

Request for OMB Clearance
Microbusiness Innovation Science and Technology (MIST)
MIST Additional Testing and Debriefings
Supporting Statement for Paperwork Reduction Act Submissions

Section A. Justification

The National Center for Science and Engineering Statistics (NCSES) of the National Science Foundation is broadly tasked with measuring the role of science and technology (S&T) in the United States' economy and abroad. A major component of this activity is its sponsorship of the Business Research and Development (R&D) Innovation Survey (BRDIS), which collects information annually on research and development and related activities performed by U.S. firms. However, businesses with fewer than five employees are excluded from this survey.

In 2004, the National Academy of Sciences' Committee on National Statistics (CNSTAT) reviewed NSF's portfolio of R&D surveys and recommended that NSF explore ways to measure firm innovation and investigate the incidence of R&D activities in growing sectors, such as small business enterprises, not currently covered by BRDIS. As a result, NSF has decided to undertake a survey of R&D and other innovation-related items among very small (i.e., micro), independent U.S. businesses with fewer than five employees. In addition to general business information - primary business activity (NAICS code), year business was formed, and number of employees - this survey proposes to collect data on R&D, innovation, employment, related activities (such as sales of significantly improved goods and services; operating agreements and licensing activities; technology transfer; patents and intellectual property; and sources of technical knowledge), measures of entrepreneurial strategies, and demographic characteristics of the entrepreneur.

Significant work has been completed towards the development of the questionnaire. In the fall of 2012 a pretest was conducted with microbusinesses. While the content of the questionnaire worked well with respondents, the response rate was far below expectations. Before proceeding with a full data collection it is requested that OMB give permission for additional testing and follow up debriefing interviews.

This clearance request is for two projects to assist in the design of the Microbusiness Innovation Science and Technology Survey (MIST): additional testing including methodological experiments with 4,000 microbusinesses and debriefings of 50 respondents and non-respondents. The debriefing interviews will be conducted at the conclusion of the testing to assist in its evaluation.

A.1. Need for Data Collection and Legislative Authorization

Approval to conduct the MIST Survey is being requested under The America COMPETES Reauthorization Act of 2010 § 505, codified in the National Science Foundation Act of 1950, as amended, at 42 U.S.C. § 1862 that under paragraph b directs the Foundation through NCSES to:

- (1) collect, acquire, analyze, report, and disseminate statistical data related to the science and engineering enterprise in the United States and other nations that is relevant and useful to practitioners, researchers, policymakers, and the public, including statistical data on--
 - (A) Research and development trends;
 - (B) The science and engineering workforce;
 - (C) United States competitiveness in science, engineering, technology, and research and development;

The MIST Survey will provide essential data on the research and development conducted by the smallest businesses in our country. The survey will allow us to have a better understanding of the innovative activities conducted by these companies.

To date, MIST has performed considerable development work to determine the relevant research questions and design the MIST questionnaire. This work includes cognitive interviews, consultation with an expert panel specially created for MIST, a pretest of 74 businesses and debriefing interviews with those respondents. Before conducting the full data collection, additional testing is required.

A.2. How, by Whom, and for What Purpose the Information Is to Be Used

Policy officials from many Federal agencies will rely on these statistics for essential information. For example, total U.S. R&D expenditures statistics have been used by the Bureau of Economic Analysis (BEA) to establish an R&D satellite account to the System of National Accounts (SNA). As of July 2013, R&D is incorporated as a core component in the SNA. Accurate and complete R&D data are needed for the SNA. With the addition of the microbusiness R&D data, the accounts will provide a more complete picture.

Government policy officials who are involved in assessing the role of the Federal government in promoting economic growth use R&D and related statistics in their decision-making processes that affect technological and economic progress. NSF staff also work closely with the Office of Science and Technology Policy (OSTP), providing R&D and related statistics and indications of emerging trends to

assist the OSTP staff in their analyses of the status of science and technology in the United States. In addition, NSF has frequent contact with the Office of Management and Budget (OMB), the Congressional Budget Office (CBO), and the Congressional Research Service (CRS) which use R&D statistics in their studies.

