O*NET

**U.S. Department
of Labor**

**Employment
and Training
Administration**

O*NET® Data Collection Program

**Office of Management and Budget
Clearance Package Supporting Statement**

December 10, 2008

Table of Contents

Section **Page**

Appendices

    A.   Questionnaires
    B.   Advisory Panel for the *Dictionary of Occupational Titles:* Final Report
    C.   Citations from the Workforce Investment Act, the Carl D. Perkins Career and
        Technical Education Act, and the *Code of Federal Regulations*
    D.   *Federal Register* Notice
    E.   Publications Referencing the O*NET Data Collection Program
    F.   Mailing Materials
    G.   Paper on Model-Aided Sampling for the O*NET Data Collection Program
    H.   Nonresponse Analysis for Analysis Cycles 4 Through 8
    I.   The Generalized Exponential Model for Sampling Weight Calibration for
        Extreme Values, Nonresponse, and Poststratification
    J.   The Employment and Training Administration's Responses to the U.S. Bureau of
        Labor Statistics' Questions About the Draft O*NET OMB Clearance Package

List of Exhibits

Number | **Page**

A. Justification

A.1 Circumstances of Information Collection

This Supporting Statement is a request and justification for a 3-year clearance from the Office of Management and Budget (OMB) to continue the Occupational Information Network (O*NET®) Data Collection Program. The O*NET Data Collection Program continually operates to populate and maintain a current database on the detailed characteristics of workers, occupations, and skills. The program uses an occupational taxonomy, the O*NET-SOC, which is based on the Standard Occupational Classification (SOC) mandated by OMB for use by all federal agencies collecting occupational and labor market information (2000). The O*NET-SOC comprises occupations at the most detailed level of the SOC and includes additional occupational specificity as needed. In addition, new and emerging occupations in high-growth sectors of the economy are being identified and added to the taxonomy. So far, data have been published for 780 O*NET-SOC occupations. This request is to continue to collect new data on selected high-growth occupations, together with new and emerging O*NET-SOC occupations, over the next 3 years (calendar years 2009 to 2011), subject to annual budget levels.

The continued population of the O*NET database is important because the O*NET database is the most current and comprehensive standard source of occupational information in the United States. The O*NET Data Collection Program remains at the center of an extensive network of occupational and skill information used by a wide range of audiences, including individuals making career decisions; public agencies such as workforce investment boards (WIBs) and One-Stop Career Centers making training investment decisions; educational institutions preparing a future workforce; and employers making staffing and training decisions.

This program provides a common language and framework to facilitate communication about industry skill needs among business, education, and the workforce investment system. The resulting O*NET database is used to develop industry competency models and occupational competency profiles. O*NET data include information about transferable skills and are used for skills gap analysis, facilitating a mobile workforce responsive to changing economic needs. The O*NET database and companion O*NET Career Exploration Tools are used by many private companies and public organizations to tailor applications to their needs and those of their

customers. The broad utility of the O*NET tools contributes to U.S. competitiveness in a global, 21st-century economy.[1]

The O*NET Data Collection Program employs a multiple-method approach to updating the O*NET database. The primary method involves a two-stage sample design to survey establishments and workers in those establishments. When necessary, this method may be supplemented with a sample selected from additional sources, such as professional and trade association membership lists, resulting in a dual-frame approach. An alternative method, based on sampling from lists of identified occupation experts, is used for occupations for which the primary method is inefficient. This method is reserved for selected occupations, such as those with small employment scattered among many industries, and those for which no employment data currently exist on which to base a sample, such as new and emerging occupations. With all methods, the O*NET survey instruments are used.[2] The rest of Section A describes the O*NET Program and reviews statutory and regulatory information.

A.1.1 What Is the O*NET Program?

The O*NET Program is a comprehensive system for collecting and disseminating information on occupational and worker requirements. The O*NET data supersede the U.S. Department of Labor's (DOL's) *Dictionary of Occupational Titles* (DOT) and provide additional occupational requirements not available in the DOT. The DOT is no longer updated or maintained by DOL.

As shown in Exhibit 1, the O*NET Program uses a data structure, the Content Model, to organize occupational information and to provide a common language of standardized and defined occupation descriptors and measures for use by all audiences. The O*NET Content Model is the result of extensive research, and its development is fully documented (Peterson, Mumford, Borman, Jeanneret, & Fleishman, 1995, pp. 2-6; Peterson, Mumford, Borman et al., 1997; Peterson et al., 2001). It comprises worker-oriented and job-oriented characteristics at both an occupation-specific level and across occupations, as the exhibit illustrates.

---

[1] For details on the uses of the O*NET Program, see Section A.2. More information about the O*NET Data Collection Program can be found at the National O*NET Program's public Web site, http://www.onetcenter.org/ (the O*NET portal page that links to several O*NET-related Web sites) and at http://www.doleta.gov/programs/ onet (U.S. Department of Labor, Employment and Training Administration, Web site).

[2] For detailed information on these methods, see Sections B.1 and B.2.

Exhibit 1.   O*NET Content Model
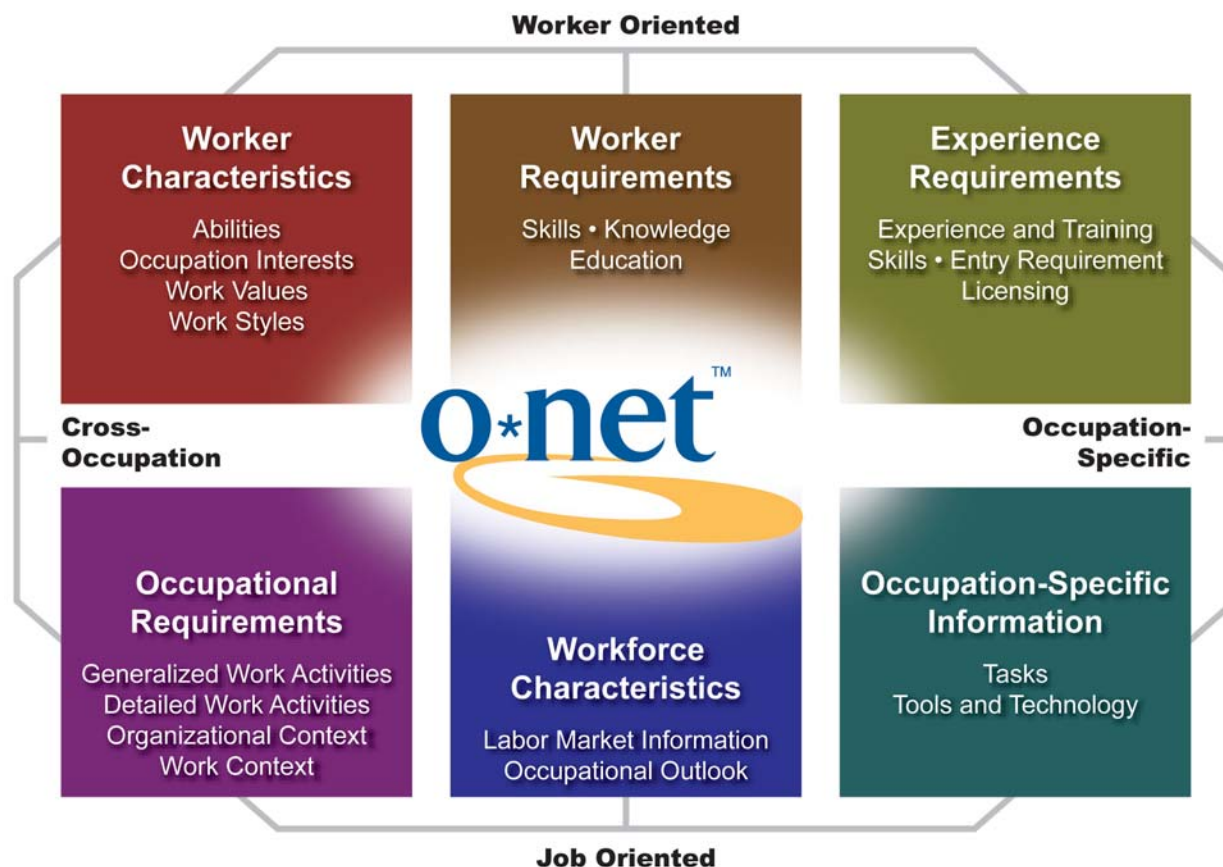
Each of the six domains of the Content Model groups information hierarchically. For example, the Worker Characteristics domain contains four types of information: Abilities, Occupational Interests, Work Values, and Work Styles. From these four, the Abilities domain, in turn, contains four types of abilities: Cognitive, Psychomotor, Physical, and Sensory. Each of these types of abilities contains further levels of detail. For example, the Psychomotor type includes Fine Manipulative, Control Movement, and Reaction Time and Speed. Finally, Fine Manipulative contains three specific descriptors: Arm-Hand Steadiness, Manual Dexterity, and Finger Dexterity. Hierarchies both organize the information and allow for different levels of specificity. By organizing occupational information hierarchically, the O*NET taxonomies of information and SOC-based occupations facilitate the use of a common language to describe the world of work.

The O*NET Program takes the best knowledge about both content and methodology from the last 60 years of research, since the first DOT. The descriptors and rating scales for O*NET data were developed through extensive research, drawing primarily from job analysis in industrial/organizational psychology and human resource management (Peterson et al., 1995).

The descriptors in the O*NET Program are meant to be exhaustive. The primary sources of data are job incumbents and occupation experts. The Standard Occupational Classification is used as the basis for classifying occupations. The use of questionnaires and rating scales reflects the most currently accepted approach to job analysis. The scales used for the O*NET ratings are Importance, Level, and Frequency. Each descriptor in the O*NET questionnaires may use one or more scales. For example, the O*NET Skills descriptor Coordination is rated on both a 5-point Importance scale and a 7-point Level scale. For the complete set of O*NET questionnaires, which include O*NET descriptors, see Appendix A.

Exhibit 2 summarizes the number of descriptors and scales in the O*NET Data Collection Program questionnaires. Descriptors are identified from O*NET Content Model domains. Data are collected by means of 239 descriptors that include 400 scales (e.g., Importance, Level, Frequency). Currently, to collect ratings for the Abilities and Skills domains, trained occupational analysts review updated information (e.g., Tasks, Generalized Work Activities) provided by job incumbents.[3] No data collection is planned for the Workforce Characteristics domain. Information for it is provided through links to the employment, wage, and long-term projections databases produced by the U.S. Bureau of Labor Statistics (BLS), the state employment security agencies, and other agencies.

Exhibit 2.   O*NET Data Collection Program Questionnaires

O*NET Data Collection Program Questionnaire	Number of Descriptors	Number of Scales per Descriptor	Total Number of Scales	Data Source
Skills	35	2	70	Analysts
Knowledge	33	2	66	Job incumbents
Work Styles[a]	16	1	16	Job incumbents
Education and Training[a]	5	1	5	Job incumbents
Generalized Work Activities	41	2	82	Job incumbents
Work Context	57	1	57	Job incumbents
Abilities	52	2	104	Analysts
Tasks[b]	Varies	2	Varies	Job incumbents
Total (not including Tasks)	239	NA	400	NA

Notes: Occupation experts use the same questionnaires as job incumbents for those occupations whose data collection is by the Occupation Expert Method. NA = not applicable.

[a] The Knowledge Questionnaire packet also contains the Work Styles Questionnaire and the Education and Training Questionnaire.
[b] All job incumbents are asked to complete a Task Questionnaire in addition to the domain questionnaire.

---

[3] For a discussion of the preferred data source, see Section A.1.2.

Versions of the O*NET Database

The first version of the O*NET database released to the public was O*NET 98. O*NET 98 moved from the DOT's more than 12,000 occupations to a more user-friendly 1,122 occupations based on the BLS Occupational Employment Statistics (OES) codes. The O*NET 98 database contained 306 descriptors and 684 scales. A review of O*NET 98–specific scales and descriptors during the preparation for pretest data collection led to some consolidation and refinement of descriptors and scales to reduce burden on the public and increase employee response rate.[4]

The O*NET 98 database was first replaced with the O*NET 3.1 database and has been updated eight times as new data have been collected and analyzed. The current database, O*NET version 12.0, contains the same descriptors used in O*NET 98; however, the occupations have been restructured and coded to encompass the most detailed level of the 1998 SOC, with more occupational specificity added as needed. Research is under way to identify additional new and emerging occupations in high-growth industries. New occupations emerge because of changes in technology, society, law, business practices, and markets. As these new and emerging occupations are identified and data are collected, they will be integrated into the O*NET-SOC occupation classification and database.

O*NET 12.0 has a Web-based accessing application, O*NET OnLine, which is available to the public at no cost at http://online.onetcenter.org/. An electronic version of the 12.0 database can be downloaded at http://www.onetcenter.org/. The O*NET 12.0 database has been restructured to incorporate improvements made to the O*NET data collection instruments and is the structure currently being offered to developers.

Data in the O*NET database include the mean ratings on each of the items (or descriptors) in the O*NET questionnaires. Ratings have been standardized to facilitate interpretation and comparison. In addition to mean rating data on level and importance for various questionnaire items, text information on occupational definitions, descriptor definitions, scale anchors, and task descriptions are also included.

The Foundation for the O*NET Program

The DOT, first published by DOL in 1939, provided the occupational classification and descriptions required for operation of the public employment service. Information for the DOT was collected primarily through observation and interview by trained occupational analysts. These collection techniques were applied to a small number of job incumbents from what can best be characterized as a "convenience" sample. The DOT has not been maintained or updated since 1991; many of its observations were conducted during the 1970s.

---

[4] See *Revision of O*NET Data Collection Instruments*, available at http://www.onetcenter.org/resData.html.

In the late 1980s the Employment and Training Administration of DOL began a review of the DOT program. The purpose was to address concerns about the cost and difficulty of maintaining the DOT with existing methods and to address the need for additional and more current information. In 1990, as part of the DOT review, the Secretary of Labor appointed the Advisory Panel for the *Dictionary of Occupational Titles.* In response to its charge to advise on improvements to the DOT, this panel called for development of a new database of occupational information, a concept that evolved into the O*NET database. An important theme throughout the panel's recommendations was the development of a common language about jobs, occupations, and skills. The panel's complete report is presented in Appendix B.

A.1.2 The O*NET Data Collection Approach

The O*NET Data Collection Program is key to the continued effort to update the O*NET database to reflect the most current occupational skills and attributes, with special emphasis now on high-growth sectors of the economy and new and emerging occupations. In the research leading to the O*NET Data Collection Program, various sources and methods for collecting occupational information were examined, including collection of data from job incumbents and supervisors and development of ratings by occupation experts and occupational analysts. On the basis of this work, the O*NET team determined that the preferred source of data for most domains (Generalized Work Activities; Work Context; Knowledge; Education and Training; and Work Styles) is job incumbents.

By contrast, occupational analysts, who are provided with updated information from job incumbents, are preferred for the Abilities domain, which tends to be more abstract than the other domains. The Skills domain, whose variables are somewhat abstract, is a strong candidate for either source of collection; it is now updated by occupational analysts. A study conducted in 2006 found no clear evidence that one source of raters provides more valid or "accurate" data than the other for the Skills domain (Tsacoumis & Van Iddekinge, 2006).[5] Consequently, considerations of relative practicality, such as cost, informed the decision to proceed with analyst ratings of both abilities and skills. Other occupation experts, such as supervisors and trainers, may be used where access to job incumbents proves difficult or where the sampling of business establishments is inefficient.

Previous studies that compare various sources of job analysis ratings suggest that incumbents are able to provide information across a variety of descriptor domains (Fleishman & Mumford, 1988; Peterson, Owens-Kurtz, Hoffman, Arabian, & Whetzel, 1990). In addition,

---

[5] The goal of the study was to compare the psychometric quality of incumbent skills ratings with that of analyst skills ratings across a large sample of O*NET-SOCs. Although some mean differences between incumbent and analyst ratings were observed, the results yielded only minimal differences between the two systems of obtaining skills information, supporting previous research.

"large samples of knowledgeable job incumbents are available, which should contribute to the reliability of the resulting descriptive system" (Peterson, Mumford, Levin, Green, & Waksberg, 1999, pp. 2-6). Furthermore, the world of work is constantly changing, and technological advancements are occurring so rapidly that an efficient, effective way to remain current and accurate is to obtain the information directly from those performing the work.

As part of a random sample, workers selected to participate in the O*NET Data Collection Program are asked to rate the requirements of their own jobs as defined by the O*NET questionnaire items. The responses are tabulated into statistics, such as mean ratings for each scale. Collecting representative information from job incumbents presents many challenges; among them is determining the best method for identifying a representative sample of job incumbents in each occupation. The Advisory Panel for the *Dictionary of Occupational Titles* had recommended using sampling techniques that would ensure the representativeness and accuracy of the occupational data. Sampling allows an estimate of the population. By contrast, no systematic sampling methods were used in the development of the original DOT. The O*NET Program is concerned, in particular, with identifying sampling approaches that minimize burden on employers and the public, achieving broad coverage of the workers in each occupation, ensuring acceptable response rates, and supporting overall cost-efficiency.

Each sampling option has been evaluated in light of these criteria; each offers advantages and disadvantages. Three types of sampling frames are available for identifying samples of workers in each occupation: (1) lists of individual workers identified through professional and trade associations, licensing agencies, and unions; (2) households; and (3) employer establishments.

Identifying sampling frames of workers through professional and trade associations and unions retains the advantage of lower response burden because contacts with a sample of employers are replaced with contact with one or a few associations. Although it adds the cost of soliciting and maintaining association cooperation, this procedure also removes the cost of soliciting and maintaining employer cooperation. However, a major limitation of using special sampling frames of professional and trade associations and unions is coverage: rarely does association membership encompass a broad coverage of employment in the occupation. In addition, the membership of many associations consists of people in multiple occupations, retired individuals, and other interested parties. Moreover, few associations keep occupational information on their membership, causing the identification of job incumbents in a specific occupation to be problematic. Where coverage of employment in an occupation by association membership is significant and members of the occupation in the association can be identified, a

special frame can sometimes be used to supplement the Establishment Method in a dual-frame sample design.[6]

The Establishment Method provides the advantage of lower response burden and cost when compared with a household survey, as well as good coverage for the large majority of occupations. Response burden and costs are lower for two primary reasons: (1) there are more workers per employer than per household, so fewer contacts are required to identify workers; and (2) employer contacts can be minimized by focusing on those most likely to employ workers in each occupation for which the sample is required. Efficient sample design is possible because the distribution of employment in an occupation is usually a function of the industry of the employer. Employment by occupation by industry is measured by the federal-state OES program national estimates provided by the BLS.

