours with 30-40 minutes spent on completing the questionnaire. The purpose of these sessions was to identify and clarify issues regarding participation, incentives, and questionnaire design. The participants were asked some general questions about their participation in the survey, incentives as an inducement to participate, and general questions about the questionnaire.

### **Focus groups for the Postcensal questionnaire**

Four focus groups were conducted with each group representing a different mix of occupations. Group 1 primarily included scientists, engineers, and mathematicians. Group 2 included administrators, managers, teachers, and professionals in scientific disciplines. Group 3 were nonscientists such as architects, lawyers, and real estate agents. Group 4 were all social scientists. Each group was scheduled for 2 and 1/2 hours and was composed of between 9-11 individuals. After completion of the questionnaire the moderated discussion followed for about 1 and 1/2 hours. Techniques used by the moderator included, direct questions, directive probes, and retrospective “think aloud techniques” (techniques used to think back and determine what strategies they used for recalling information or formulating an answer).

### **Improvements to Postcensal questionnaire from cognitive laboratory and focus groups**

Major changes were made to the questionnaire as a result of analyzing the focus group and cognitive lab results. The cover letter was revised to have a clear statement of purpose. Many of the focus group participants felt that the questionnaire needed more purpose. Many focus group participants felt that it was important to show potential respondents what the questionnaire was used for so a copy of summary results will be offered to participants in the survey.

The focus group participants believed that the sequencing of questions on the National Survey of College Graduates should begin with questions on employment during the reference week and questions on education should be included in the background section and not prominently displayed as an independent section.

The work-related training questions were separated from the academic coursework questions; focus group participants felt that there was confusion when they were included in one section. The work-related training questions were placed after the questions on occupation and the academic coursework questions were placed after the questions on education. Questions on college education were included in one matrix that resulted in a short and less crowded questionnaire format.

The work related training activity question was simplified and reduced and the Dept. of Energy was listed as a sponsor in the cover letter to help in the transition to energy questions in the questionnaire.

The two coding lists were attached to the back of the questionnaire and the code lists were renamed from Reference Lists A and B to List A - Occupational Codes and List B - Field of Study Codes.

### **Focus groups for the New Entrants questionnaire**

Two focus groups have been held so far in 1992. Since the sample for the 1993 National Survey of Recent College Graduates will include bachelor's and master's graduates in the social and natural sciences and engineering from 1990, 1991, and 1992, the focus group reflected this mix of degrees, years, and majors. The participants included graduates from a broad spectrum of universities having a balanced mix of men and women, and a balanced mix of those who were working and those who were attending graduate school. The techniques used by the moderator included the same techniques used in the focus groups for the Postcensal questionnaire.

### **Improvements to the New Entrants questionnaire from focus groups**

Sequencing of questions on the National Survey of Recent College Graduates was maintained with education questions prominently displayed as the first section of the questionnaire. However, questions on past employment for the New Entrants survey were dropped because few of the recent graduates had extensive previous employment history. A personalized letter on National Science Foundation letterhead was mentioned by the focus group participants as a positive aspect of the questionnaire material.

Three other major improvements were made to the New Entrants questionnaire as a result of the focus group comments and discussions. One change resulted in more accurate response categories for sources of financial assistance for degrees.

The question on undergraduate grade point average (GPA) included in the draft questionnaire after the college degree information matrix (collecting detailed data on all college degrees) was moved to a location preceding the matrix because some participants wanted to report their graduate school GPA in that question, because it was higher than the undergraduate GPA.

Finally some focus group participants were uncomfortable about providing information on a job they had when they received their degree because it wasn't a career path job. A question was placed before the question on the job held during the period the graduate received their degree to determine whether the graduate held or was seeking a career path job.

### **Pretest of Postcensal questionnaire**

A pretest of the survey is currently underway of 3200 sample members, equally distributed by the following four versions of the questionnaire. A long form was sent to 800 sample members. The same long form was sent with a \$5.00 check as an incentive to complete the questionnaire to 800 individuals. Finally a screener form was sent to 800 individuals with follow-up forms sent to those identified as scientists and engineers in the screener. A mid-length form was also sent to 800 individuals. In addition to information on response rates the pretest will provide more insight on infrequently used response categories, and frequent write-in responses.

## **8. Selected features of the proposed redesigned questionnaires**

The development of the questionnaires is an iterative process that has not ended. The results of the postcensal pretest and additional focus groups for all surveys will undoubtedly provide information that will affect the final questionnaires. The questionnaires that have emerged at this point in time have some features in common. The proposed formatting features of these questionnaires are the same, with standard size booklets with no respondent letter and instructions or survey letter on the front cover (a cover letter is proposed for these surveys). A double column format with response categories below the questions and a vertical flow of questions from top of page to bottom of the page is used in all surveys. Finally, as has

already been discussed, the content areas of each survey have been enriched and where feasible made uniform among the individual surveys of the STPDS.