A.3. Use of Automated, Electronic, Mechanical or Other Technological Collection Techniques

During the debriefing interviews respondents reported they would prefer a web survey (one was not available during the pretest). A web survey has been developed and will be used during this testing. A paper questionnaire will still be available for those who prefer that mode. The web instrument will simplify certain aspects of the survey process for the respondents, such as automatically skipping questions that do not apply based on previous answers, performing arithmetical operations, and including edit checks that will help to limit certain types of respondent errors (such as providing percentages that do not add to 100).

An electronic survey management system will be used to manage the data collection process, including tracking all responses, making assignments to interviewers, managing all modes of contacts with the sampled businesses, and storing paradata that will be used to evaluate the effectiveness of various aspects of the data collection process. Computer edits will be performed to verify the completeness and consistency of the data that are collected. For example, computer edits include whether the detailed responses sum to the totals, whether skip patterns have been followed correctly, whether some values fall outside of expected ranges, and whether some responses might be logically inconsistent. The web version of the survey has a real-time monitoring system, which allows NSF to monitor data, response status, and comments from respondents. From the perspective of the respondents, the web version is more convenient and simplifies the survey (e.g., by automatically calculating totals). NSF benefits from the use of the web version by receiving improved data quality.

A.4. Efforts to Avoid Duplication

The data collected through MIST are new data not collected elsewhere and not available through administrative data. NSF does collect comparable information on larger companies through BRDIS, but this is the first survey of microbusinesses. While data on the number of employees are included on IRS Form 941, and are used for sampling purposes, the data collected by the survey questionnaire on the number of employees ask for updated data not yet available from the IRS. In addition, businesses are asked to enumerate all who work for the company, including volunteers and subcontractors who are not counted on Form 941.

A.5. Collection of Data from Small Businesses

As a survey of microbusinesses, the survey will directly have an impact on small businesses, but the questionnaire has intentionally been kept short and designed to be easy to complete. In addition, the survey will be offered over the web, which participants in the debriefing interviews indicated was their preferred mode. It will not be necessary for respondents to consult their financial records when answering questions. This was confirmed in the debriefing discussions following the pretest.

Although initial drafts of the survey content contained questions that asked for actual dollar figures, many of these questions now request estimated percentages, which are easier to provide. These changes will make it easier for a respondent to complete the survey in one continuous session, without the need for multiple sessions or through an initial data-gathering stage followed by a web-reporting stage. In addition, the questions that ask for dollar amounts now ask for the answers to be rounded to the nearest thousand, which should lessen the need to consult records.

A.6. Consequences of Less Frequent Data Collection

This is a new request and future data collections have not yet been determined. This request is for clearance of additional testing and debriefings only. A separate request will be submitted for any future MIST collections.

A.7. Special Circumstances

This information collection will be conducted in a manner consistent with OMB guidelines and there are no special circumstances.

A.8. Consultation Outside NSF

In September 2010, NCSES held a one-day workshop for prospective data users of microbusiness, R&D and innovation-related data in order to help gain a better understanding on data user needs, to understand how microbusiness data will be used, and to obtain guidance on priorities and strategies for content to be included in the survey. The potential data users included academics, government employees and other stakeholders. The findings from the data user workshop were incorporated into the questionnaire.

In February 2012, NCSES convened an expert panel of academics and other stakeholders to provide input and feedback to NCSES on the development of the MIST survey. Panel members reacted to the draft survey questionnaire available at the time and discussed issues encountered in the earlier testing, such as questions on innovation and defining the number of employees.

In late 2011 and early 2012, the survey questionnaire was tested with 48 businesses using cognitive interviews. After the cognitive interviews, a pretest was conducted with 74 microbusinesses followed by debriefing interviews with 20 of those respondents. Following the debriefings an evaluation of all data collection procedures was conducted. During that evaluation, it was determined that additional testing was necessary before conducting a full data collection. The questionnaire required only minimal changes. However, the response rate was far below expectations.¹ As a result of the pretest and the debriefing interviews, the following changes are being incorporated as part of this additional testing.