The Establishment Survey provides good coverage of wage and salary employment so long as an acceptable employer sampling frame is available. Coverage of self-employment is more difficult, although the sampling frame used in O*NET surveys includes many establishments operated by self-employed workers. Coverage of unpaid family workers in an employer sampling frame is minimal or zero, but the number of workers involved is quite small (121,000 nationwide in 2006; U.S. Bureau of Labor Statistics, 2007).

The Establishment Method, pretested in 1999 and 2000 remains the primary way to update the O*NET database; most data are currently collected this way.[7] Achieving high response rates with the Establishment Method can be problematic, however, because it requires cooperation at two levels: that of the employer and that of the sampled worker. To date, high levels of cooperation nonetheless have been attained, and this method has proved successful. Although the resulting response rates have been acceptable, the O*NET team continually works to enhance response rates.[8]

An alternative method using occupation experts is employed to optimize the use of burden hours and resources because some occupations are difficult to sample efficiently. This situation occurs when it is difficult to locate industries or establishments with occupation incumbents, employment is low, or employment data are not available, as is the case for many new and emerging occupations. With the Occupation Expert (OE) Method, persons considered experts in the target occupation are surveyed. These experts include supervisors and trainers, as well as experienced job incumbents. The limitation of the OE Method is that locating experts can be difficult. For some occupations, identifying a professional association proves difficult; in

---

[6] This dual-frame sample design is further discussed in Section B.1.1
[7] For a description of the pretest, see Section B.4.1.
[8] For a discussion of current and future efforts to improve response rates, see Section B.3.

other cases the association may lack membership information sufficient to identify experts for a specific occupation.

Using the most appropriate sources of information (e.g., workers, occupation experts, and analysts) and a multiple-method approach, the O*NET Data Collection Program collects and yields high-quality occupational data efficiently.

A.1.3 Summary of the O*NET Data Collection Program

The O*NET Data Collection Program is an ongoing effort to populate and maintain the O*NET database with valid, reliable, and current occupational data. The primary data collection method used to update the O*NET database is the Establishment Method, a survey of establishments and workers within those establishments. The Establishment Method uses a two-stage design: (1) a statistical sample of establishments expected to employ workers in each specific occupation, and (2) a sample of workers in the occupations within each sampled establishment. The sampled workers are asked to complete the survey questionnaires.

Three domain questionnaires are used to collect data from sampled workers: Knowledge (including Work Styles and Education and Training), Generalized Work Activities, and Work Context. Each sampled worker is randomly assigned one of the three questionnaires. The workers are also asked to provide basic demographic information and to complete a brief task inventory for their specific occupation. Workers may complete the paper questionnaire and return it by mail, or they may choose to complete the questionnaire online at the project Web site. Questionnaires are available in Spanish for selected O*NET-SOC occupations. Data for two domains, Abilities and Skills, are provided by trained analysts because of the more abstract nature of the questions.[9]

As described in Section A.1.2, in addition to Establishment Method data collection, the OE Method is used for selected occupations, such as those with few employees scattered among many industries or those for which no employment data exist from which to sample, including new and emerging occupations. With the OE Method, occupation experts are identified and asked to complete all three domain questionnaires, the demographic items, and the task inventory for the specific occupation being surveyed.[10]

A.1.4 Steps in the Data Collection Process

The O*NET Data Collection Program involves the following steps, broadly summarized here and detailed later in separate sections of this Supporting Statement.

---

[9] For the list of questionnaires, number of items and scales, and data sources, see Exhibit 2.

[10] For more details on these data collection approaches, see Section B.2.

Sample Design

      The O*NET Program sampling approaches are designed to create and update the O*NET database in a highly cost-efficient and timely manner while maximizing the reliability of the information in it. The primary method for collecting this information is the Establishment Method, a survey of workers employed in a national probability sample of establishments. This method essentially uses a stratified two-stage design: (1) businesses (the primary stage) are selected with probability proportional to the expected number of employed workers in the specific occupations being surveyed, and (2) a sample of workers (the secondary stage) is selected in the occupations within the sampled businesses.

      For selected occupations that are difficult to complete, a special frame is sometimes used to supplement the Establishment Method of data collection when additional observations are required to complete data collection and a suitable supplemental frame exists, such as a professional or trade association membership list. To be selected for O*NET data collection, an association must (1) represent the O*NET-SOC occupation in terms of the nature of the work performed by its members, (2) contain a high percentage of the total occupational employment, and (3) be willing to provide a list of its members in a form usable as an O*NET sampling frame. The sample selection procedures vary across associations, depending on the type of information available on association members. In general, association lists are sampled with a single-stage, stratified simple random sampling approach. Stratification by geographic location and occupation subspecialty is considered if it is appropriate for the occupation. Approximately 75% of occupations are completed by the Establishment Method.[11]

      The OE Method is considered for use when the Establishment Method would likely be problematic because of occupations with very low rates of employment, new and emerging occupations lacking industry employment data, or occupations whose incumbents are in remote locations that are difficult to access. The OE Method can be used only if the occupation is well represented by one or more professional or trade associations that are willing and able to identify experts in the target occupation. For this method, stratified samples of experts are selected from lists of potential respondents. These potential experts are questioned to determine whether they meet the specified criteria to serve as an occupation expert for their respective occupation. Approximately 25% of the O*NET-SOC occupations are completed by the OE Method.[12]

Data Collection

      Data collection operations are conducted at the contractor's Operations Center in Raleigh, North Carolina, and its Survey Support Department in Research Triangle Park, North Carolina.

---

[11] For additional information about sampling with the Establishment Method, see Section B.1.2.
[12] For additional information about sampling with the OE Method, see Section B.1.3.

For the Establishment Method, the Operations Center's Business Liaisons (BLs) contact sample business establishments, secure the participation of a point of contact (POC), and work with the POC to carry out data collection in the target occupations. The data are provided by randomly selected employees in the occupations of interest. All within-establishment data collection is coordinated by the POC; the BLs do not contact employees directly. After the POC agrees to participate, informational materials and questionnaires are mailed to the POC, who distributes the questionnaires to the sampled employees. Completed questionnaires are returned directly to the survey contractor for processing. Respondents also have the option of completing the survey online. Only when using a special frame do the BLs contact job incumbents directly. BLs also contact occupation experts (the OE Method) directly and attempt to secure their participation; no POC is involved.

Survey support staff mail materials to POCs, job incumbents, and occupation experts, and they receive and process completed questionnaires returned by respondents. Both the telephone operations of the BLs and the mailing and questionnaire-receipt operations of the survey support staff are supported by a Case Management System (CMS). Data-entry software supports the keying and verification of incoming survey data.

With the Establishment Method, each sampled worker is randomly assigned and mailed only one of the three domain questionnaires. For the OE Method, identified experts are asked to complete all three domain questionnaires. All respondents are also asked to provide basic demographic information and to complete a brief task inventory for their specific occupation. Respondents can either complete and return the paper questionnaires by mail, or complete the questionnaires online at the project Web site.[13]

Data Cleaning; Identification and Evaluation of Anomalous Cases

Data cleaning procedures eliminate completely blank questionnaires and insert consistent analysis codes for legitimate skips, blank items, and invalid responses. Anomalous cases are identified so respondents may be removed whose responses either suggest that they are not working in the occupation of interest, or are highly inconsistent with those of the others responding for the occupation. Unusable cases are identified according to prescribed eligibility criteria, such as percentage of items completed. Cases with certain questionable characteristics are flagged for further analysis. These flagged cases include cases with response patterns deviating from those of other cases in the occupation and cases with write-in job titles that do not appear to match the occupation. Responses judged invalid by expert reviewers are excluded from the analysis file.[14]

---

[13] For details about the data collection procedures, see Section B.2.
[14] For a description of both processes, see Section A.16.1.

Weighting and Estimation

Estimates generated from O*NET survey data are computed with sampling weights that compensate for the unequal probabilities of selecting establishments, occupations within establishments, and employees within each selected occupation. In addition, these base weights are adjusted to further compensate for multiple subwaves of sampling, sample adjustment, population under- and overcoverage caused by frame imperfections, and nonresponse at both the establishment and the employee levels.

These weight adjustments can lead to weights that are very large or very small compared with the weights for other sample units. Such weight variability may increase the standard error estimates. When the variation in the weights is large, it is desirable to trim the weights to reduce the variation. For the O*NET estimates, the weighting process involves a weight trimming procedure in which extremely large or small weights are truncated to fall within a specified range. Although trimming weights can introduce bias in the estimates, the variance reduction it achieves usually offsets the potential bias, resulting in estimates with smaller net mean squared errors.

Based on a pooled sample of all completed waves, final estimates are produced. Estimates are computed by summing the weighted observations and dividing by the sum of the weights. Standard errors are estimated with the first-order Taylor series approximation of deviations of estimates from their expected values. These design-based variance estimates are computed with SUDAAN® software (RTI International, 2004). These estimates properly account for the combined effects of clustering, stratification, and unequal weighting—all of which are present in the O*NET data. In addition, estimates with questionable precision are flagged "recommended for suppression" in the O*NET database.[15]

Analysis

Nonresponse is analyzed at multiple levels. Establishment nonresponse in the O*NET Data Collection Program can occur at the verification, screening, recruiting, or sampling stages of selection. Employee-level nonresponse occurs when a selected employee fails to complete and return a questionnaire. Because these same characteristics of establishments and employees were used to adjust the analysis weights, compensation is made to further reduce the size and impact of any nonresponse bias. In addition, item-level nonresponse is assessed, and item response rates are high—generally higher than 90% for Likert scale items, with most exceeding 95%.

For each O*NET-SOC occupation, the degree of interrater reliability (the covariation among ratings) and the level of interrater agreement (the absolute difference among ratings) is

---

[15] For more information on the calculation of weights and variance estimates, see Section B.1.1. For information on the suppression of estimates, see Section A.16.1.

calculated annually. The results of the analyses are used to examine the potential sources of variability across respondents in a specific occupation. As part of a continuous improvement process, these results also inform an evaluation of the O*NET-SOC occupational taxonomy, Content Model descriptors, and scales.[16]

A.1.5 Summary of Response Rate Experience to Date

Establishment Method

Data collection for the O*NET Data Collection Program began in June 2001 with Wave 1.1. Since then (as of September 30, 2007), 76 waves have been completed, and more than 95,000 establishments and 128,000 employees have responded to the survey request. This results in a 75% establishment response rate and a 64% employee response rate. Although it is difficult to make response rate comparisons between the O*NET Data Collection Program and other establishment surveys, these rates compare favorably with those of other establishment surveys.

Some of the design characteristics that make the O*NET Program unique are as follows:

- ***Voluntary rather than mandatory participation.*** Experiments conducted by the U.S. Bureau of the Census suggest that simply informing respondents that their response is mandatory adds about 20 percentage points to overall response (Tulp, Hoy, Kusch, & Cole, 1991; Worden & Hoy, 1992). Because the survey is voluntary, O*NET response rates would be expected to be 10 to 20 percentage points lower than those for the average federally mandated survey.

- ***No direct, personal contact with the survey respondents by the survey organization conducting the data collection.*** The O*NET Program requirement of respondent anonymity means that participation at the employee level relies exclusively on the interactions between the POC and the employee. The survey organization is not allowed to speak to the employee to respond to questions, motivate responses, or follow up on noncompliance. In their review of establishment mail survey response rates, Paxson, Dillman, and Tarnai (1995) found that establishment surveys featuring anonymous mailings typically have lower response rates (by as many as 30 percentage points) than surveys featuring direct, personal contact with the respondents.

- ***Participation required at three stages of response—establishment level, point of contact level, and employee level.*** The typical establishment survey requires participation at only one or two levels: the establishment level and, in some cases, the POC level. By contrast, for the O*NET Program three often distinct entities must agree to participate: the establishment administration, the POC identified in the screening interview, and the employee who is asked to complete the questionnaire. Because very few surveys incorporate such a design, survey methods literature is essentially devoid of examples on which to base a reasonable response rate expectation for the O*NET Data Collection Program. However, it is possible to compare O*NET response rates at each stage with other establishment surveys that

---

[16] For a description of analysis processes, see Sections A.16.1 and A.16.2.

incorporate these stages either separately or in combination. For example, the O*NET establishment-level response rate can be compared with other mail establishment surveys having only one response stage at the establishment level. In addition, the O*NET employee-level response rate can be compared with the response rate of other establishments' self-conducted employee surveys.

A high level of nonresponse is a threat to the accuracy of the estimates. For this reason, the O*NET data collection protocol (described in Sections B.2 and B.3) was designed to achieve the highest response rates possible within respondent burden and survey cost constraints. Paxson et al. (1995) analyzed the response rates for 46 surveys conducted by both governmental and nongovernmental organizations. Among the surveys in their study, 26 were conducted by the Social and Economic Sciences Research Center (SESRC) at Washington State University, and 20 were conducted by the U.S. Bureau of the Census. The SESRC, directed by Dr. Donald Dillman, is well known for its development of the Total Design Method (TDM) approach to mail surveys (Dillman, 1978, 2000) and its high response rates in implementing this method. Furthermore, the 20 U.S. Census Bureau surveys include 12 well-established, ongoing mandatory surveys. The average response rate for all 46 surveys is 63%, but, if only voluntary surveys are considered, the average response rate drops to 55%. These results suggest that, for this type of survey, the O*NET establishment-level response rate of 75% exceeds expectations.

Comparisons of the O*NET employee response rate can be made with surveys that directly sample employees within an establishment. Because federally sponsored surveys of employees within organizations are rare, the literature on their response rates is sparse. Most surveys of this type are employee satisfaction surveys. For example, one well-documented government-sponsored survey of employees is the Public Service Employee Survey, administered to more than 190,000 employees of the federal Public Service of Canada and conducted by Statistics Canada in 1999. A questionnaire was delivered to each employee by a government agent who personally requested that the employee complete the questionnaire and return it by mail. Multiple follow-ups of nonrespondents were made by e-mail and interoffice mail to maximize the response rate. No incentive was used, however; because all sample members were also employees of the organization conducting the survey and could fill out the survey on government time, the use of an incentive was thought to be unnecessary. The final overall response rate for the survey was 55%.

The Office of Personnel Management (OPM) developed the Organizational Assessment Survey (OAS) and has encouraged all federal agencies to survey their employees in order to evaluate organizational performance, benchmark best practices, and align performance with important and measurable outcomes. The experience OPM has had in implementing these surveys in numerous federal agencies provides some evidence of response rates for employee surveys conducted by the U.S. federal government. The OAS design closely resembles the

Canadian Public Service Employee Survey design. The surveys are self-administered and are conducted by each agency for its own employees. Furthermore, like the Public Service Employee Survey, the OAS request to participate is personalized and made directly to the employee by his or her employer. The features of the design offer a significant advantage over the O*NET survey design, as previously noted.

Although the results of the OAS surveys are not publicly available, an official at OPM was able to provide some general information regarding OAS response rates (C. Simons, personal communication, March 21, 2002). According to OPM, response rates for OAS surveys vary considerably by agency, from 30% to 80%. However, the average response across all agencies is approximately 57%. This result is similar to the Canadian experience. It is further supported by Roth and BeVier (1998), who conducted a meta-analysis of 173 surveys in the field of human resource management and organization behavior. They found that surveys implementing many of the response-rate-enhancing features of the O*NET survey had response rates in the range of 23% to 78%, with a median of 51%. These results suggest that the O*NET employee response rate of 64% exceeds expectations for mail surveys of employees within their organizations.

Although results from the survey methods literature and from other federal surveys in the United States and in Canada suggest that a 75% establishment response rate and a 64% employee response rate exceed expectations for federal surveys that use similar data collection approaches, methods for further improving response rates will continue to be explored, and response rates are fully expected to continue to increase as the O*NET Data Collection Program progresses.[17]

Occupation Expert Method

The OE Method is a much smaller but still important component of the O*NET Data Collection Program protocol. As of September 30, 2007, 82 occupations have been completed with this data collection method, and almost 2,400 occupation experts have participated, for an overall OE response rate of 82%. This response rate is higher than the employee response rate for the Establishment Method and may be due to factors such as personalized direct contact, greater total incentive, and generally higher education levels of the eligible population. Some factors may counteract these positive effects, including greater respondent burden and task complexity, as well as potentially less available time for occupations than for the average employee.

---

[17] For the enhancements that have been implemented (since the last OMB Supporting Statement was submitted in September 2005) to maximize response rates and further reduce the impact of potential nonresponse bias in estimates produced from the O*NET data collection effort, see Section B.3.

A.1.6 Statutory and Regulatory Information

Although the O*NET name is not referenced specifically in statute, it is cited at least seven times in the *Code of Federal Regulations.* Furthermore, O*NET information is the foundation, the common language, for carrying out the responsibilities related to the statutory and regulatory sections discussed in this section.

Section 309 of the Workforce Investment Act, Pub. L. 105-220 (Employment and Training Administration [ETA], DOL, 2000) requires the Secretary of Labor to oversee the "development, maintenance, and continuous improvement of a nationwide employment statistics system," which shall include, among other components, "skill trends by occupation and industry." The O*NET Program is the primary response vehicle for collecting skills information across all occupations. Updating the entire O*NET database is a critical component of the nationwide labor market information system to support employer, workforce, and education information needs.

The Workforce Investment Act (WIA) contains numerous references identifying the need for information on the skill requirements of jobs; for example, section 154 requires local determinations of the "skills and education that are necessary to obtain the employment opportunities" in the local area, and section 134 requires provision of "information on job skills necessary" for obtaining jobs listed for the local area.

The WIA also contains references to identifying and assessing the skills of individuals; for example, section 134 provides for assessment services to identify "the skill levels and service needs of adults and dislocated workers." Section 129 allows funds to be used to provide youth with an assessment "which shall include a review of basic skills, occupational skills, prior work experience, employability, interests, aptitudes (including interests and aptitudes for nontraditional jobs), supportive service needs, and development needs of such participant. . . ." In section 136, the "attainment of basic skills and, as appropriate, work readiness or occupational skills" may be included as performance indicators for youth programs. The O*NET Career Exploration Tools, including the O*NET Interest Profiler and O*NET Work Importance Locator, are such assessment tools, designed specifically to relate a person's interests and work values to the information on education and skill requirements for occupations in the O*NET database.

Section 508 of the Rehabilitation Act of 1973 as amended specifies that persons with disabilities shall have access to and use of the same information that persons without disabilities have. To comply with this section, the O*NET Data Collection Program designed its online application, O*NET OnLine, to ensure that the data and information are accessible to the widest possible audience, including persons with disabilities. O*NET OnLine also provides links to

several accommodation and disability resources on the Internet. Furthermore, the O*NET Program Web site, the DOL's ETA O*NET Web site, and O*NET OnLine have all been "Bobby Approved," indicating that the Center for Applied Special Technology has deemed them accessible to persons with disabilities.