## BIBLIOGRAPHY

Citro, Constance F. and Kalton, Graham, Surveying the Nation's Scientists and Engineers: A Data System for the 1990s, Panel to Study the NSF Scientific and Technical Personnel Data System, Committee on National Statistics, National Academy Press, Washington, D.C., 1989

Eisenhower, Donna, Report on the Cognitive Laboratories and Focus Groups for the 1993 Postcensal Questionnaire, Mathematica Policy Research, Inc., January, 1992

Engineering Personnel Data Needs for the 1990s, Committee on Data Needs for Monitoring Labor Market Conditions for Engineers, National Research Council, Washington, D.C., 1988

Mathematica Policy Research, Inc. Unpublished Notes from the January 14-15 Meeting of the Technical Design Group

Mathematica Policy Research, Inc. Unpublished Notes from the December 10, 1990 Meeting of the Survey Content Group

Mooney, Geraldine, A Review of the STPDS Survey Questions of the 1980s, Task 3.1 Report, Mathematica Policy Research, Inc., October 5, 1990

Mooney, Geraldine and De Saw, Cheryl, 1993 New Entrants Questionnaire Spring Focus Groups Summary, Mathematica Policy Research, Inc

Moshman Associates, Inc., Proceedings of the NSF Conference on the Scientific and Technical Personnel Data System: Strategy and Design, October 16-17, 1979, Washington, D.C.

Wallman, Katherine, Scientific and Technical Personnel in the 1990s: An Examination of Issues and Information Needs, Proceedings of the May 9-10, 1985 Conference, Council of Professional Associations on Federal Statistics, Washington, D.C.



## Appendix A. Sequencing of Major Topic Areas Among STPDS Questionnaires in the 1980s

<b>1982 Postcensal/Experienced Sample Survey</b>	<b>1982 New Entrants Survey</b>	<b>1981 Survey of Doctoral Recipients</b>
1. Some Facts About Yourself	I. Demographic Characteristics	I. Employment Profile
II. Education and Training	II. Education and Training	II. Demographic Characteristics
III. Employment Status	III. Employment Status	
IV. Employment Profile	IV. Employment Profile	
V. Other Information		
<b>1984 Postcensal/Experienced Sample Survey</b>	<b>1984 New Entrants Survey</b>	<b>1983 Survey of Doctoral Recipients</b>
1. Education and Training	I. Demographic Characteristics	I. Demographic Characteristics
II. Employment Status	II. Education	II. Employment Profile
III. Employment Profile	III. Employment Status	
IV. Other Information	IV. Employment Profile	
<b>1986 Postcensal/Experienced Sample Survey</b>	<b>1986 New Entrants Survey</b>	<b>1985 Survey of Doctorate Recipients</b>
I. Education and Training	I. Demographic Characteristics	I. Demographic Characteristics
II. Demographic Characteristics	II. Education	II. Employment Profile
III. Employment Status	III. Employment Status	
IV. Employment Profile	IV. Employment Profile	
V. Other Information		
<b>1989 Postcensal/Experienced Sample</b>	<b>1988 New Entrants Survey</b>	<b>1987 and 1989 Survey of Doctoral Recipients</b>
I. Education and Training	I. Education	I. Employment Profile
II. Employment Status	II. Employment Status	II. Demographic Characteristics
III. Employment Profile	III. Employment Profile	
IV. Demographic Characteristics	IV. Demographic Characteristics	
V. Other Information	V. Other Information	

## Appendix B. Formatting Features of Selected Questionnaires of the STPDS Surveys of the 1980s

Feature	1989 Postcensal/Experienced Sample Survey	1988 New Entrants Survey	1988 Survey of Doctorate Recipients
<b>Overall Format</b>	Booklet	Booklet	Booklet
<b>Size</b>	Standard - 8.5" x 11"	Undersized 6 1/8" x 8 1/4"	Standard - 8.5" x 11"
<b>Paper and Print Color</b>	White with dark blue ink	White with black ink	White with brown ink
<b>Front Cover</b>	Respondent letter and instructions, No survey questions	No respondent letter or instructions, No survey questions	No respondent letter or instructions, Survey questions
<b>Page Format</b>	Double column with response categories under questions  Vertical flow of questions from top of page to bottom of page	Varies between double column and no columns with response categories under questions  Vertical flow of questions from top of page to bottom of page	Double column with response categories under questions  Horizontal flow of questions across the page from left to right

## Appendix C. Proposed Features of the STPDS Surveys of the 1990s

Feature	1993 National Survey of College Graduates (formerly Postcensal/Experienced Sample Survey)	1993 National Survey of Recent College Graduates (formerly New Entrants Survey)
<b>Overall Format</b>	Booklet	Booklet
<b>Size</b>	Standard - 8.5" x 11"	Standard - 8.5" x 11"
<b>Paper and Print Color</b>	White paper with teal background shading with black and blue ink	Not yet determined
<b>Front Cover</b>	No respondent letter or instructions, No survey questions	No respondent letter or instructions, No survey questions
<b>Page Format</b>	Mostly double column with response categories under questions  Vertical flow of questions from top of page to bottom of page	Mostly double column with response categories under questions  Vertical flow of questions from top of page to bottom of page
<b>Sequencing of Major Topic Areas</b>	I. Employment Status during Reference Week (Includes Employment Profile and Employment Status)  II. Past Employment  III. Other Work Related Information  IV. Background Information (Includes Education and demographic Characteristics)	I. Education  II. Employment Status during Reference Week (Includes Employment Status and Employment Profile)  III. Other Work Related Information  IV. Background Information (Includes Demographic Characteristics)