- The sample for the microbusiness survey is based on data from the IRS Compliance Data Warehouse (CDW). As part of the sample selection, a business income tax return is matched to its corresponding quarterly payroll tax return. For the pretest the two returns (income and payroll tax) were from the same tax year. However, the quarterly payroll tax return was in many cases more than two years old. It is believed that this contributed to a larger than anticipated percentage of out of business respondents.² As a result, a more recent quarterly payroll tax return will be used for this testing. At the time of the pretest it was thought that it was necessary to ensure that the two returns were from the same tax year. However, since more recent data is available it will be used.
- For the pretest, due to IRS restrictions on outside vendors the sample data were not compared to the National Change of Address database. It was also not possible to append phone numbers or other information. However, with IRS cooperation, the sample will be sent to a third party

¹ During the pretest a sample of over 300 respondents resulted in 74 completed questionnaires.

² Following the pretest, an analysis of out of business microbusinesses was conducted. The proportion of microbusiness that were no longer in business varies dramatically by form type (typically between 20% and 30%).

vendor for updating addresses and obtaining additional information including telephone numbers for potential follow up calls.

- During the debriefing interviews respondents stated that the preferred method for completing the survey was the web. A web survey has been added and will be used during this testing.
- Respondents also indicated during the debriefings that the NSF name helped and that the subject matter was of particular interest. As a result, the NSF logo will be prominently displayed on communication to the respondent, the questionnaire and the web survey. A brochure has also been created that emphasizes the NSF name and the importance of the survey.
- Unlike in the pretest, for this testing, experiments will be used, including an incentive and one page screening questionnaire to see what methodology is more effective. Additional information on these experiments is provided in section B.4.
- The initial questionnaire testing included very limited follow up. That will not be the case with this additional testing. The data collection procedures to be followed will allow for a close examination on the number of telephone calls made as well as other contacts. Following the data collection testing a full evaluation including a comprehensive non response analysis will be conducted.

On April 9, 2013, the National Science Foundation published a notice in the Federal Register (78 FR 21160 - <https://federalregister.gov/a/2013-08177>) inviting the general public and other Federal agencies to comment on plans to submit this request. No comments were received.

NSF plans to continue to obtain outside advice on the content, design, and conduct of the survey from potential respondents, data users, academics, government employees as well as other stakeholders.

A.9. Payments to Respondents

Achieving a high response rate will be a challenge for the MIST survey. It is well known that response rates for government surveys have been declining. Studies have shown that incentives are an effective way to boost response rates (Dillman et al., 2009). However, the literature on incentives is mostly on incentives for household studies. The literature on incentives for establishment surveys tends to focus on larger establishments. While Luo and White (2005) note three reasons why incentives for establishment surveys do not translate as easily as they do for individual surveys, these reasons do not appear applicable

to microbusinesses because of their small size. The survey will be directed towards the business owner, making the following three concerns appear less relevant.

- Confusion over the appropriate recipient of incentive: with large businesses, it is sometimes difficult to know who the proper respondent is, but with microbusinesses that will not be difficult.
- Lack of social exchange: the person who completes the survey in a large organization may be completing the survey only because their supervisor directed them to do so or, with a prepaid incentive, the person who receives the incentive may not be the person who actually completes the survey; with a microbusiness this is unlikely to happen because the number of employees is so few.
- Organization may forbid accepting incentive: some businesses may prohibit their employees from accepting money from any outside sources. However, microbusinesses are so small it is unlikely that they would have such a policy in place.

NSF proposes a methodological experiment to test the effectiveness of using incentives to increase responses among microbusinesses. Each sample unit will fall into one of three conditions: a \$2 prepaid incentive, a \$20 incentive upon completion, or no incentive at all (control). By testing these different types of incentives, which incentive levels appear to be most effective and whether they would be justified for the full survey can be measured. A longer discussion of the experiments is provided in section B.4.

A.10. Assurance of Confidentiality

The sampled businesses will be provided the following assurance concerning confidentiality to be located on the cover of the paper questionnaire and the introductory screen (after sign-in) of the web version:

Your responses are confidential under law. See the statement at the bottom of this page for more details.

This information is solicited under the authority of the National Science Foundation Act of 1950, as amended. All information you provide is protected under the Confidential Information Protection and Statistical Efficiency Act of 2002; it will be used only for statistical purposes. Any information publicly released (such as statistical summaries) will be in a form that does not personally identify you or your company. Your response is voluntary and failure to provide some or all of the requested information will not in any way adversely affect you or your company.