The Secretary of Labor's Strategic Plan for Fiscal Years 2006–2011 includes O*NET within Strategic Goal 2—A Competitive Workforce: "Meet the competitive labor demands of the worldwide economy by enhancing the effectiveness and efficiency of the workforce development and regulatory systems that assist workers and employers in meeting the challenges of global competition." Specifically, within Strategic Goal 2, O*NET OnLine is a component of Performance Goal 2E, which states in part,

> Current, comprehensive, relevant, accessible information on occupations and labor markets is critical for planning and delivering demand-driven workforce services. The Department contributes to the State and national workforce information system by creating and broadly disseminating such information, particularly on high-growth industries and occupations, via several types of products collectively referred to as National Electronic Tools. Customers include employers, job seekers, educators, economic developers, State and local Workforce Investment Boards, and participants in the One-Stop delivery system who find this service useful for business, education, economic development and individual career decisions. Specifically, these National Electronic Tools, which include America's Career Information Network, America's Service Locator, Career Voyages, and O*NET OnLine, deliver strong workforce and economic information and career guidance resources and models for two main purposes: 1) to support individuals in making education and training decisions and investments, and 2) to support business and community needs for a prepared and globally competitive workforce.

Providing the most current information on the O*NET skills—the purpose of the proposed data collection effort—is key to supporting the needs of customers through the wide variety of public- and private-sector products that depend in part on the availability of O*NET information. O*NET supplies the common language on skills, knowledge, abilities, and other work requirements and characteristics.

Finally, the predecessor to the O*NET database, the DOT, is cited either directly or indirectly in a number of places as a source of occupational information in support of federal programs. The DOT is cited in the *Code of Federal Regulations* (C.F.R.) in at least 28 places. These citations are generally related to determining disability, administering DOL employment and training programs, or administering immigration, civil rights, and labor standards law. DOL officials responsible for the O*NET Program have been and continue working with DOT users in the federal government, some of whom have made the transition either in regulatory changes or in practices and procedures. Some of those making such transitions include State Department

officials responsible for visas, the U.S. Bureau of Apprenticeship Training, Office of Foreign Labor Certification, and others. Copies of the pertinent sections of the WIA, the Perkins Act, and the specific C.F.R. citations appear in Appendix C.

A.1.7 *Federal Register* Notice

The *Federal Register* notice asking for public comment before submission of this request to OMB is presented in Appendix D. No comments were received from the public.

A.2   Uses, Products, and Services Based on the O*NET Program

The O*NET Program provides essential tools and services for numerous critical federal and state workforce investment functions. These tools also serve job seekers, businesses, and educational institutions across the country. The common language used in O*NET occupational and skill descriptions facilitates communication among the various user groups. This facilitation contributes to our nation's talent development and promotes U.S. competitiveness in a global 21st-century economy.

Users of the O*NET database can access and use the information in various ways. The data are available to the public free of charge through the O*NET OnLine Web application, or the database may be downloaded for use by developers who provide applications targeted to specific communities or audiences. The O*NET database is now the foundation for many programs serving the American workforce, providing information to build skills gap analyses and competency profiles, and to facilitate cross-functional team building. These examples are just a few of the ways the O*NET Program supports the activities critical to maintaining a mobile workforce responsive to changing regional and national economic needs.

A.2.1 The O*NET Database, O*NET OnLine, O*NET Code Connector, O*NET Knowledge Site, and O*NET Career Tools

The O*NET Database

With the June 2007 version of the database, 780 occupations have been updated through the O*NET Data Collection Program. The O*NET database currently provides

- detailed occupational and skill information for more than 800 occupations;
- information on standardized descriptors of skills, abilities, interests, knowledge, work values, education, training, work context, and work styles; and
- occupational coding based on the 2000 SOC.

Since September 2005 there have been four database updates as listed in Exhibit 3.

Exhibit 3.   **Database Updates**

Update	Number of Occupations Updated
December 2005	100
June 2006	100
December 2006	100
June 2007	100

Future updates of the O*NET database will include identification of and information about in-demand and new and emerging occupations.

O*NET OnLine

The O*NET database is provided free of charge to the public through O*NET OnLine, a Web-based application at http://online.onetcenter.org. O*NET OnLine offers users multiple ways to search for O*NET occupations and related data. O*NET OnLine includes

- regularly updated occupational information;

- tiered search algorithms that maximize successful results of searches by keyword (title), occupational code, or partial code;

- a Crosswalk search function that enables users to easily identify O*NET equivalents of other occupational systems, such as the Military Occupational Classification or the Classification of Instructional Programs;

- the ability to Browse by O*NET Descriptor (this search enables users to make cross-occupational comparisons by viewing an occupation's rank order based on a selected knowledge, skill, ability, work activity, interest, or work value variable);

- a sample of reported job titles for each occupational report, providing the user with a broader understanding of the O*NET-SOC;

- report display options (aside from the default Summary Report, users can choose to view a comprehensive Details Report or to build a Custom Report); and

- inclusion of Wages and Employment Trends information (derived from BLS data) and links to Web sites of professional associations, giving the user access to additional information on related specialties, industries, and education and training resources.

O*NET Code Connector

The Code Connector (http://www.onetcodeconnector.org/) was developed to assist workforce professionals needing to code jobs. One-Stop Career Centers, other government workforce agencies, and college career services offices are the most prevalent users of the Code Connector. The Code Connector uses the O*NET database to help users determine the correct occupational code for their job orders. To access an occupation, the user may type in a keyword

or select an occupational group from the home page. On subsequent pages the user is able to refine his or her search to select a specific occupation. The final report contains information to help determine whether the selected occupational code is the best match for the particular job order. This information includes the O*NET-SOC description, Tasks, Related Occupations, Occupation Family, and Detailed Work Activities.

O*NET Knowledge Site and O*NET Academy

Awareness-building and training support for the O*NET system can be found on the Web at http://www.onetknowledgesite.com/ and www.onetacademy.com. The O*NET Knowledge Site provides various user communities with access to O*NET-based collaborative tools. It serves as a forum for users and developers to share O*NET knowledge and gain new insights on user needs and real-world applications of O*NET data. Through the O*NET Academy, O*NET users gain electronic access to live Webinars, self-paced training courses and "how to" tutorials on using the O*NET system and on best practices for its application on the job. To date, more than 5,000 participants have taken advantage of live O*NET Webinars, while many others have accessed the archived recordings of Webinars. Combined, the O*NET Knowledge and Academy Sites have more than 40,000 registered users, including educators, employers, WIB members, national association representatives, students, and job seekers.

O*NET Career Exploration Tools

The Career Exploration Tools are based on a "whole person" concept and are designed for career counseling, career planning, and career exploration. They include the Ability Profiler, the Interest Profiler, and the Work Importance Locator. They also include electronic versions of the interest and work importance assessments—the Computerized Interest Profiler and the Work Importance Profiler. These electronic versions can be downloaded onto a computer system and used at no cost by schools, One-Stop Centers, and others. Users exploring careers may use the Ability Profiler to discover what they do well, the Interest Profiler to identify the types of work they may like to perform, and the Work Importance Locator to determine which occupations will likely be satisfying, according to the their values and needs. The tools enable users to discover important information about themselves and use the information to explore the world of work. Workers may use these tools as aids in exploring career options, in exploring career transitions, and in preparing for career change. The assessments also are designed for use by students exploring the school-to-work transition.

A.2.2 O*NET Web Site Statistics

Use of O*NET products has increased dramatically over the past few years. O*NET OnLine currently averages 775,000 visits per month, up from the 180,000 visits per month 3 years ago. An August 2007 comparison of career information delivery Web sites showed that

O*NET OnLine ranks number one in use among its peer sites, including America's CareerInfoNet and CareerOneStop. The O*NET Resource Center (http://www.onetcenter.org) averages 78,000 visits per month—double the average number of visitors 3 years ago. The O*NET Code Connector Web site averages 63,000 visits per month, another large increase from the average of 29,000 just 3 years ago. In addition, use of career information systems, Web site linkages, user certifications, and O*NET product downloads is widespread. By design, the primary dissemination strategy of the O*NET Program is for the private sector to build O*NET-based products that are audience-specific or tailored for particular user needs.

Career Information Systems

Private and governmental online career information systems using O*NET data and career tools reach millions of customers annually. For example, Bridges (http://www.bridges.com/us/home.html), a private provider of career information services built O*NET data into its system and is used each year by more than one million graduating high school students in the United States and Canada. IntoCareers (http://www.intocareers.com/products.htm) is another career information delivery system (CIDS), which serves more than 7,000 schools and state agencies in the United States. The content components of IntoCareers include data on 480 O*NET-based occupations. State CIDSs also provide widespread access to O*NET-driven career exploration. Many state CIDSs report between 100,000 and 200,000 annual users and more than 100,000 individual career portfolios. The O*NET Program is designed to be accessible to multiple users, in keeping with the directives from the APDOT committee. The O*NET Program encourages these and other service providers and product developers to create applications that deliver O*NET information to the public.

Internet Web Site Linkages

According to an exploratory search conducted in September 2007,

- more than 5,309 sites link to O*NET OnLine,
- more than 400 sites link to the O*NET Code Connector, and
- more than 1,100 sites link to the O*NET Resource Center.

There are many different types of organizations linked to the O*NET Web sites, including the following:

- libraries and career centers based in higher education;
- higher education institutions' schools of business, labor and industrial relations, psychology, education, and counseling;
- government agencies (primarily state labor-related agencies, some federal and local);
- public libraries (especially those offering career and job search assistance programs);

- career counselors, coaches, and recruiters (mostly private-sector vendors);

- career exploration or job search assistance (both private and public sectors);

- public school systems, educational associations, and secondary schools (often recommending the O*NET database as a resource for faculty, parents, and high school juniors and seniors);

- human resources management;

- law firms specializing in immigration law;

- vocational rehabilitation, or occupational medicine and health; and

- international sites in Turkey, Canada, the United Kingdom, Holland, Japan, Bangladesh, and elsewhere.

User Certifications

When individuals or organizations download the O*NET database or intend to use all, any one, or any part of one of the O*NET Career Exploration Tools to develop value-added products, they are asked to voluntarily register their use by completing a certification form. As of September 30, 2007, the certification database holds 2,179 entries. Because registration is voluntary, this number likely reflects only a portion of those actually using O*NET products for their own applications. An overview of those user groups that have submitted certification forms for their use of O*NET products appears as Exhibit 4.

Exhibit 4.   Main Organization Types Submitting O*NET Certifications

Organization Type	Percent
Educational services	18
Government; public administration	12
Computer systems design; programming services	8
Employment services	8
Vocational rehabilitation services	6
Temporary help services; employment placement agencies; human resource and executive search consulting	5
Information and Internet publishing; broadcasting; software publishers	5
Health care	3
Individual and family services	2
Professional, scientific, and technical Services; legal services	2
Military	1
Other	30
Total	100

O*NET Product Downloads

From January 2002 to September 2007, downloads of O*NET products totaled 298,800 (Exhibit 5).The use of O*NET products and tools continues to increase. The O*NET Program, through continuous improvement efforts based on user needs and advancing technology, works to efficiently develop products that meet customer demands in both the public and private sectors.

Exhibit 5.   O*NET Product Downloads

Product	Number of Downloads
Database	59,232
Career Exploration Tools	
Ability Profiler	85,794
Interest Profiler	88,054
Work Importance Locator	47,554
Computerized Interest Profiler, Work Importance Profiler Software	55,689
Total Career Exploration Tools	277,091
Other (e.g., Toolkit for Business)	44,706
Total O*NET Products	343,506

A.2.3 Examples of O*NET Data and Products in Use

The O*NET Program provides comprehensive, up-to-date occupational information that is used directly by the public through access to the O*NET Web sites, as well as indirectly through value-added products delivered by both governmental and private developers. Discussed under the following general categories of product developers and service providers are examples of how O*NET-based products benefit the public:

- government organizations and workforce investment systems,
- assessment and career information delivery systems,
- business and staffing agencies,
- educational and research institutions,
- private companies and commercial products,
- U.S. armed forces,
- community and faith-based organizations, and
- international users.

Government Organizations and Workforce Investment Systems

Federal and state agencies are using O*NET products to achieve their goals of service to employers and the public. Business development specialists, WIBs, career counselors, and others rely on the common language of O*NET products to build connections required for a strong workforce. Discussed here are examples of O*NET products used by One-Stop Career Centers, by state department of labor Web sites, in state responses to mass layoffs, and in a special program sponsored by the National Institutes of Health.

The state of Maine used O*NET information to help dislocated workers make informed decisions about employment opportunities after a layoff. Based on a comparison of knowledge, skills, and abilities, the jobs the workers were leaving were first linked to O*NET occupations in order to identify likely opportunities requiring similar skills. O*NET information was supplemented with state-level wage information, job openings, and information about retraining possibilities so workers could choose reemployment opportunities best suited to them.

California Labor Market Information (LMI) has used O*NET information as part of the state's rapid response to mass layoffs. LMI's guide for workforce development professionals assists dislocated workers in reevaluating their skills and relating those skills to other occupations, particularly in high-demand fields and among new and emerging jobs. The guide outlines a step-by-step process enabling a dislocated worker to build an occupational profile with Skills Search and Detailed Reports in O*NET OnLine, together with multiple LMI data sources. From the Detailed Reports pages, an applicant copies and pastes Tasks statements, Detailed Work Activities and other descriptors used to enhance resumes. Skills Search is used to compare the applicant's job skills to similar skills in new careers.

To address the pressing need for talented and skilled computer 3-D graphic artists and traditional animators, California's Employment Development Department conducted an industry study with the endorsement of the California Skillsnet Consortium. O*NET's survey data collection instruments and other materials were used in the study. Relying on O*NET data, human resources personnel modified local training and education initiatives to help close the gap between industry needs and local workers' skills. Training programs defined the skills and requirements for the evolving occupations of computer 3-D graphic artists and traditional animators. State and local leaders forged important partnerships with local establishments, educators, WIBs, community-based organizations, and other stakeholders to meet industry demands and develop needed talent.

The Lancaster County Workforce Investment Board in Pennsylvania uses O*NET information in research and service delivery activities designed to support economic development efforts in the county. Using an industry clusters model, The Lancaster WIB

cooperates closely with the economic development community. After the industry clusters are defined, the primary occupations are profiled, and a Skills Map is developed. Information for all occupations and the various career ladders in the cluster are analyzed to identify common skills and attributes, as well as those unique to a particular occupation or career ladder. The resulting Skills Map is a helpful way to involve the educational community in addressing skills gaps, taking a systemic approach to the identification of industry's skill needs. O*NET data are a key part of the Skills Map used to develop the profiles, providing extensive information about skills and knowledge requirements of occupations in each cluster.

Wisconsin's Department of Workforce Development (DWD) works with a network of 78 local Job Centers (the equivalent of One-Stops) across the state. Each has a resource area and a mission to provide high-quality career information to job seekers and other customers. In many of the centers, local staff wanted or needed to become more knowledgeable about the career development needs of diverse clients and the array of resources available to assist them. To help staff improve their career development skills and services, DWD's Division of Workforce Solutions developed a 3-day, competency-based training curriculum that includes an introduction to O*NET information. Students use O*NET OnLine, the O*NET Work Importance Locator, and the O*NET Interest Profiler. Staff have discovered that these O*NET tools provide a good foundation for approaching the career exploration process.

The Connecticut Department of Labor used O*NET data to respond to the Workforce Investment Act requirements that states assess (1) current and future job opportunities in the state, (2) the skills necessary to obtain these jobs, and (3) the skills necessary to meet the economic development needs of the state. To meet these requirements, it collected and analyzed labor market information and published an extensive report. O*NET Skills and skill descriptions were used in the report sections describing skills necessary for Connecticut's high-demand occupations and industry sectors important for state economic development.

The Workforce Oklahoma Career Connection Center One-Stop partner agencies are using the O*NET Program in a broad effort to identify and meet specific employer needs in a timely, effective, and cost-efficient way. The center uses a survey of employer needs in customizing a service mix to meet these specific needs. O*NET information is built into this survey to help employers clearly define their employment needs by job title and skill bundles. For one such employer, center staff developed a series of job descriptions for skilled welders. These job descriptions were used as a basis for an in-house certified training program for welders.

Designed by the State Department of Labor, the New York CareerZone Web site is an O*NET delivery system that has revolutionized the way New Yorkers access career information. The NYCareerZone system is now a statewide resource for teachers, counselors, and career

center staff. The system engages students and job seekers in career exploration and helps them develop their skills in career planning. Users build electronic portfolios linked to state learning objectives and O*NET-SOC occupational information. NYCareerZone developers incorporated the O*NET Interest Profiler and Skills Search capabilities into the portfolios to help students identify their interests and relate these interests to appropriate career possibilities. Using job profiles drawn from the O*NET database, students can explore specific occupations. The profiles include state-specific wage and outlook information related to the O*NET-SOC occupations and include a link to a state college database.

California's Labor Market Information Division (LMID) gathers, analyzes, and produces occupational information to inform people about California's economy and to aid in labor-related decisions. In response to layoffs at Pillowtex (a large household textile manufacturer that closed plants employing 6,000 workers across eight states), California's LMID created an instruction kit that incorporates parts of O*NET OnLine. Including the O*NET Find Occupations and Skills Search capabilities, this toolkit enables use of O*NET data and other resources to make dislocated workers aware of their current skills, and it suggests occupations that match these skills.

Assessment and Career Information Delivery Systems

Organizations are routinely relying on O*NET products as the basis for their assessment and career information delivery systems. Community colleges, university career centers, and many higher education Web sites also have integrated O*NET products into their career services to students. From nonprofit organizations to high-powered consulting firms, specific examples discussed here indicate the range of organizations using O*NET products.

Adaptiqs provides customized survey and assessment applications that capitalize on the technology of artificial intelligence for efficient, accurate administration. Adaptiqs has developed a job match system using O*NET data and questionnaires in conjunction with a Bayesian network model. The goal of this recent project was to create an adaptive instrument using responses to O*NET Knowledge, Skills, Abilities, and Work Context Questionnaires to identify the most relevant job types for job candidates.

Assessment Associates International (formerly Human Resource Consultants) has developed the Work Behavior Inventory (WBI), an assessment to measure Work Styles in employees. Work Styles are personal characteristics that can affect how well someone does a job. The WBI assesses Work Styles as defined in the O*NET system and provides the user with information that can be used to

- focus and guide leadership development efforts,

- pinpoint specific strengths that may help with career advancement,

- identify areas for training and skill enhancement,

- identify work environments that may support success for the individual, and

- help guide career choice and transition decisions.

Keys2Work, a private, nonprofit organization, offers a community-based assessment and career development system that helps students, job seekers, employers, and providers of education and training. Using work-related skills assessments and online tools powered with O*NET data, the system demonstrates the relevance of school subjects to future careers and earnings, thereby contributing to talent development in schools. Using ACT's WorkKeys®, students assess important work-related skills and match them to occupations. Keys2Work links these occupations to O*NET data, expanding students' perspectives, allowing them to explore a wider range of occupations and obtain extensive information about them. Students can learn about common tasks and work activities; other knowledge, skill, and educational requirements; and interest profiles of workers. O*NET data provide the connection to related information on wages and employment outlook. Using O*NET data, Keys2Work shows students that what they earn in the future depends directly on the skills they acquire in school.

Profiles International (PI) is one of several assessment companies that incorporates O*NET data in customized workforce development tools for public- and private-sector use. PI's products are designed to help businesses improve their hiring practices, reduce turnover rates and costs, and enhance workforce harmony and performance. O*NET information is used extensively with the company's clients, including a number of public workforce agencies. In PI's system, job seekers begin with an assessment of their "soft skills," such as job behavioral traits, thinking style, and occupational interests. Their results are matched to O*NET occupations and presented in a Career Compatibility Report. The report lists occupations that may be a good fit, explains how to use O*NET OnLine to obtain more information about the occupations, and explains the results of the assessment. PI sets up for its clients private Web sites on which managers can post their job openings and screen for suitable applicants. In addition, the site describes O*NET OnLine and how to use O*NET information in drafting position descriptions and in other human resource activities. This resource is especially helpful to small and mid-size companies without large human resource departments or formal position descriptions.

The SkillsNET Foundation provides open-resource solutions for human-capital management; WIB regional skill systems; and other large-scale knowledge, skills, and abilities as well as competence framework systems. The SkillsNET Foundation has used the O*NET Content Model in the development of SkillObject Designer®, a Web-based job analysis tool.

Included in this product are O*NET data on knowledge, skills, and abilities; generalized work activities; and tools. SkillsNET Designer is a resource for large, complex organizations, requiring inputs from incumbents, subject matter experts, supervisors, and strategic decision makers.

Health and medical science career exploration with LifeWorks™ is accomplished through an interactive career development Web site operated by the National Institutes of Health's Office of Science Education. Driven by O*NET data, the LifeWorks search engine, or Career Finder, offers an array of information on more than 100 health and medical science careers. Staff designed the Web site for middle school and high school students, parents, mentors, teachers, and guidance counselors. As a first step, students scan a list of O*NET Job Families and select the ones that most interest them. Second, they identify the kinds of jobs that suit their interests, using the O*NET interest categories. Third, they select skills they have or want to acquire. The Career Finder then generates a customized list of health-related careers, with brief descriptions, matching the students' selections. By clicking on a title, students can view job-specific information on the summary page. If they like, they can access details about the occupation, including employment outlook, salary, suggested high school courses, related careers, and more.

Higher education institutions are one of the most prevalent types of user of O*NET products for assessment and career information systems. A Web search shows O*NET products at work in most state educational systems. A few examples include AlabamaMentor.org, Arizona State University CRESMET, CaliforniaColleges.edu, Cascadia Community College (Washington), University of Central Florida, CTMentor.org (Connecticut), IllinoisMentor, Middle Tennessee State University, MississippiMentor, Normandale Community College (Minnesota), PennsylvaniaMentor, TexasMentor, and Texas State University. Web links to these and other online sites offering career services based on O*NET products can be found in Appendix E.

Business and Staffing Agencies

Some of the best-known staffing services are using the O*NET database in their applications. Brick-and-mortar locations, as well as Web sites such as Monster.com, recognize that O*NET data are key to their mission of connecting job opportunities to job seekers and thereby enhancing workforce mobility in response to changing economic needs.

The HR-XML Consortium has used the O*NET database in developing specifications that allow organizations to capture occupational information and use it efficiently in their established human resource processes, such as building competency models, conducting job analyses, and developing performance appraisal systems. These firms include Adecco, Aetna,

CareerBuilder, Cisco Systems, IBM, Kelly Services, Manpower, MetLife, Microsoft, and Monster.com, as well as many other technologically innovative companies. The O*NET Program is the featured content provider in an HR-XML schema explaining to member organizations and to the public how to capture and exchange skills, abilities, and related human resources information.

Manpower, a worldwide provider of high-value staffing services with nearly 1,100 offices in North America and 4,400 offices in 73 countries, provides jobs to more than 2 million persons every year and services more than 400,000 customers worldwide. The O*NET occupational and skills taxonomy helps Manpower match the right person to the right job. The O*NET system also offers a systematic structure that enhances Manpower's analysis of the employment marketplace and its tracking of staffing trends. By incorporating O*NET structure into its procedures, Manpower has benefited by being able to

- identify accurately the types of placements each field office makes;

- locate field offices where the highest need exists;

- more accurately consolidate information for various types of analysis, including marketing analysis; and

- begin the process of having Manpower offices in other countries map their occupations to O*NET, enabling Manpower to consolidate information for global reporting more efficiently.

Educational and Research Institutions

With the help of the occupational information contained in the O*NET database, colleges are developing educational curricula, and research organizations are conducting a broad array of research projects. Presented here are a few examples.

The Association of College and Research Libraries, a division of the American Library Association, provided the following information on the O*NET Web sites in their *C&RL News* (http://www.ala.org/ala/acrl/acrlpubs/crlnews/internetreviews/200402/onet.cfm):

Job seekers, students, workers, employment professionals, counselors, and others interested in exploring occupations and careers will find this site a great place to start. . . .

O*NET Online is a well-designed search engine for exploring the database. . . .

Beyond O*NET Online, this site is chock-full of information on the Consortium, other easily accessible O*NET products, career assessment guides and tools, research and technical reports, data collection methods, planned products and ways to contribute to the process. . . .

O*NET Online alone, however, is worth the stop for students searching for relevant careers at any point in their academic journey. Through it they can explore occupations,

match skills, find out about salary and trends, or just see what is out there. Researchers, employment specialists, and others can dig deeper and discover a wealth of information about the world of work likely available nowhere else. (Valentine, 2004, February)

CareerJournal.com, the *Wall Street Journal* Executive Career Site, conducted a study designed to identify the "best careers," those occupations for which workers report a high degree of satisfaction. CareerJournal contracted the polling firm Harris Interactive to survey workers about their jobs, whether they were satisfied, and what job qualities contributed to their satisfaction. Four factors emerged for jobs held by highly satisfied workers: intellectual stimulation, strong job security, high level of control and freedom of action, and extensive direct contact with customers or clients. After identifying the common factors reported by highly satisfied workers, CareerJournal used occupational data in the O*NET database, looking for O*NET occupations with high ratings related to three of the four contributing factors. (Job security was evaluated with BLS employment projections.) CareerJournal identified eight occupations providing the most satisfying work experience:

- curriculum and instructional coordinators,

- high school special education teachers,

- hospital and clinic managers,

- management consultants and analysts,

- medical researchers,

- physical therapists,

- sales, marketing, and advertising managers, and

- social workers, counselors, and related managers.

The National Academies of Sciences is evaluating O*NET information as a tool for making important human-capital decisions. As part of its research on changing worker requirements, the National Academies of Sciences commissioned a paper by the Human Resources Research Organization to investigate the feasibility of O*NET information to assess changing worker skill requirements (Tsacoumis, 2007, May). The paper describes the key elements of the O*NET system, and provides examples of the many ways O*NET has been used by the education, public, and business communities to improve workforce decisions. The paper concludes that O*NET is a rich and important data source that can be used in many different ways to assess changing skills necessary for workers to be successful in today's workplace.

Lehigh Carbon Community College in Pennsylvania has a federal demonstration grant to coordinate and develop curricula in three areas in nanotechnology at the postsecondary level and in two areas at the secondary level. Working primarily with Lehigh Career and Technical Institute, project staff find O*NET OnLine a handy, practical tool for faculty at both institutions.

It offers an ideal starting point for occupational research essential to creating programs that respond to contemporary business and industry needs and that contribute to talent development. O*NET OnLine provides detailed descriptions of occupations, including Semiconductor Processors, which is related to nanofabrication. Researchers used the tasks, activities, and other elements for this occupation as a basis for further research and discussions with industry experts and representatives. In addition, it was a good starting point for developing task statements that teachers could turn into performance objectives for courses. Moreover, project staff could identify and obtain detailed descriptions of other occupations related to those that they already target and that could become part of the new training. The use of O*NET OnLine is also recommended for reviewing existing secondary and postsecondary programs.

Temple University's Center for Professional Development in Career and Technical Education in Pennsylvania supports professional growth and learning among all levels of educators who work with career-bound students. The center offers certification and degree programs for secondary career and technical education teachers, curriculum supervisors, cooperative education coordinators, and career and technical school directors in the 17 counties of eastern Pennsylvania. Center faculty and staff incorporate O*NET OnLine in their courses on program planning and evaluation, curriculum development, and cooperative education. They also use it as an aid in structuring occupational competency assessment committee reviews. With its wealth of data on occupations, O*NET OnLine provides a reliable and readily accessible resource for planning and evaluating programs and developing curricula in career and technical education.

Rotman Research Institute is using O*NET data for a National Institutes of Health study on occupation attributes and dementia. The O*NET database has proved to be an invaluable resource, providing comparable variables across a wide range of occupations. The database provides a basis for the assessment of the characteristics of patients' occupations and for the subsequent task of relating these characteristics to cognitive style and brain degeneration in various forms of dementia.

The John J. Heldrich Center for Workforce Development is located at the Edward J. Bloustein School of Planning and Public Policy at Rutgers, the State University of New Jersey. The center teaches the use of O*NET information in its Working Ahead: The National Workforce and Career Development Curriculum. This credentialed 120-hour curriculum teaches career counseling and guidance to front-line staff in workforce development, community organizations, and community colleges. Working with a network of state departments of education and state career resource networks, the Heldrich Center is developing a training guide and student manual so that counselors and education staff can learn to use the O*NET database as a career exploration tool for middle school, high school, and college-level students

(http://www.aacc.nche.edu/Content/NavigationMenu/NewsandEvents/Workforce_Development_
Institute/AACC_Working_Ahead.PDF).

Private Companies and Commercial Products

Private companies are relying on O*NET information for in-house purposes such as job description writing and employee development, as well as for product development. Presented here are a few examples of how companies are benefiting from O*NET's broad range of detailed occupational data and tools.

The corporation Business and Legal Reports uses O*NET data in the Job Description Manager module of its HR.BLR.com Web site (http://hr.blr.com/tour.aspx?view=jdm). O*NET descriptors included in the Job Description Manager are occupation title, position summary, tasks, skills, attributes, and experience and education.

Piedmont Natural Gas is using O*NET tools to help reduce turnover among entry-level employees. It has used O*NET Career Exploration Tools and occupational data to create a formal system to better match entry-level job applicants' talents, abilities, and preferences with the duties and working conditions of the jobs for which they are applying. Part of the project included conducting a job analysis on each job family. In addition to employee surveys, interviews, focus groups, observations, and job-specific training materials, Detailed Reports from the O*NET OnLine Web site provided information about the necessary knowledge, skills and abilities for the targeted positions.

iNQUATE Corporation has incorporated O*NET data into its human-capital software application. With a patent pending, iNQUATE is an integrated seven-module application that accommodates the development and communication of corporate goals and objectives, automates the performance management process, and engages managers in the best use of human capital while these managers analyze costs associated with maintaining a workforce. iNQUATE integrates with onsite databases or information available from external providers through Web services. All of iNQUATE's seven modules work together to simplify the process and eliminate duplication of information and effort (http://www.inquate.com/).

The U.S. Armed Forces

The U.S. military has recognized the value of O*NET data and career tools in its various transition programs, recruiting activities, and human systems development projects. Presented here are a few examples of the range of O*NET products being put to work in the armed forces.

The Army is currently evaluating the O*NET system as a tool for military job analysis and classification. The Human Resources Research Organization, one of the O*NET's partners, is working on a pilot test, using O*NET questionnaires and the O*NET database to describe and

classify positions in the Army. The Army believes that the success of this project will allow it to map required competencies and performance criteria and to evaluate and enhance personnel development practices.

The U.S. Department of Defense's Armed Services Vocational Aptitude Battery (ASVAB) Career Exploration Program offers tools to help high school and postsecondary students learn about career exploration and planning. Developers of the ASVAB Program wanted to change its philosophy to emphasize wider career exploration and decision making among its participants. Completely redesigned, the program now uses O*NET data to broaden occupational choices for nearly a million ASVAB participants at more than 14,000 of America's high schools annually. Students are encouraged to explore a variety of career possibilities suited to their interests and learn to make decisions based on information about themselves and about careers, instead of exploring just a few occupations that match their current abilities.

The North Carolina Military Foundation teamed with the North Carolina Military Business Center to create a database and interactive Web site that enables businesses to link their needs to the competencies of troops exiting the military. One of the challenges faced by the troops and business leaders alike is identifying the knowledge, skills, and abilities shared by military and civilian jobs. Using a keyword related to a job opening, employers are able to search to find related military occupations and information on how many military personnel in these occupations are returning annually to civilian jobs. The user can view additional information about these occupations, including a list of related civilian job titles. Further exploration is available through a link to the related occupations in O*NET OnLine. This Web site helps employers and transitioning military personnel come together through the common language of O*NET.

Military Personnel in Transition at Offutt Air Force Base in Omaha, Nebraska, provides a comprehensive 5-day Transition Assistance Program (TAP) for personnel leaving military service. As part of this program, the Nebraska Workforce Development Labor Market Information Center presents information on career exploration, employment searches, and training opportunities in the civilian job market. O*NET OnLine is prominently demonstrated as a tool to help TAP participants relate their skills and military experience to appropriate occupations and learn about a wider range of career possibilities as they transition to civilian life and employment. Military personnel can thus tap into the entire range of O*NET occupational information to explore career possibilities in the civilian sector. O*NET OnLine is a key element of the TAP for military personnel in transition.

HireVetsfirst.gov uses O*NET OnLine as a military skills translator to help veterans returning from active military duty reenter the workforce. Using their military occupational

classification in O*NET OnLine's crosswalk function, they can identify equivalent civilian occupations. They can see typical tasks, skills, knowledge, and abilities; learn wage levels and employment outlooks both nationally and in their own localities; and search for actual job openings in areas of their choosing.

Community and Faith-Based Organizations

Community and faith-based organizations can easily provide career services to their clients with O*NET products, which are in the public domain. The online database of occupational information, as well as the free downloads of computerized Career Exploration Tools, is helping to broaden the career outlook for special populations.

The American Foundation for the Blind (AFB) uses O*NET data to help the visually impaired expand their employment possibilities. CareerConnect™ is a free employment planning resource for persons who are blind or visually impaired. Sponsored by the AFB, it helps them learn about the range and diversity of occupations available in the labor market. It also provides mentors and information about assistive technology that can help them do the work. This practical, user-friendly resource incorporates O*NET data to supply essential information for career exploration and to expand the universe of jobs for persons with visual impairments.

Catholic Community Services in Baton Rouge, Louisiana, offers a free Senior Employment Program for workers aged 55 years or older in the city and 17 other civil parishes, mostly rural. With more than 25 years of experience, the agency is now a mandated partner in Louisiana's Workforce Investment Act initiatives. The program uses the O*NET Career Exploration Tools to help retirees and other older adults find jobs. The O*NET Interest Profiler is a key tool in the process that helps open clients' minds to what the workplace offers them and what they can offer the workplace.

The Boys and Girls Clubs of America developed a career preparation program to help teens make sound educational decisions, explore a variety of careers, and develop skills for success in the workplace. A system, CareerLaunch™, supports this program. This expanded career preparation Web site incorporates O*NET data, enabling teens to find O*NET career profiles related to their interests.

At Edgewood Terrace in northeast Washington, DC, the Community Preservation and Development Corporation (CPDC) is helping unemployed and underemployed adults build career plans, develop job skills, and find employment. In one component of the CPDC career and skills enhancement program, staff use O*NET OnLine, among other resources, to help participants identify potential career goals. Participants leave the class with a long-term career plan fashioned as a résumé builder. Some go on to specialized training programs in information

technology, others seek employment or further education in other fields, but all have an action plan, with next steps toward a better future.

The Traumatic Brain Injury Program at the Tennessee Rehabilitation Center relies on O*NET OnLine to help clients identify jobs they want to do and can do well. Team members refer to detailed reports of the skills, knowledge, abilities, and other characteristics of occupations as they tailor their services to the individual's interests and goals. Employers benefit from the program's Job Placement Specialist's assistance in finding competent workers, those workers well matched to the employers' particular needs.

International Users

The international interest in O*NET products has been phenomenal. Governments, researchers, human resources consultants, and career guidance programs in more than 20 countries have requested permission to make use of the O*NET database, questionnaires, and career tools. Each year, the O*NET Web sites log thousands of visitors from virtually every geographic region in the world.

O*NET Customer Service has received inquiries from Romania, Japan, Nigeria, the Netherlands, Singapore, Hong Kong, and many other countries requesting information, assistance, and permission to use O*NET products. China and Mexico are developing national occupational information systems based on O*NET data and O*NET questionnaires. Australia uses the O*NET database of information to provide content to their own Australian Standard Classification of Occupations.

In addition to international interest in the O*NET database, career guidance professionals in foreign countries are adapting the O*NET Career Exploration Tools for their own use. The Technical Education and Skills Development Authority (TESDA), a government agency attached to the Philippine Department of Labor and Employment, has employed the O*NET Ability Profiler in its aptitude assessments of hundreds of thousands of high school students. In addition, TESDA is distributing the O*NET Computerized Interest Profiler to regional, provincial, and district offices, as well as to their training institutions, as part of their e-Profiling program, targeting out-of-school youth, career shifters, course shifters, and unemployed adults. This example is only one of how the O*NET Career Tools are being used by governments and private developers of career assessments around the world.

A.2.4 Examples of the O*NET Program in Published Literature

Presented here are some examples of references to the O*NET Program in publications. For a thorough list of research articles, books, book chapters, technical reports, and presentations referencing the O*NET Program, see Appendix E.

Job Seeker's Online Goldmine: A Step-by-Step Guidebook to Government and No-Cost Web Tools, by Janet E. Wall, Ed.D., is a comprehensive resource of step-by-step directions for dozens of excellent job-seeking, career transition, and career exploration tools. It includes references to the O*NET Program and O*NET online tools and products throughout its 222 pages. It was published by JIST Works in 2006 (Wall, 2006).

O*NET data were used to rank occupations for their potential offshorability, as described in a working paper by Alan S. Blinder of Princeton University. O*NET data on tasks, knowledge, skills, abilities, work activities, work context, interests, work styles, and work values were used to assign an index number to each occupation, indicating in rank order how easy or difficult it would be to move the occupation offshore. This initial paper assesses the potential movement of specific jobs offshore and may inform public policy (Blinder, 2007).