The Confidential Information Protection and Statistical Efficiency Act of 2002 states “Data or information acquired by an agency under a pledge of confidentiality and for exclusively statistical purposes shall be used by officers, employees, or agents of the agency exclusively for statistical

purposes" and "Data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes shall not be disclosed by an agency in identifiable form, for any use other than an exclusively statistical purpose, except with the informed consent of the respondent."

All procedures in IRS Publication 1075 to protect the confidentiality of tax data received from the IRS will be followed. In addition, the survey contractor understands and applies to all projects the security standards and requirements established by the Federal Information Security Management Act (FISMA) and related regulations and guidance documents, treating this as a survey presenting a moderate level of risk. The survey contractor will work closely with NSF and the IRS to ensure compliance with their established security policies and procedures and to meet specified goals for confidentiality, integrity, and availability of information and information systems.

A.11. Sensitive Questions

MIST is a voluntary survey, and no persons are required to respond to the questionnaire. In addition, respondents may decline to answer any questions in the survey. This voluntary aspect of the survey is clearly stated in the introduction. During the debriefing interviews, no respondents said that they found the questions to be sensitive.

A.12. Estimates of Response Burden

Based on the cognitive interviews and pretest, it is estimated that the survey will require approximately 30 minutes (on average) to review the materials and complete the questionnaire. This is based on answers respondents provided during the pretest debriefing interviews. There are two experiments to be included as part of this testing (both are described in depth in section B.4). The first experiment will involve dividing the sample equally and administering a one-page questionnaire to 2,000 potential respondents followed by the rest of the questionnaire in a separate step. The remaining 2,000 respondents will be asked to complete the full questionnaire. The full questionnaire can be found in Attachment A and the one page screening questionnaire is in Attachment B. All respondents will receive the survey brochure (Attachment C) and introduction letter (Attachment D).

Regardless of which group the respondent is in the time necessary to complete the questionnaire is the same. With a goal of an 80% response rate, an estimate of 3,200 respondents will complete the questionnaire for a total of 1,600 burden hours (3,200 x 30 minutes).

In addition, debriefings of 50 respondents will be conducted, with the debriefings lasting no more than 60 minutes per respondent, for a total of 50 burden hours. The combined burden for the survey and debriefings is estimated at 1,650 hours (1,600 hours for the questionnaire and 50 hours for the debriefing interviews.).

At an estimated cost of \$68.76 per hour (based on the Bureau of Labor Statistics May 2011 average hourly wages for “Top Executives” within NAICS Sector 54 - Professional, Scientific, and Technical Services and NAICS Sectors 31, 32, and 33 - Manufacturing), the total annual cost to respondents is \$110,016 (or approximately \$34.38 per respondent for 30 minutes).

A.13. Estimate for the Total Annual Cost Burden to Respondents or Recordkeepers

Non-applicable. Respondents are not expected to incur any capital and startup costs or systems maintenance cost in responding.

A.14. Annualized Estimates of Cost to the Federal Government

The total contractual cost of the additional testing and debriefings to the government is approximately \$1.59 million over a period of 24 months, for an annualized cost of \$796,653. Including labor costs for NSF staff, approximately \$250,000 (project manager (full time), mathematical statistician (part time), program director (part time) and other staff) the total annualized cost to the federal government for the additional testing and debriefings is \$1.04 million.

A.15. Program Changes

Not applicable. The MIST survey has not been conducted previously.

A.16. Plans for Tabulation and Publication

This testing is being conducted to work on establishing the procedures for data collection such as developing contact strategies and testing different procedures that should enhance the response rate. There are no plans to publish survey data from this testing except from a methodological perspective. Following the additional testing and the debriefings an evaluation will be conducted and a report produced from that evaluation. Depending on the evaluation phase, NCSES may pursue a full scale survey to take place in 2015.

A.17. Approval to Not Display Expiration Date

Not applicable. The survey will display the expiration date for OMB approval of the information collection.

A.18. Exceptions to the Certification Statement

Not applicable. No exceptions to the certification statement are being sought.