The Career Pathways Handbook (2006) provides a comprehensive tool for guidance counselors and individuals by incorporating information from the O*NET database. Written by Jim Cassio (http://www.cassio.com), it includes the latest U.S. employment statistics, as well as O*NET information including occupation titles and definitions, common job tasks, key skills and abilities, and related occupations (Cassio, 2006). Cassio's organization, the Workforce Information Group (http://www.work-info.com) provides workforce and labor market information research, publishing, and systems development, using O*NET as a resource.

The O*NET Program is featured in a chapter of *Unfocused Kids—Helping Students to Focus on their Education and Career Plans* (Schmelter-Davis, 2004). Author Harvey Schmelter-Davis has demonstrated the value of O*NET information in dozens of workshops for counselors, career development specialists, teachers, and students. At Rutgers University he developed a training package to introduce O*NET resources to educators. In his chapter Schmelter-Davis explains the O*NET system from a counselor's perspective and suggests how to use it in career counseling and teaching. He describes how others are using O*NET resources to help younger persons focus on their strengths and interests and relate them to career opportunities. In a handbook written for practitioners, *Job and Work Analysis: Methods, Research, and Applications for Human Resource Management* (2007), authors Michael T. Brannick, Edward L. Levine, and Frederick P. Morgeson describe the O*NET Program as a hybrid model including descriptors of both the work and the worker, as compared with other systems focused on either the worker or the work itself. The authors tout the O*NET Content Model as comprising virtually any information about a job that a user could want. The various job analysis systems available are discussed in light of various job analyses objectives, with consideration for real-world constraints, such as time and money (Brannick, Levine, & Morgeson, 2007).

"The Big 'O' in Your Job Search: O*NET," by Jane M. Lommel, Ph.D., is the title of a "NetWorking" article that appeared on Gary Johnson's Brave New Work World Web site (http://www.newwork.com). This article provides an extensive overview of the entire O*NET system, including descriptions of the skills and competency information available at http://online.onetcenter.org. The article highlights ways different workforce professionals, counselors, and employers use O*NET to accomplish their goals (Lommel, 2004). Additional Web sites, such as http://www.careerclusters.org/16clusters.htm, developed by the National Association of State Directors of Career Technical Education Consortium, are linked to this article. At the career clusters Web site, O*NET concepts are related to 16 major clusters into which almost all occupations can be categorized. Other O*NET Program–related sites linked to the "Big 'O'" article offer online training and information (http://www.onetacademy.com) and access to the O*NET Career Exploration Tools (http://onetknowledgesite.com/jobseekers.cfm).

O*NET data are frequently used for scientific research. For example, one recent study was conducted to determine the validity of O*NET descriptors for predicting mean literacy scores. This article, *Linking O*NET Descriptors to Occupational Literacy Requirements Using Job Component Validation* (LaPolice, Carter, & Johnson, in press), will be published in the summer 2008 issue of *Personnel Psychology*. Results showed that literacy was highly predictable from O*NET descriptors: the authors were able to predict estimates of literacy scores for all O*NET-SOC occupations.

A.2.5 Looking to the Future with Competency Models

DOL has identified industry and occupational competency models as a resource that is key to businesses and industries articulating their workforce needs. The O*NET Content Model and database provide a valuable resource for the development of reliable and valid competency models. In turn, these models form the basis on which curriculum developers and training providers develop objectives and outcomes to ensure that workers have the right skills. The ETA acts as a broker of information in the area of competency models and skill standards development by promoting, supporting, and disseminating information (http://www.careeronestop.org/competencymodel/). The information gathered through the O*NET Data Collection Program continues to provide the workforce investment system partners with tools necessary for investment in human-capital development.

To date, three high-growth industry competency models have been developed, and more are in progress. For some industries the information available from industry sources is neither comprehensive nor readily available to developers. In these cases, model developers can access occupational information in the O*NET database to identify the critical work requirements, worker skills, and knowledge for occupations in the industry. The O*NET data are an essential

source of information, reducing the time and effort required to develop industry competency models. When industry information is available or competency models have been developed, a comparison with the O*NET data validates information from industry sources.

The University of Central Michigan has used O*NET Content Model domains to identify the competencies for a leadership competency model for students. The model consists of five dimensions, each with multiple competencies drawn from the O*NET Content Model and an existing management taxonomy of work activities. Work requirements and worker characteristics were reviewed to identify skills, knowledge, abilities, work styles, generalized work activities, work context, and organizational context relevant for each dimension. Although the resulting model was intended for students and was distributed to members of the campus community, it was also distributed to interested employers because of likely relevance to a variety of organizations.

The Tennessee Valley Authority used O*NET information as presented in the Career Pathways tool, which was designed by Economic Modeling Specialists, Incorporated, to evaluate workforce competencies in northeastern Mississippi. By showing Toyota that competencies of workers in the declining furniture industry in northeastern Mississippi were 98% compatible with the needs of the auto industry, O*NET data helped bring a Toyota plant to northeastern Mississippi.

A.2.6 Workforce Innovation in Regional Economic Development

WIBs at both the state and the local levels are a critical component of the Workforce Innovation in Regional Economic Development (WIRED) regions. They work together with civic, investment, academic, entrepreneurial, and philanthropic leaders to provide leadership by aligning and leveraging their respective activities, goods, and resources.

The Central Western Maine WIRED initiative includes a plan to build a smoother interface with O*NET assessments. This objective is stated in the draft plan submitted to the Maine Jobs Council, WIA Title I Strategic Plan Modification for Central Western Maine Local Workforce Investment Board, Local Workforce Investment Area III.

Michigan's West Coast WIRED Initiative uses the Career Readiness Certificate, based on WorkKeys[®] and O*NET Job Zone information, as a means of linking worker readiness to jobs available in the local economy (ACT, 2006). WorkKeys, an online assessment linked to O*NET occupations, measures foundational skills of Reading for Information, Applied Mathematics, and Locating Information. It assigns a score and grants a Career Readiness Certificate. The certificate is then related to jobs available in the local economy.

The Career Readiness Certificate is being used in several states as a part of state, regional, and local workforce development initiatives, including North Carolina, Virginia, Indiana, Kentucky, and New Mexico. Many others are in the process of implementing a Career Readiness Certificate program. (See http://www.crcconsortium.org/state-news.htm.)

A.3 Uses of Information Technology

The O*NET Data Collection Program employs the latest in information technology systems and procedures to enhance the quality of the data, minimize the burden on the responding establishments and job incumbents, and reduce the overall cost of the data collection effort.

A.3.1 Web Questionnaires

Electronic versions of the O*NET questionnaires are available via the Internet to sampled job incumbents and occupation experts. Many of the benefits of the paper questionnaires are replicated in the electronic questionnaires. Specifically, users are able to start and stop multiple times without losing data. They can return to their partially completed questionnaire at any time during the survey period and resume where they stopped. A respondent may also review and edit previous answers as necessary. In addition, an on-screen progress meter keeps respondents informed of their movement through the questionnaire.

Advances in Web technologies and security, as well as the prevalence of establishments' and employees' access to Web browsers, have made Internet-based data collection both feasible and practical. Internet use continues to accelerate, and the use of the Web questionnaire has increased with it. For 2004, 15.8% of employee respondents used the Web questionnaire, a substantial increase over the 3.7% who responded via the Web during the 1999 pretest. In 2007 the number of job incumbent respondents choosing the Web version increased to 20.3%; for occupation experts, the Web usage rate was 24.0%. The paper questionnaire cover and informational materials mailed to respondents continue to highlight the option of completing the survey on the Web.

The paper and Web versions of the questionnaires were designed to be optimal for their respective modes of administration. The questionnaire design literature suggests that this approach is essential to reduce mode effects. That is, if each questionnaire is designed to minimize measurement error in its particular mode of interview, mode effects are also minimized. For example, in the paper version, multiple questions appear on a single page of the questionnaire. However, in the Web version, only one question appears per screen (although the respondent can navigate at will through the instrument). The literature on Web survey design see (see, e.g., Fuchs, 2002) suggests that one question per screen for Web surveys reduces measurement error and therefore the effects of administration mode. This change was the only

important one necessary for the Web version, because both instruments are self-administered. In fact, to ensure comparability between the paper and Web responses, the formats and wordings of the questions and response categories for the two versions remain identical.

A.3.2 Project Web Site

An O*NET data collection Web page application has been developed to support and assist with the O*NET Data Collection Program (http://onet.rti.org). This site is divided into two major sections: the public and restricted-access sections. The restricted-access section is further subdivided into two areas: the online questionnaires area and the project management area.

The goal for the public section is to support the establishment-recruiting process. This section is accessible to the public, without restrictions. The public section includes O*NET background information, endorsement letters, frequently asked questions, copies of the questionnaires, and links to other O*NET Program–related Web sites. The purpose of this section is to provide establishments, sampled workers, and occupation experts with readily accessible information about the data collection effort and uses of the data.

The restricted-access section contains "sensitive" information available only to certain groups, such as survey respondents, project managers, and data-entry staff. This section is controlled by a user ID and password authentication scheme. The Web server includes a secure socket layer (SSL) certificate to allow encrypted transmission of all information over the Internet. This same technology is used by e-commerce Web sites to secure credit card numbers. In addition, no "cookies" are used. (A cookie is data given to a user's Web browser so that the browser will return the data to the server or Web site during subsequent requests.) Some Internet users distrust Web sites that deposit cookies on users' computers and may even configure their computers to prohibit cookies, so the "cookie-less" techniques are to ensure that the site will perform as expected, whether or not a user has disabled cookies.

The online questionnaire area of the restricted-access section provides sample members an alternative to pencil-and-paper self-interviewing. Only persons who have been selected to participate in the survey have access to this area. Unique user IDs and passwords are assigned to each job incumbent and occupation expert by a central office computer system at the time they are selected into the sample; names and other personally identifiable information are not obtained. The ID and password are provided with other survey materials to the sample member. Before allowing access to the online questionnaire area, the Web site confirms the validity of the ID and password and verifies that a completed survey form has not already been received.

Having entered this portion of the site, respondents are

- informed that participation in the survey is voluntary,

- assured that their survey responses will remain confidential,

- limited to seeing only the questionnaire they have been asked to complete,

- permitted to stop at any point and continue responding later,

- permitted to skip questions they choose not to answer, and

- permitted to review and change previous responses.

On the last page of the survey, respondents confirm that they have completed the questionnaire; then they exit the questionnaire area of the site and are thanked for their participation. Their user ID is automatically deactivated at this time. Any further attempts to log in will not be possible, because the system recognizes these users as having completed the survey.

The database containing the survey data are not accessible via the Internet; it resides on a server inside the contractor's firewall. The Web data collection application encrypts and transmits data from the respondent's computer into the survey database. Only authorized project staff, operating from inside the firewall, have access to the survey database.

The project management area of the restricted-access section contains data collection management reports and information. Access accounts for this area are created for managers of the O*NET Data Collection Program. This portion of the Web site serves as an intranet for the O*NET Program, facilitating communication among contractor staff, program staff at the National Center for O*NET Development, and DOL. Production reports are posted nightly and include summaries of the progress of establishment recruiting, questionnaire shipment and receipt, and overall data collection status. Additional applications include a secure, centralized data file repository used by statisticians, analysts, and others to share results of specialized, nonroutine analyses and reports. The project management area of the Web site reduces the cost of the data collection effort by centralizing and streamlining features used by the project team.

A.3.3 The Case Management System and Data Collection Utilities

The O*NET CMS is a Web-based control system that supports and monitors the data collection activities of the BLs, the mailing of informational materials and questionnaires, and the receipt of completed paper and Web questionnaires. Enhancements have been made to the CMS that allow greater flexibility and effectiveness in the communications between the BLs and the POCs in sample establishments. For example, visual cues (icons) in the CMS allow the BLs to prioritize and customize their approach for certain types of establishments that require special procedures. Another enhancement is the ability to do ad hoc package modifications, such as the inclusion of special endorsement letters, based on O*NET-SOC occupations. Because the packages for specific O*NET-SOC occupations are customized, the POCs receive targeted

materials that help communicate the data collection mission. In addition, tools have been developed to help operations management load-balance the number of cases assigned across the team of BLs. Another feature permits shipping of questionnaires to more than one POC in an establishment, for use with multisite organizations. These and other enhancements have improved the study's effectiveness and efficiency.

In an effort to help manage supplies (envelopes, brochures, questionnaires, etc.) associated with data collection, an inventory tracking system, integrated with the CMS, has been developed and deployed. As the BLs place orders for informational materials and questionnaires to be shipped to survey participants, the inventory system updates reports to show the expected remaining inventory. As stocks run low, the staff are alerted to replenish supplies. The system provides a means to reconcile physical and expected inventory. As a result, the system has improved the efficiency of ordering, storing, and shipping data collection materials.

Questionnaires are prepared for each respondent by an automated order-fulfillment system. The system detects the questionnaire domain type and occupation assigned to the respondent, and it prints a complete package, including the individually customized and labeled questionnaire, ready to pack and ship.

The CMS automatically assigns questionnaire domain types sequentially as new respondents are added to the sample. Depending on how many of each type are actually returned by respondents, it is possible to receive enough responses of a particular questionnaire type for an occupation before receiving the total desired number of questionnaires of all types. The system continually monitors the number of returned questionnaires by type and overrides the sequential assignment of questionnaires as appropriate so that only the questionnaire types that are still required are sent. This feature enhances efficiency by eliminating the shipping of questionnaires that are no longer necessary to complete the data for an occupation. Moreover, by continually focusing resources only on the remaining questionnaire types required, it reduces the total time required to complete any occupation.

A.3.4 Additional Uses of the Internet for Data Collection

The O*NET Program uses the Internet to gather additional occupational information, such as lay titles and tools and technology. This use expands the O*NET database, providing easily maintained current information about tools and technology and enhancing users' ability to find occupations relevant to their training and expertise by giving them a wider range of search terms. This enhancement is accomplished without additional burden to the public and at less cost than other means of data collection.

"Web scraping" is employed to gather tools and technology linked to in-demand occupations in high growth industries, with particular emphasis on identifying rapidly changing

information technologies, such as database user interface and query software and analytical or scientific software. Tools and technologies linked to specific occupations are found on job posting, career education, professional association, and other Web sites. Occupational analysts link the tools and technologies to the United Nations Standard Products and Services Code®, an online classification system for tools and services. Before the information is published in O*NET OnLine, rigorous reviews are performed to ensure the quality and usefulness of the data. Information collected from the Web in this way represents a significant enhancement to the data available to O*NET users.

O*NET lay titles are collected from unmatched search terms submitted to America's Career InfoNet (ACINet), a DOL-sponsored online resource for job seekers that helps individuals explore career opportunities and make informed choices about employment and education. O*NET lay titles, as well as tools and technology, are also collected from the ACINet Job Description Writer, an online tool offered by ACINet to help employers and human resource specialists write job descriptions. Search terms, such as Sports Medicine Specialist and Nanotechnologist, that do not result in occupational matches in ACINet are captured electronically. These terms, together with job titles collected through the Job Description Writer, are reviewed by O*NET occupational analysts for potential inclusion in the O*NET Lay Title database. Lay titles greatly enhance the keyword search functions for O*NET OnLine, America's Career InfoNet, and O*NET Code Connector, a Web application to assist workforce professionals in matching job titles to O*NET-SOC codes.

Use of Internet web sites further enhances the O*NET database, provides greater search capabilities to users, and reduces burden. It also allows for rapid update of the data through user input and at minimal cost.

A.4   Efforts to Identify Duplication

To avoid duplication and reduce cost, several portions of the O*NET Content Model are provided from existing data sources. Specifically, as discussed in Section A.1, the domain of Workforce Characteristics—including information on industries, job opportunities, and pay—is obtained through links to existing labor market information databases. Information about occupational licensing, certifications, national industrial skill standards, and related instructional programs is provided from existing sources, such as the Manufacturing Skills Standards Council, the National Retail Federation, and several Web sites, including http://www.careertools.org/ certification and http://www.acinet.org/acinet/lois_start.asp.

The exhaustive reviews of existing labor market and occupational information, conducted by the Advisory Panel for the *Dictionary of Occupational Titles* and DOL's DOT review staff, as well as subsequent current research, identified no other comprehensive, valid, and reliable

sources that could be used for the data items included in the O*NET database. The development of the O*NET Program has also involved staff and advisors with many years of experience in labor market and occupational information who are familiar with existing data sources. In fact, as discussed in Section A.2, many existing systems that provide detailed occupational information are actually using information based on O*NET data or the predecessor DOT.

The few existing sources with similar measures are too limited to be used in the O*NET database. Some existing sources are valid and reliable—for example, information from the Office of Personnel Management and the U.S. Department of Defense—but are not comprehensive, because they represent only those jobs in federal civilian employment or the military. Some private sources of job analysis information exist; however, they are based on job analyses conducted on an as-needed basis rather than on a representative sample of employers and workers. They are therefore limited in their coverage and not representative of the entire workforce. Furthermore, these analyses are not comparable because they do not use the prescribed O*NET common language to describe occupational requirements; it is not practical to combine them, because they include dissimilar components. Finally, these private data sources are not available to the general public.

A.5   Efforts to Minimize Burden on Small Establishments

All sizes of establishments are represented in the O*NET estimates for most occupations. For some occupations, the targeting strategy used in selecting an efficient sample may lead us to omit some small establishments from the sampling frame, but this omission occurs for few occupations. The omission is allowed when it is clear that sampling small establishments will greatly reduce the efficiency of the data collection or when it is clear that incumbents from small establishments are not working in the mainstream of the occupation.

Given that establishments of all sizes should be represented in the samples for most occupations, specific design provisions have been undertaken to avoid overly burdening small establishments. For example, Exhibit 6 shows the distribution of establishments by number of employees on the Dun and Bradstreet (D&B) list used as the sampling frame for O*NET data. The exhibit also shows the expected distribution of the O*NET sample of establishments by number of employees. O*NET sampling selects small establishments at a much lower rate than that at which they occur in the population: although 80.5% of establishments employ fewer than 25 employees, the O*NET sample will consist of only 39.7% of such small establishments. On the other hand, large establishments (with 250 or more employees) will make up 25.8% of the O*NET sample but make up only 0.4% of all establishments. Thus, the O*NET sample relies more heavily on large establishments in order to reduce the burden on small establishments with

Exhibit 6.   **Distribution of Frame and Sample Establishments by Employment Size**

Number of Employees	Total Number of Frame Establishments[a]	Frame Distribution (Percent)	Actual Distribution of O*NET Sampled Establishments[b] (Percent)
1–4	9,053,605	60.3	23.9
5–9	1,760,339	11.7	6.5
10–24	1,275,844	8.5	9.3
25–49	482,385	3.2	14.2
50–99	264,782	1.7	10.2
100–249	140,541	0.9	6.6
250–499	33,539	0.2	14.7
500–999	12,607	0.1	6.4
1,000–4,999	6,943	0.1	4.4
5,000+	532	0.0	0.3
Unknown	1,976,604	13.2	3.4
Total	15,007,721	100.0	100.0

[a] Data based on October 2007 Dun and Bradstreet (D&B) frame of establishments.

[b] Data based on distribution of prior O*NET samples that used the D&B frame. Future O*NET samples will be similarly designed.

few employees. The disproportionate sampling of large and small establishments is properly accounted for in the analysis weighting, resulting in statistically consistent estimates.

In addition, data collection procedures place lower burden on small establishments than on large establishments. When a small establishment is selected, it likely employs fewer of the targeted occupations and has fewer employees working in the occupations. Thus, a POC at a small establishment generally spends less time preparing sampling lists and distributing questionnaires than a POC at a large establishment, which is more likely to employ several of the targeted occupations, with a large number of employees working in the occupations.

A.6   Consequences of Collecting the Information Less Frequently

The O*NET database is the most comprehensive source of occupational information in the United States. No other similar, comprehensive, reliable, and valid source exists. If O*NET data are not collected, U.S. citizens, industry, business establishments, military, government and educational institutions, and the workforce investment system will have few options that meet their needs for occupational information. O*NET data are used to develop industry competency models and occupational competency profiles. O*NET data also include information on transferable skills and are used for skills gap analysis, facilitating a mobile workforce responsive to changing economic needs.

The use of O*NET data by industry, employers, software developers, job seekers, students, educators, and workforce development specialists supports a well-functioning U.S. labor market and workforce investment system—both essential to U.S. competitiveness in a global 21st-century economy. O*NET-SOC occupations conform to the SOC, permitting O*NET data to be linked to and analyzed with sources of information on current occupational employment and trends, wages, and demographic data. Its electronic format is freely accessible, either through O*NET OnLine (http://online.onetcenter.org), or as a download from the O*NET Resource Center (http://www.onetcenter.org).

The initial 3.1 version of the O*NET database was populated with information derived from the DOT. So far, updated data have been published for 780 O*NET-SOC occupations. Additional data releases are planned through 2012 to update selected high-growth occupations and release new data on new and emerging occupations. The consequences of discontinuing data collection would be that the millions of users who rely on O*NET data for business and career decisions, for educational programming, and for work in human resources or workforce development would instead be using portions of information that are out of date and incomplete. If data were collected less frequently, the currency of some data would become questionable, especially for occupations that are changing as a result of new technologies. The focus of data collection on high-growth and new and emerging occupations reflects the need to provide current information in a demand-driven economy that is rapidly changing.

A 3-year extension of the O*NET Data Collection Program is being requested for the period spanning calendar years 2009 through 2011. This extension will provide for the updating of selected high-growth occupations and for data collection activities for new and emerging occupations. A dynamic and progressive U.S. economy requires continuous improvement to the data on which so many decisions are based. Millions of people are currently using O*NET information and the numbers continue to expand as public agencies and private developers integrate O*NET data into their systems and products. The O*NET database provides valid, reliable, and current occupational information crucial to a strong U.S. workforce. O*NET database updates are scheduled to occur once a year to incorporate newly collected information on recently surveyed occupations. A schedule for data analysis is provided in Section A.16; schedules for data collection and analysis are subject to annual appropriations.

A.7   Special Circumstances

The study will be conducted in a manner consistent with the guidelines in 5 C.F.R. § 1320.9. There are no special circumstances that might require deviation from the guidelines.

A.8   Consultation Outside the Agency

The *Federal Register* notice initiating the 60-day public comment period was published on May 16, 2008. A copy of the notice appears in Appendix D. No comments were received from the public.

The OMB clearance package was reviewed by two expert consultants, Drs. John P. Campbell and Michael A. Campion. Dr. Campbell is professor of psychology and industrial relations at the University of Minnesota, and Dr. Campion is professor of management at Purdue University. Both are industrial/organizational psychologists who are widely published in their field and who are intimately familiar with the O*NET data collection process and the types and uses of the data collected. Revisions responsive to their comments have been incorporated, as appropriate, in the clearance package.

The data collection contractor, RTI International, has several mechanisms in place to obtain ongoing feedback from study participants. The BLs have multiple telephone contacts with POCs and occupation experts; they are careful to document in their call notes significant comments or suggestions they receive. Letters and brochures sent to POCs and occupation experts offer a toll-free number they can call with questions or comments. In addition, the project Web site, https://onet.rti.org/, which both study participants and the general public can access, has a "contact us" tab that offers both a toll-free number and an e-mail link. Furthermore, the online Web questionnaire provides respondents an opportunity to submit comments with their questionnaire responses. All feedback received from these sources is promptly reviewed by project management staff and has resulted in several modifications to the data collection protocol.

A.9   Payments or Gifts to Respondents

Since the origins of the survey, incentives have been offered to POCs, establishments, and employees to encourage their participation in the O*NET Data Collection Program. Although the current procedures are designed to minimize respondent burden, the effort for the company and the POC participation is not insignificant.

A.9.1 Incentives for the Point of Contact and the Employer

The POC's responsibilities include the following:

- reading the introductory package to become familiar with the purpose of the O*NET Data Collection Program and the role of a POC;

- seeking permission within the company, as necessary, to participate in the O*NET Data Collection Program;

- making a roster of all employees at the location who work in as many as five different occupations;

- participating in a sampling process that selects as many as 20 employees from these occupations and maintaining this sample roster for future reference;

- distributing questionnaires to the sampled persons within the company and addressing their questions and concerns about the survey; and

- distributing follow-up materials to employees, including thank you–reminder cards, replacement questionnaires, and e-mailed requests for cooperation.

Because POCs are the only link with the O*NET respondents, they must be fully committed to the data collection process. The POC is the O*NET Program's representative within the establishment who communicates the importance of the O*NET Program.

The employer is also critical because he or she is being asked to

- support the O*NET Data Collection Program by agreeing to the company's participation in data collection,

- provide support and encouragement to the POC to carry out his or her responsibilities,

- allow the POC to provide information regarding the number of persons employed in the establishment in the occupations of interest, and

- support the participation of the sampled employees.

Incentives for both POCs and employers are essential to encourage the highest response rates possible. POCs receive a clock and a framed Certificate of Appreciation from DOL. The clock is battery operated, has the O*NET logo on its face, and is suitable for display on a desk or shelf. The Certificate of Appreciation is printed in color on card stock, bears the DOL and O*NET logos, displays the POC's name, and is signed by a high-ranking DOL official. It has a solid oak frame and Plexiglas cover and is suitable for displaying on an office wall. The total cost of both gifts is approximately $10, although their perceived value likely far exceeds this amount.

The employer receives the O*NET Toolkit for Business. The Toolkit is an O*NET Program information packet, including a guide for writing job descriptions, that managers can use for human resource planning. These materials are attractively organized in colored, glossy folders.

The continuation of these incentives is planned for both the POC and the employer because these incentives seem to work quite well and are the products of considerable thought and experimentation. For example, in the early years of O*NET data collection, the POC received a mouse pad with built-in calculator in place of the Certificate of Appreciation. The mouse pad was not well received by the POC, and its perceived value, with the clock, was

judged to be inadequate for all the POC's efforts. Although not formally tested, the perceived value of the Certificate of Appreciation has significantly exceeded its actual cost, and it has been very well received by the POCs.

In addition, an experiment was conducted from 2002 to 2004 to evaluate an additional incentive of $20 to the POC. Essentially, the incentive had no effect on POC cooperation rates or employee response rates while it significantly increased data collection costs. The experimental findings suggested that the current POC and employer incentives are adequate for maximizing response rates at a reduced cost (Biemer, Ellis, Pitts, & Robbins, 2006).[18]

A.9.2 Incentives for the Employee

In keeping with what has been done since 2001, each employee is offered a prepaid incentive of $10 to ensure that a high percentage of the job incumbents respond by completing the questionnaire.

Monetary incentives have the greatest potential impact when the respondent has to exert some special effort, such as taking a test or filling out a multi-item questionnaire. The incentive encourages respondents in a task requiring higher levels of involvement and commitment than the typical one-time, face-to-face interview. Although the O*NET questionnaires are not tests, the cognitive demands they place on respondents resemble test-taking demands in that the respondents must assess the requirements of their job. The monetary incentive is instrumental in impressing upon the respondent the importance of this rating task. Respondents who perceive the rating task as important will likely make thoughtful, carefully considered assessments rather than hasty ones, thus improving the reliability of the data.

In addition, the monetary incentive is important because respondents are encouraged to complete the questionnaire, which takes about 30 minutes, on their own time rather than on the job. This encouragement not only minimizes the burden on employers, but also improves the quality of the data because, otherwise, busy workers might be underrepresented in the data, which would bias the estimates for job performance.

The monetary incentive may at least partially offset its inherent cost through efficiencies created in the data collection process as a result of higher response rates. For the job incumbent respondents especially—though they are not viewed as a difficult-to-reach population in the usual sense—considerable effort and resources are expended to identify and reach them through the sampling process. They represent a worker in a specific occupation in a specific establishment in a specific industry. The expense of reaching that particular respondent justifies the cost of a monetary payment to ensure a high rate of response. With regard to the size of the

---

[18] For a more detailed discussion of the Biemer et al. (2006) results, see Section B.4.

employee incentive, payment amounts were evaluated in the pretest to determine the optimal means to maximize the response rate. On the basis of these data, a $10 cash incentive has been used since the initial wave of data collection and will continue to be used in future waves.

A.9.3 Incentives for Occupation Experts

Occupation experts provide data for approximately 25% of the O*NET-SOC occupations. Each occupation expert is offered a $40 monetary incentive to complete three domain questionnaires, or about $13.33 per questionnaire. This incentive is slightly higher than the $10 offered to Establishment Method respondents for completing one domain questionnaire. This increased incentive is necessary to gain cooperation from what is often a rare group of experts for an occupation. Moreover, it seems commensurate with the effort involved in responding to multiple questionnaires, given that the average occupation expert's salary is higher than the average employee's salary.

A.10 Assurance of Confidentiality

The contractor for the data collection task, RTI International, has extensive experience in protecting and maintaining the confidentiality of respondent data collected from surveys. To ensure confidentiality, RTI has drawn from its experience in designing the data collection procedures incorporated in this program.

Survey data are collected from job incumbents (Establishment Method) and from occupation experts (OE Method). Informational materials mailed to potential respondents contain essential program information and assurances of confidentiality that enable the person to make an informed decision about his or her voluntary participation in the data collection effort. Examples of informational materials provided to survey participants appear in Appendix F.

Employees sampled at establishments are asked to complete their questionnaires on their personal time, not company time. This stipulation enables the employee to select a comfortable and private setting, if desired, in which to complete the questionnaire. All respondents have a choice of completing paper questionnaires or completing the questionnaires online at the project Web site. Paper questionnaires are mailed directly to RTI in a business-reply envelope provided by RTI. The individual responses are processed according to a study ID number. All O*NET Data Collection Program staff are required to sign a confidentiality pledge that assures each respondent that the confidentiality of responses to the questionnaire will be maintained. Only authorized staff have access to the completed instruments and data files. The completed and processed questionnaires are stored in a secure document-control area until federal authorization has been granted to destroy them. All computer files, including those associated with the control system, are password protected.

The Internet-based system that allows respondents to provide their survey responses electronically has restricted access, including a user ID and password authentication scheme for respondents. The Web server includes an SSL certificate to allow encrypted transmission of all information over the Internet, the same technology used by e-commerce sites. The database containing the survey data are not accessible via the Internet; it resides on a server inside the RTI firewall. Only O*NET Data Collection Program staff have access to the master survey database.

The O*NET questionnaires (see Appendix A) collect very little personal information about the respondent, and what is collected contains no identifiers, such as personal name or place of employment. No individual-level data are published, nor are they accessible or provided to anyone except the O*NET Data Collection Program staff. Published results are made available only in aggregate, as one set of estimates for an entire occupation. Furthermore, no demographic data (e.g., sex, race) are released, even in aggregate form. Finally, estimates are not produced for any subpopulations within an occupation, such as geographic region or sociodemographic group, which otherwise may have allowed the identification of an individual.

Before publishing the O*NET tables on the public Web site, the O*NET Program team thoroughly examines the tables for any risk of disclosure of confidential information. In particular, each table is analyzed to identify any "sensitive" cells (i.e., cells that may reveal too much information about an individual employee). Willenborg and Waal (1996) have recommend using an (n,k)-dominance rule that a cell be regarded as sensitive if the sum of the largest n contributions account for more than k% of the total cell value. Willenborg and Waal further recommend $n = 5$ and $k = 80$. Because every sampled employee contributes only one response, these recommendations translate into a minimum cell size of 5/.80, or about 7. In fact, the minimum cell size for the O*NET tables is 10 respondents; any sensitive cells with fewer than 10 are suppressed. In addition, the (n,k)-dominance rule assumes a complete census. As Willenborg and Waal note, when applied to tables based on samples and where the cell entries are weighted averages, the (5,0.80) rule affords even greater disclosure control. This extra control ensures that these O*NET tabular data pose no disclosure risks to any individual respondent.

A.11 Questions of a Sensitive Nature

Only one question in the O*NET questionnaires may be considered to be of a sensitive nature. In the Background section, the survey asks respondents if they consider themselves to have a disability, using the disability questions developed for the 2000 Census of Population. Completion rates for this item indicate that the great majority of participants (96%) elect to complete the question.

The O*NET sampling strategy is to randomly select participants at the individual level. The disability question, together with the demographic questions, provides descriptive information about the sample of respondents. In addition, these data may allow for some broad comparisons with the characteristics of the general population working in the occupation, using information from the 2000 Census.

Respondents to the O*NET survey are informed that responding to all questions, which includes disability status and the other demographic characteristics, is voluntary. They complete the survey on their own time, in a private setting if they choose. No identifying information, such as the respondent's name or place of employment, is recorded on the questionnaire.

A.12 Estimates of Annualized Hour Burden

As described in Section A.1.2, there are two protocols for O*NET data collection—the Establishment Method and the OE Method. An estimated 75% of occupations are completed by Establishment Method. The OE Method is used for occupations as necessary to improve sampling efficiency and avoid excessive use of burden, as when it is difficult to locate industries or establishments with occupation incumbents, employment is low, or employment data are not available, as is the case for many new and emerging occupations.

The Establishment Method uses a two-stage sample, with establishments selected at the first stage and employees selected at the second stage. Thus, there are burden hours associated with both establishments and employees. For each participating establishment, a POC is identified to coordinate data collection activities in the establishment. In Exhibit 7, the first and second columns of the Establishment Activity section present the number of responses and assumed average burden per response for the POC's activities; these averages were obtained from previous years of O*NET data collection experience. When the total establishment burden was estimated, the estimated number of establishments that will complete each activity was multiplied by the average burden and summed across the activities.

The Employee Activity section of Exhibit 7 is based on the fact that selected employees under the Establishment Method will complete one domain questionnaire requiring an average of 30 minutes of effort, while occupation experts complete all three domain questionnaires, requiring an average of 90 minutes of effort.

Exhibit 7.   Estimate of Hour and Cost Burden by Year

	Number of Responses per Sample Unit	Average Burden per Response (Minutes)	2009		2010		2011	
			Sampling Units	Burden Hours	Sampling Units	Burden Hours	Sampling Units	Burden Hours
Establishment Activity								
Verification calls to initial contact at establishment	1	2	16,083	536	14,667	489	13,833	461
Screening call to POC	1	3	13,349	667	12,173	609	11,482	574
Initial recruitment call to POC	1	12	7,743	1,549	7,061	1,412	6,659	1,332
POC Creation of occupation lists for Sampling	1	20	4,336	1,445	3,954	1,318	3,729	1,243
Call to POC to sample workers	1	10	4,336	723	3,954	659	3,729	622
POC's distribution of questionnaire packets	1	15	4,076	1,019	3,717	929	3,505	876
Follow-up calls to POC	4	2	4,076	543	3,717	496	3,505	467
Total, establishment[a]	NA	NA	NA	6,482	NA	5,912	NA	5,575
Employee Activity								
Establishment Method employee respondents	1	30	16,321	8,161	16,525	8,262	14,430	7,215
Occupation Expert Method respondents	3	30	500	750	500	750	500	750
Total, employee[b]	NA	NA	NA	8,911	NA	9,012	NA	7,965
Total Respondents								
Private sector	NA	NA	12,548	NA	11,443	NA	10,793	NA
Federal government	NA	NA	534	NA	487	NA	459	NA
State/local/tribal governments	NA	NA	267	NA	243	NA	230	NA
Subtotal, establishment respondents	NA	NA	13,349	NA	12,173	NA	11,482	NA
Individuals/households	NA	NA	16,821	NA	17,025	NA	14,930	NA
Total, all respondents[c,d]	NA	NA	30,170	NA	29,198	NA	26,412	NA
Total Burden Hours								
Private sector	NA	NA	NA	6,093	NA	5,557	NA	5,241
Federal government	NA	NA	NA	259	NA	236	NA	224
State/local/tribal governments	NA	NA	NA	130	NA	118	NA	111
Subtotal, establishment burden hours	NA	NA	NA	6,482	NA	5,911	NA	5,576
Individuals/households	NA	NA	NA	8,911	NA	9,012	NA	7,965
Total, all burden hours[e]	NA	NA	NA	15,393	NA	14,923	NA	13,541

	Cost Burden		
Respondent Type	**2009**	**2010**	**2011**
Establishments[f]	$304,654	$277,864	$262,025
Employees[g]	$175,814	$177,807	$157,149
Total	$480,468	$455,624	$419,174

Note: NA = not applicable.

[a] Includes total burden time for all establishments (private sector, federal government, and state/local/tribal governments).

[b] Includes total burden time for individuals/households.

[c] The total number of respondents across all three years = 85,780.

[d] The total number of respondents was derived by summing the number of POCs (row name is "Screening call to POCs"), Establishment employees (row name is "Establishment Method employee respondents"), and occupation experts (row name is "Occupation Expert Method respondents").

[e] The total number of burden hours across all three years = 43,857.

[f] Assumed hourly rates: $47.00 for 2009–2011.

[g] Assumed hourly rates: $19.73 for 2009–2011.

Exhibit 7 also displays the estimated number of sampling units and the estimated burden hours. As shown in Exhibit 9 in Section A.16, we assume that 75 occupations will be completed under the Establishment Method and 25 will be completed under the OE Method each year. For 2009, establishments are estimated to expend 6,482 burden hours, and employees are estimated to expend 8,911 burden hours, for a total of 15,393 burden hours. In 2010, establishments are estimated to expend 5,912 burden hours, and employees are estimated to expend 9,012 burden hours, for a total of 14,924 burden hours. In 2011, 5,575 burden hours are estimated for establishments, and 7,965 burden hours are estimated for employees, for a total of 13,540 burden hours. The slight decline in total burden hours across the 3-year period of 2009 to 2011 (15,393; 14,923; 13,541) results from initiating data collection for most occupations during 2009, which causes many of these occupations to complete data collection during the final 2 years.

The Total Respondents section of Exhibit 7 shows the estimated annual number of respondents by category and overall. The respondent totals include the number of POCs (the row name is "Screening call to POC"), the number of Establishment Method employee respondents, and the number of Occupation Expert Method respondents. The category totals are based on prior sampling experience. The total number of respondents across all 3 years is 85,780.

The Total Burden Hours section of Exhibit 7 shows the annual number of burden hours by category and overall. The burden hour totals include both establishment burden and employee burden. The category totals are based on the estimated number of respondents for each category, the number of responses, and the average burden per response. The total number of burden hours across all 3 years is 43,857. This 3-year total reflects a slight reduction in burden compared with the previous 3-year period of 2006 to 2008, during which 48,737 burden hours were requested (Employment and Training Administration, 2005).

O*NET Establishment Method data collection has been designed to minimize the burden on the selected establishments:

- Establishments are asked about no more than 10 occupations each, with questioning terminated once 5, or sometimes fewer, occupations are identified at an establishment.

- Establishments are asked to complete rosters of employees only for the five or fewer occupations identified.

- Establishments are selected no more than once within a 12-month period.

- No more than 20 employees are selected from an establishment across all of its selected occupations.

- If an occupation proves difficult to complete under the Establishment Method, the dual-frame approach may be used to supplement it. For occupations that are difficult to sample, the alternative OE Method may be used.

A.13 Annual Reporting Burden Cost

Exhibit 7 displays the estimated annualized cost to respondents for burden hours by year. The cost burden was estimated with use of average wage rates obtained from the June 2007 Employer Costs for Employee Compensation Summary issued by BLS on September 20, 2007. The average wage per hour for private industry and state and local government was $19.38, which was inflated to a December 2007 hourly wage rate of $19.73 based on the Employment Cost Index. This is the wage rate used for estimating the employee cost burden. Given that the POC will often be a human resources manager, the June 2007 wage rate of $46.17 for the Management, Professional and Related category was inflated to a $47.00 hourly wage rate for December 2007 for estimating the establishment cost burden.

The exhibit shows that the combined establishment and employee total cost burden is $480,468 for 2009, $455,624 for 2010, and $419,174 for 2011. As with the burden hours, the slight decline in total cost burden across years results from initiation of data collection for most occupations during 2009, which causes many of these occupations to complete data collection during the final 2 years.

There are no respondent costs for capital or start-up or for operations, maintenance, and purchase of services. There are no costs to the employers, POCs, or sampled employees other than the time it takes them to participate in the survey.

A.14 Estimates of Annualized Cost to Government

The estimated annual cost to the government for the O*NET Data Collection Program for 2009 to 2011 is approximately $4.6 million. This estimate includes all direct and indirect costs of conducting the sampling, data collection, and analysis activities of the O*NET Data Collection Program.

A.15 Reasons for Program Changes or Adjustments Reported in Sections A.13 and A.14

Exhibit 8 compares the projected burden hours for calendar years 2009 to 2011 with the average annual burden as estimated for FY2006 to FY2008 in the OMB Supporting Statement dated September 2, 2005. The projected total annual burden hours for years 2009 to 2011 range from 13,541 to 15,393. These annual projections are slightly less than the average annual burden of 16,246 hours requested for the previous 3-year period (2006 to 2008). In addition, as indicated in Exhibit 7, the total burden hours for the 2009 to 2011 period, 43,857, reflects a slight reduction in burden compared with the 2006 to 2008 period, for which a total of 48,737 hours were requested (Employment and Training Administration, 2005). Over the years of conducting the O*NET Data Collection Program, experience gained in selecting efficient samples of

establishments has led to reduced burden hours. Continuous improvement efforts include refinements in occupational clustering, industry identification, and the design of data collection waves, as well as the ability to capitalize on previous sampling experience and history for some of the occupations. The Skills domain is now rated by analysts, leading to a reduction in the number of survey types from 4 to 3, thus reducing the employee sample size by 25%. Calculations also reflect a decreased time needed to complete surveys. However, despite the fact that the Supporting Statement for the 2005 ICR correctly requested 48,737 burden hours, ROCIS incorrectly reflected 24,621 instead, which were the burden hours for the single year of 2006 and did not correspond to the number of respondents figure in ROCIS, which was the total across all 3 years: 2006 through 2008. Thus, the previous NOA continued the error by approving 24,621 hours for this collection. The current Supporting Statement correctly requests 43,857 hours, a real reduction of 4,880 hours, rather than the apparent, but erroneous, added burden.

Exhibit 8 compares the estimated costs to respondents for 2009 to 2011 with the average annual cost estimated for FY2006 to FY2008. The annual costs have increased slightly since FY2006 to FY2008, primarily because of inflation.

A.16 Time Schedule, Publication, and Analysis Plans

The major components of the O*NET Data Collection Program include sampling, data collection operations, and analysis. Exhibit 9 provides the expected schedule for annual data analysis cycles and data publications for the next several years.

A.16.1 Data Analysis Tasks Conducted for Each Cycle

Described here is the approach used for data cleaning and editing, as well as the analyses that are performed annually.

Data Cleaning

Paper questionnaires are manually and machine edited so that completely blank questionnaires are dropped; responses to items that should have been skipped are blanked out; multiple responses are blanked out; and codes indicating missing data, multiple responses, and legitimate skips are inserted. Codes for legitimate skips and missing responses are also inserted in the records for questionnaires obtained through the Web. In addition, an electronic check is conducted to detect duplicate questionnaires from the same respondent.

Exhibit 8. Hour and Cost Burden for FY2006–FY2008 Compared with Estimated Hour and Cost Burden for 2009–2011

	Annual Average FY2006–FY2008[a]		2009		2010		2011	
	Sampling Units	Burden Hours	Sampling Units	Burden Hours	Sampling Units	Burden Hours	Sampling Units	Burden Hours
Establishment Activity								
Verification calls to initial contact at establishment	18,667	622	16,083	536	14,667	489	13,833	461
Screening call to POC	15,493	775	13,349	667	12,173	609	11,482	574
Initial recruitment call to POC	8,986	1,797	7,743	1,549	7,061	1,412	6,659	1,332
POC creates occupation lists for sampling	5,032	1,677	4,336	1,445	3,954	1,318	3,729	1,243
Call to POC to sample workers	5,032	839	4,336	723	3,954	659	3,729	622
POC distributes Questionnaire Packets	4,730	1,182	4,076	1,019	3,717	929	3,505	876
Follow-up calls to POC	4,730	631	4,076	543	3,717	496	3,505	467
Total establishment [a]	NA	7,523	NA	6,482	NA	5,912	NA	5,575
Employee Activity								
Establishment Method employee respondents	14,591	7,296	16,321	8,161	16,525	8,262	14,430	7,215
Occupation Expert Method respondents	713	1,427	500	750	500	750	500	750
Total, employee	NA	8,723	NA	8,911	NA	9,012	NA	7,965
Total, Establishments and Employees								
Total		16,246		15,393		14,924		13,541
Cost Burden								
Respondent Type	Annual Average FY2006–FY2008[a]		2009		2010		2011	
Establishments[b]	$255,590		$304,654		$277,817		$262,025	
Employees[c]	$159,980		$175,814		$177,807		$157,149	
Total	$415,570		$480,468		$455,624		$419,174	

[a] From the OMB Supporting Statement dated September 2, 2005.

[b] Assumed hourly rates: $33.97 for 2006–2008 and $47.00 for 2009–2011.

[c] Assumed hourly rates: $18.34 for 2006–2008 and $19.73 for 2009–2011.

Exhibit 9. Data Analysis and Publication Schedule

Analysis Cycle	Analysis Cycle Start Date	Analysis Cycle End Date	Estimated Establishment Method O*NET-SOC Occupations	Estimated Occupation Expert Method O*NET-SOC Occupations	Estimated O*NET-SOC Occupations Published	Publication Date
10	Jul-08	Jun-09	75	25	100	Jun-09
11	Jul-09	Jun-10	75	25	100	Jun-10
12	Jul-10	Jun-11	75	25	100	Jun-11
13	Jul-11	Jun-12	75	25	100	Jun-12

Identification and Evaluation of Anomalous Cases

Analyses of the ratings are based on the assumption that raters were qualified, willing, and able to engage in the rating tasks. This task is accomplished by requiring each case to first pass through a series of machine edits using prescribed eligibility criteria, including having at least one task rated important and having at least 50% of the domain questionnaire items completed. Cases not meeting these criteria are excluded from the analysis file. Cases with certain questionable characteristics are flagged in this editing process, and analysts review these cases to determine their completion status. Flagged for review are all cases for which the respondent (1) indicated in the "global match" item that the target O*NET-SOC occupation description did not at all describe his or her own job, and (2) rated fewer than one third of the tasks as important. Analysts review the self-reported job titles of these cases to determine if they are at all likely to belong in the O*NET-SOC occupation. If a case does not belong, it is removed from further analysis. If, in the judgment of the analysts, there is a reasonable chance that the case belongs in the O*NET-SOC occupation, it is sent to the next stage of review.

Finally, cases that pass the machine edits and the analyst review are subjected to a deviance analysis designed to identify cases that are outliers relative to other cases in their occupation. The deviance analysis involves two procedures: (1) a statistical procedure to quantitatively identify potential outliers and (2) an analyst review of these potential outliers to make the final decision for each case. Cases that do not pass the analyst review are deemed deviant within their occupation and are removed from the data set. On average, these activities eliminate about 10% of all returned questionnaires. The cases passing all data cleaning criteria are used to create the estimates for publication.

Computation of Sampling Statistics

Basic sampling weights are applied to the data to make inferences to the population of incumbents for each occupation.[19] These weights are computed as the inverses of the overall selection frequencies and the selection probabilities for each selected establishment and each individual participant. The analysis weights for the eligible sample units are adjusted to compensate for unit nonresponse for both establishments and employees, multiple subwaves of sampling, and sample adjustments. In addition, when the variation in the weights is large, the weights are trimmed to reduce the variation. In order to maximize comparability of O*NET estimates to estimates from other federal sources, the final sample weights are also ratio-adjusted to occupation estimates obtained from the OES survey.

---

[19] For details about the weighting and estimation procedures, see Section B.1.1.

Sampling errors are computed. The analysis weights used in the sampling error computations, as noted, have been adjusted for nonresponse and are consistent with the complex sampling design.

Calculation of Descriptor Values and Reliability

For each occupation, the sample size, mean, standard deviation, and standard error of the ratings for each descriptor are calculated, together with the 95% confidence interval around the mean. Estimates with questionable precision are flagged and recommended for suppression if any of the following conditions is true:[20]

- The sample size (i.e., number of respondents who answered the question) is less than 10.
- The variance is zero and the sample size is less than 15.
- The relative standard error (RSE) is greater than 50%.[21]

In addition, estimates of sensitive data items with a cell count (i.e., the number of respondents who selected a specific response option) of less than 10 are suppressed to ensure that O*NET data pose no disclosure risk to any individual respondent.[22] On average, approximately 4% of the estimates are flagged or suppressed.

Interrater Reliability and Agreement

Interrater reliability and agreement are assessed with three different analyses. The first two measure reliability in terms of the covariation among ratings, while the third analysis reflects rater agreement. The intraclass correlation coefficients (ICCs) for each item across all occupations are computed. These results allow one to compare respondent rating variance within an occupation with respondent rating variance across occupations. In addition to the ICCs, the mean interrater correlations (Pearson's *r*) are calculated for all pairs of raters within each occupation. Finally, to assess the absolute difference among ratings of each item within each occupation, an average deviation index is calculated for each item within each occupation. For any given item and occupation, the average deviation index measures the average extent to which each individual rating deviates from the item mean. Some differences in ratings within occupations are expected because O*NET-SOC occupations comprise a range of different jobs in most cases.

---

[20] O*NET suppression criteria are based on substantive expert recommendations (Peterson, Mumford, Levin, Green, & Waksberg, 1997) and on other large government surveys (see Klein, Proctor, Boudreault, & Tuczyn, 2002).

[21] RSE = the standard error of the mean divided by the mean.

[22] A sensitive data item is one that represents a value of a survey variable that an individual would not like to be revealed and provides an unacceptable risk that an individual in the cell could be identified.

Each of the reliability analyses conducted (rater, standard errors) are influenced by the number of respondents. The O*NET data collection methods include a sufficient number of respondents in each occupation to ensure reliability (Peterson, Mumford, Levin et al., 1997).

A.16.2 Creation of the Occupation Database

The O*NET database is scheduled to be updated annually. Each update will include data for those occupations collected and analyzed during the previous 12-month period. Consequently, a database update includes occupations from multiple data collection waves, depending on the number of prior waves for which analysis was completed that year. For each occupation collected, the newly calculated means data replace existing analyst-based or incumbent-based data in the database. Provided to users are metadata regarding when the data were collected and any other pertinent information that will assist the users in interpreting the data.

The O*NET database is designated with a version number denoting each update (e.g., from O*NET 4.0 to O*NET 5.0). The database is developed and administered with the MySQL database management system. Once the MySQL database is updated, it is used to generate the database for public release as a series of flat text files. It is accessible to the public on the O*NET Program Web site at http://www.onetcenter.org/.

A.17 Display of Expiration Date

The expiration date will be displayed on the cover of the survey questionnaires.

A.18 Exceptions to Certification Statement

There are no exceptions.

B. Collection of Information Employing Statistical Methods

B.1 Sampling Universe, Sampling Methods, and Expected Response Rates

A multiple-method data collection approach for creating and updating the O*NET database has been developed to maximize the information for each occupation while minimizing data collection costs. The primary source of information for the database is a survey of establishments and sampled workers from within selected establishments, which is referred to as the *Establishment Method* of data collection. With its two-stage sample design, employees are sampled in their workplace, the establishments being selected in the primary stage and employees being selected in the secondary stage.

Although the Establishment Method provides the best approach for most occupations, a special frame (e.g., a professional association membership list) is sometimes used to supplement the Establishment Method in a dual-frame approach when additional observations are required. When this supplementation to the Establishment Method is used, a dual-frame adjustment is made to the sampling weights to account for the coverage overlap between the two sources of collected data.

A second method entails recruitment of appropriate occupation experts, who can supply the information required for an occupation (*Occupation Expert Method,* or *OE Method*). An occupation expert is someone who has worked in the occupation for at least one year and has 5 years of experience, as either an incumbent, trainer, or supervisor. Additionally, an occupation expert must have had experience with the occupation within the most recent 6 months. The OE Method is used for occupations as necessary to improve sampling efficiency and avoid excessive use of burden, as when it is difficult to locate industries or establishments with occupation incumbents, employment is low, or employment data are not available, as is the case for many new and emerging occupations.

B.1.1 Establishment Method

Establishment Method Sampling Universe

The central goal of the O*NET Data Collection Program is to provide data for each of the O*NET-SOC occupations prevalent to varying degrees in different industries in the United

States. Estimates from this program are designed to assist users in distinguishing among occupations and are not necessarily designed to capture all of the subtle differences between jobs in different industries. Nonetheless, the O*NET sampling universe for each occupation is generally a subset of all employees in the occupation who are working in the United States. This subset, or target population for the occupation, is defined by two criteria: (1) its workers represent a majority of job incumbents in the occupation, and (2) data among this set of establishments can be gathered with reasonable efficiency.

Previous O*NET experience has shown that trying to build a sampling frame that covers 100% of an occupation is inefficient and poses undue burden on some establishments. For example, the occupation-by-industry matrix data suggested that a very small number of bricklayers could be found in establishments in the hospital industry; however, asking a point of contact (POC) in a hospital about bricklayers led to some difficulties. In addition to the questioning's being unduly burdensome, often the Business Liaison (BL) lost credibility when a POC was asked about occupations not likely associated with his or her establishment, as with bricklayers in hospitals. Moreover, the establishment POC may give some false negative responses because the POC simply does not know whether some rare occupations exist in his or her establishment. This situation would be particularly likely for larger establishments. To address these concerns, the target population is defined so that it includes establishments in industries and size categories in which the occupation is most prevalent.

When less than complete population coverage is allowed, it is possible that some bias may be introduced into the study estimates if the covered and noncovered population members would give substantially different responses to the survey questions. To evaluate this potential bias in the O*NET estimates, a sensitivity study was conducted. The study considered 18 randomly selected occupations for which at least 80% population coverage had been achieved. The linkages of these 18 occupations to industries were then reconsidered, and reduced sets of industries were determined that covered only 50% of workers in each occupation. Estimates for a selected set of outcomes were then computed from the reduced data set, which simulated estimates at the 50% level of population coverage. When the original data with at least 80% coverage was compared with the reduced data with 50% coverage, no systematic differences in the estimates were observed. The great majority of the differences between the two sets of estimates were very small and symmetrically distributed around zero. The observed pattern could be explained by random sampling error alone and provided no evidence of bias due to reduced frame coverage. The study concluded that no systematic bias is introduced with the use of a population coverage minimum of 50% for each occupation as compared with a minimum coverage of 80%. On the basis of these results, O*NET Establishment Method data collection now maintains a population coverage of at least 50% for each occupation.

Sampling Waves

To help identify industries in which particular occupations are employed, the O*NET sampling method uses employment statistics published by the U.S. Bureau of Labor Statistics (BLS) and supplemented by empirical information learned during O*NET data collection. Groups of approximately 50 occupations each, called *primary waves,* are formed so that the occupations in a primary wave are employed in a similar set of industries. For example, carpenters and plumbers are both employed by establishments in construction-related industries. Thus, when establishments are selected from the industries associated with a primary wave of occupations, a selected establishment is much more likely to employ one or more of the occupations in the wave than it would have been to employ one or more occupations not grouped this way. This method minimizes the number of establishments that must be contacted for selection of the required number of employees for an occupation. For example, when construction trades, such as those of carpenters and plumbers, are grouped together in a primary wave of occupations, it is much more likely that an establishment selected from construction-related industries will employ at least one of the 50 related occupations in the wave than would be the case if sampling had been from a broader set of industries associated with a group of unrelated occupations.

Each primary wave of occupations is scheduled to be fielded in three subwaves of establishment samples. The subwaves are identified as X.1, X.2, and X.3, where X represents the set of primary occupations and where the accompanying number represents the order in which the subwaves of establishment samples occur. For example, Subwave 3.1 denotes the first sample of establishments for the occupation set known as *Wave 3,* and 3.3 denotes the third sample of establishments for the occupation set. Any occupation that requires additional respondents is included in the next subwave. The first subwave of establishments uses the Occupational Employment Statistics (OES) data to indicate those industries most likely to employ the occupations. It is designed to include a wide range of industries and to cover at least 50% of the target population. As each subwave establishment sample is selected, the experience gained from the previous subwaves is used to more effectively target the sample to industries in which the occupations have been demonstrated to be found.

If, after being fielded in its X.3 subwave, an occupation lacks a sufficient number of completed respondents, then it is fielded in a completion wave. Completion waves combine the difficult-to-complete occupations from several waves and are designed to target industries with a high probability of employing the occupations. In a completion wave, only industries that the O*NET team's experience suggests might contain the occupation of interest are sampled. The goal of a completion wave is to ensure that the number of establishments selected for each occupation is sufficient to complete all occupations in the wave. Statistically, a completion wave

is no different from the X.1, X.2, and X.3 subwave sampling process, with the same sampling, weighting, and estimation methods being used to conduct the completion wave. Essentially, a completion wave adds a fourth subwave of sampling for some difficult-to-complete occupations. Packaging together some of these occupations in a combined wave maintains operational efficiency.

Sampling steps are carried out for the *primary* occupations associated with a wave. The primary occupations are those occupations selected for a wave as a result of the clustering of occupations likely to be employed in the same industries. Once the sets of industries to be targeted in a wave are identified, additional *secondary* occupations likely to be found in these industries are added to the wave and allocated to the selected establishments. To improve efficiency, if a selected establishment employs fewer than the maximum number of allowed primary occupations, secondary occupations are included for that establishment.

This method yields two major benefits. First, by prescribing sampling from industries that have been determined by empirical data to employ the occupation of interest, it maximizes the efficiency of the sample. Second, it minimizes the oversampling of any one occupation. Because the establishment sample size for a particular set of occupations is spread over three subwaves, if an occupation is found more easily than expected, the sample for future subwaves can be used to find other occupations in the wave rather than on this particular occupation.

In order to minimize both the cost of conducting the O*NET Data Collection Program and the burden placed on the establishments, the number of employees selected into the sample and the number of returned questionnaires are monitored carefully on a daily basis. Once it becomes clear that at least 15 of the goal of 20 completed respondents will be available for each of the three domain questionnaires for an occupation, consideration is given to terminating further sampling of employees for that occupation. This step is taken because of the difficulty of estimating the rate at which employees will be encountered during employee sampling. In some cases, employees from an occupation are much easier to locate than anticipated, and the desired number of responding employees is quickly exceeded; continuing to sample employees for such an occupation would use resources inefficiently and burden the establishments unnecessarily.

The method used to control employee sample selection is called *Model-Aided Sampling* (MAS). With this method, target numbers of sampled employees are defined, before data collection begins for each occupation, by Census region, business size, and industry division. MAS ensures that the resulting sample of employees is distributed across the target cells approximately in proportion to the population distribution across the target cells. MAS sample size targets are based on information from the OES survey conducted by BLS and from establishment information provided by Dun and Bradstreet (D&B). Once data collection begins,

progress toward the targets is monitored closely for each occupation on a daily basis. Once a cell's MAS target is achieved, the selection of employees in that cell is stopped. This cessation allows future data collection subwaves to focus on the cells that have not yet reached the MAS targets.

The use of MAS to enhance the efficiency of the O*NET sample is considered one of the most important sample design innovations in the history of the program. This approach has dramatically reduced data collection costs, with minimal effects on the accuracy of the estimators. Development of the MAS methodology began in early 2004 and continued through the end of 2006. As part of this research effort, various sample cutoff or stopping rules were investigated by means of Monte Carlo simulation. A cutoff rule determines the point at which efforts to interview certain types of establishments and employees are discontinued because the prespecified cell sample size criteria have been satisfied. These studies showed that, under a fairly wide range of cutoff rules, MAS does not bias the O*NET estimates or their standard errors. At the same time, MAS dramatically reduces the number of establishment contacts required in order to satisfy the random sample allocation for an occupation. This finding has resulted in substantial reductions in respondent burden, data collection costs, and time required to complete an occupation (Berzofsky, Welch, Williams, & Biemer, 2006). For further details of this MAS work, see Appendix G.

Sampling Steps for the Primary Occupations Associated with a Wave

As mentioned, the Establishment Method involves multiple sample selection steps. Establishments are selected during the first steps of selection and employees are selected during the later steps. This sample selection process is diagrammed in Exhibit 10 and is further detailed in text.

Step 1: Create establishment sampling frame. Two major sources of information are used to create the establishment sampling frame. First, a list covering nearly 12 million establishments in the United States is constructed from the D&B list of U.S. establishment locations. Use of the D&B frame is based on an evaluation that showed it to be the least costly data source that had the essential establishment-level information, including industry type and location-specific employment. The frame is updated quarterly to ensure that the most current and accurate establishment information possible is used for selecting the sample for each subwave. Additional information from the OES survey, conducted by BLS, is merged with the D&B list of establishments. The OES provides current industry employment estimates by occupation, using North American Industry Classification System (NAICS) industry codes, whereas the D&B list uses Standard Industrial Classification (SIC) codes. To complete the merge, the OES data are translated to SIC codes by means of a translation matrix available from BLS. SIC industry codes

Exhibit 10.   Summary of Sample Selection Process

```
┌─────────────────────────────────────────────────────┐
│   Step 1: Create Establishment Sampling Frame.        │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 2: Determine Industries to Target for Each     │
│                Occupation in Subwave.                 │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 3: Select Initial Sample of Establishments from│
│                        Frame.                         │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 4: Select Final Set of Establishments for Subwave.│
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 5: Assign Occupations to Selected Establishments.│
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 6: Specify Employee Sample Selection Algorithm │
│           for Each Occupation/Establishment.          │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 7: Specify Occupation-Specific Algorithm for   │
│   Random Assignment of Questionnaire Types of Sampled │
│                       Employees.                      │
└─────────────────────────────────────────────────────┘
                          │
                          ▼
┌─────────────────────────────────────────────────────┐
│   Step 8: Randomly Assign Selected Establishments to  │
│          Business Liaisons for Data Collection.       │
└─────────────────────────────────────────────────────┘
```

are used for the rest of the sampling steps, when the D&B list is used. NAICS industry codes are used again, however, during subsequent analyses of the data.

From the combined file a matrix is created for each occupation; the matrix lists the industries in which the occupation is found and, for each industry, the number of employees and associated establishments by Census region and by each of four size categories.[23]

---

[23] The establishment size category has four levels based on the number of employees working at an establishment. These levels are unknown number of employees (*none*), 1–9 employees, 24–249 employees, and 250 or more employees.

Step 2: Determine industries to target for each occupation in subwave. Using the matrix developed in Step 1, the industry–region–employee-size categories for each occupation are classified into one of four concentration groups: *high*, *medium*, *low*, or *none*. This classification helps to target industries in which an occupation is likely to be found. For example, if it is believed that establishments in a particular industry or industries have a good chance of containing a particular occupation, then the industry is classified as *high*. Similarly, industries with a low chance of having an occupation are classified as *low*. To increase efficiency in the sampling process, the sample is designed to oversample the *high* industries and undersample the *low* industries for each occupation. *None* denotes those industries that were not expected to contain any employees in the occupation of interest or that had a negligible proportion of the employees.

Step 3: Select initial sample of establishments from frame. First, for each industry, the number of establishments to be selected is determined. However, because only limited access to some D&B list data fields is possible before sample selection of establishments, a stratified simple random sample of establishments is first selected from the list, with a sample size larger than the number that will ultimately be contacted. For this large simple random sample of establishments, the O*NET team has access to the full D&B establishment information that will be used in Step 4 of selection. Within each stratum, the selection probability for the ith selected establishment is

$$p_{1i} = n_1 / N_1,$$ (1)

where n_1 and N_1 are the number of establishments selected and number of establishments in the population, respectively. The associated sampling weight for this step is

$$w_{1i} = 1 / p_{1i}.$$ (2)

Step 4: Select final set of establishments for subwave. A subsample of establishments is selected with probability proportionate to a composite size measure (Folsom, Potter, & Williams, 1987) from the simple random sample of establishments selected in Step 3. If a single occupation were being studied, then common sampling practice would be to select a sample of establishments with probabilities proportional to the number of employees working in the occupation at each establishment. Because several occupations are to be sampled from the selected establishments, a cross-occupation composite size measure is used to lessen the variation in selection probabilities and in the final analysis weights. The composite size measure accounts for the estimated number of employees in the occupations of interest within an industry

group, as well as for the overall sampling rates. The composite size measure for i^{th} establishment selected in Step 3 is

$$S_i = w_{1i} \sum_j f_j M_{ij} ,$$

(3)

where the summation is over the occupations, (j), $f_j = m_j / M_j$ is the overall sampling fraction for the jth occupation, m_j is the desired national sample size of employees from the jth occupation, $M_j = \sum_i w_{1i} M_{ij}$ is the estimated total number of employees in the jth occupation on the frame, and M_{ij} is the estimated number of employees in the jth occupation from the ith establishment.

For each occupation, the sampling rate is generally greatest for those establishments classified in the *high* industry group for the occupation and successively lower for each of the remaining *medium*, *low*, and *none* industry groups. Once the composite size measure is calculated, all nongovernmental industries are sorted in descending order by industry so that similar industries are proximate to each other before they are split into four equal groups based on the composite size measure. These four groups form the nongovernmental industry strata. Governmental establishments are not classified into separate industries on the D&B frame. Consequently, all governmental establishments constitute a fifth industry stratum in the design.

In addition, some occupations are highly concentrated in only one or two industries. These industries are split from the previously defined industry strata groupings to form separate "special" industry strata. These special strata often prove valuable for particular occupations because otherwise a small number of establishments would be selected from these industries. Forming the special strata for occupation-valuable industries with comparatively small size measures ensures that a minimum number of establishments from the "special" industry strata are selected into the final sample. Establishments are further stratified by number of employees in the establishment and by Census region. Establishments with a large number of employees are oversampled, while establishments with few employees are undersampled. The degree of over- or undersampling varies for each subwave.

Chromy's (1979) *probability minimum replacement* (PMR) selection procedure is then used to select a probability-proportional-to-size sample of establishments within each stratum, with probabilities of selection

$$p_{2i} = n_2 S_i / S_+ ,$$

(4)

where n_2 is the number of establishments selected from the stratum and S_+ is the sum of the composite size measures for establishments from the stratum. The sampling weight associated with this step is

$$w_{2i} = 1/p_{2i}.$$

(5)

Step 5: Assign occupations to selected establishments. To limit the burden for a particular establishment, each establishment selected in Step 4 is assigned a maximum of 10 occupations randomly selected with probability proportional to size. Here the size measure is the product of the sampling rate for the occupation (f_j) and the establishment's estimated number of employees within the occupation, M_{ij}. Before selection of a sample of the occupations, occupations certain to be selected because of their large sizes are included in the sample and removed from the frame, and the number of times they would have been "hit" (which, by the probability minimum replacement [PMR] method, can exceed 1) is recorded. Then the remaining (noncertainty) occupations are sampled and placed in a random order. The certainty occupations are listed first, followed by the randomly ordered noncertainty units. For each establishment, both the set of up to 10 occupations and the number of times each occupation was selected (which could be greater than 1 for certainty occupations) is entered into the Case Management System (CMS).

As before, we used Chromy's (1979) PMR selection method to select occupations with probability proportional to size. To understand how his method is applied here, one may suppose the ith establishment has J_i occupations associated with it. The size measure for the jth occupation is defined as

$$O_{ij} = f_j M_{ij},$$

so that

$$O_{i+} = \sum_j O_{ij} = \sum_j f_j M_{ij} = S_i / w_{1i}.$$

(6)

A sample of 10 occupations for each establishment will be selected with the expected number of times that the jth occupation is selected being $10 O_{ij}/O_{i+}$, which may be greater than 1 for some occupations. For an occupation j where it is greater than 1, the occupation is selected with certainty and assigned an O*NET point value equal to v_{ij} by randomly rounding $10 O_{ij}/O_{i+}$ to one of its two adjacent integers. That is,

$$v_{ij} = \text{Int}(10 O_{ij}/O_{i+}) \text{ with probability 1-Frac}(10 O_{ij}/O_{i+})$$

and

$$v_{ij} = \text{Int}(10 O_{ij}/O_{i+}) + 1 \text{ with probability Frac}(10 O_{ij}/O_{i+}),$$

(7)

where Int and Frac are the integer and fractional parts of a decimal number. This rounding provides an integer number of selections associated with each selected establishment while retaining the correct noninteger expected value for the number of selections. The certainty occupations appear at the top of the list used by BLs to inquire about occupations at the establishment. From among the remaining occupations, a probability-proportional-to-size sample is selected. If C_i is the number of certainty-occupation sampling hits from the ith establishment, $C_i = \sum_j v_{ij}$, summation over the certainty occupations, then the remaining occupations are selected with probabilities $(10 - C_i) O_{ij} / \sum O_{ij}$, summation over noncertainty occupations. The selected occupations are assigned an O*NET point value $v_{ij} = 1$. As noted previously, the selected noncertainty occupations are then placed in a random order and follow the certainty occupations on the list of occupations used by the BLs.

When a POC is identified within each establishment, the RTI BL reviews the list of occupations with the POC, asking the POC to estimate the number of employees at that establishment in each occupation. Each time the BL receives a response that is greater than zero, a counter within the CMS is increased by the associated O*NET point value, the randomly rounded number of times the occupation was selected. If the counter reaches 5 before the BL has completed the list of occupations, the BL stops. After the maximum 5 occupations are identified, the POC is asked to roster all individuals in the selected occupations.

To determine the final occupation selection probabilities, one must adjust for the occupations remaining on the establishment's sampling list at the point where the BL stopped as a result of having found the maximum number of occupations to be included in data collection. It is assumed that the resulting sample of up to 5 occupations is a random sample of the originally selected occupations for an establishment. This assumption is supported by the random ordering of the noncertainty occupations. Let a_i be the total number of sampling hits among all of the occupations about which the BL inquired before stopping; then, $a_i = \sum_j v_{ij}$, summation over the occupations inquired about by the BL. The final selection probability for the jth occupation from the ith establishment is

$$p_{3ij} = \frac{a_i}{10} \times \frac{10 O_{ij}}{O_{i+}} = \frac{a_i O_{ij}}{O_{i+}}. \tag{8}$$

The associated sampling weight is

$$w_{3ij} = 1 / p_{3ij}. \tag{9}$$

This method accomplishes two important goals:

- It results in an approximate random sample of occupations with known probabilities of selection.

- It limits POC and establishment burden. This goal is achieved because the number of positive POC responses is limited to a maximum of 5. If the company is large and happens to have employees in all 10 occupations, then stopping after 5 occupations minimizes the perceived burden on the POC, as opposed to the alternative of asking for employment estimates for all 10 occupations and then subselecting 5.

Step 6: Specify employee sample selection algorithm for each occupation/ establishment. In this step of selection, the algorithm is specified for randomly selecting employees from an employee roster provided by the POC. The resulting number of employees selected from each occupation is proportional to the number of times the occupation was selected in Step 5, its O*NET point value. However, to further minimize burden on an establishment, the total number of employees selected within any single establishment never exceeds 20, and the total number of employees selected in each occupation (within each establishment) never exceeds 8. If fewer than 20 employees are rostered and fewer than 8 are rostered for each occupation, then all rostered employees are selected. Otherwise, a random sample of employees is selected, subject to the constraints just described. If n_{ij} and N_{ij} are the number of employees selected and the number of employees listed, respectively, from the jth occupation at the ith establishment, then the selection probability for an employee k from the jth occupation is

$$p_{4ijk} = n_{ij} / N_{ij} \qquad (10)$$

and the associated sampling weight is

$$w_{4ijk} = 1 / p_{4ijk}. \qquad (11)$$

Step 7: Specify occupation-specific algorithm for random assignment of questionnaire types to sample employees. In this step, the algorithm is specified for assigning each selected employee to a domain questionnaire type. The survey is designed to collect data for each occupation from 20 respondents to each of three different domain questionnaires (Generalized Work Activities, Work Context, and Knowledge). At this step of selection, all employees selected in Step 6 are randomly assigned to one of the three questionnaire types. The questionnaire assignments are made in proportion to the number of employee respondents required for each questionnaire type in a subwave. In order to implement this algorithm, a queue is established that lists the order in which questionnaire types are assigned to employees from a specific occupation. The questionnaire types are listed in random order, with each type occurring in the queue at a rate proportional to the number of completed questionnaires required from that type. When an employee is selected, he or she is assigned the questionnaire type at the head of

the queue. The next listed questionnaire type is then moved to the top of the queue for use with the next selected employee.

Although employees are randomly assigned to a questionnaire, analysis weights are computed for each occupation, not separately for each of the three O*NET questionnaires used within each occupation. Many of the questions are the same across the three questionnaires; the estimates for these questions are produced from all the respondents for the occupation. Selected incumbents are assigned randomly, with equal probabilities, to each of the three questionnaires, with approximately the same number of incumbents responding to each of the questionnaires. Consequently, bias is not introduced for estimates of means and proportions specific to a single questionnaire. Producing the analysis weights for the entire occupation effects greater stability because the entire occupation offers a larger available sample size than each of the three individual questionnaires offers.

Step 8: Randomly assign selected establishments to Business Liaisons for data collection. The final step of sample selection randomly assigns selected establishments to a BL. In order to make this process more efficient and ensure that the BL workloads are equal, establishments are stratified by industry grouping, and assignments are made within strata. To do so, the Operations Center manager assigns up to three BLs to each stratum and indicates the number of establishments to be assigned to each BL. The specified number of establishments is then assigned to each BL. The only exception to the randomization process occurs when multiple establishments are sampled from the same company. In this situation, assignments are adjusted so that the establishments are assigned to the same BL. Although establishments are randomly assigned to BLs, this step does not enter into the overall selection probabilities or analysis weights: it is an administrative step to reduce any potential bias associated with the BLs.

The various weighting components associated with many of the Establishment Method sampling steps are combined to produce the final analysis weights, as will be described in the subsection on weighting.

Supplemental Frames

If the sample yield for an occupation proves to be far less than expected from the Establishment Method, the use of a special frame is considered for completing an occupation when additional Establishment Method subwaves would likely be nonproductive or inefficient. In this situation, if a suitable supplemental frame can be obtained, then the occupation is removed from the Establishment Method wave and sampled separately. Supplemental frames were used for 4% of the 780 O*NET-SOC occupations published to date. A supplemental frame may be either a listing of establishments highly likely to employ a particular occupation, or a listing of incumbents working in an occupation. For example, a trade association of business

establishments highly likely to employ an occupation would be appropriate when the occupation is highly concentrated in a particular type of establishment. In addition, some occupations' workers tend to be members of associations or are licensed. In such situations, it is often possible to obtain a frame listing either establishments or incumbents working in the occupation.

When a listing of establishments is obtained, a random sample of the establishments is selected from the frame. The sample usually will be stratified by geographic location and any other available establishment information that may be related to type of incumbents working in the occupation. Simple random samples are usually selected, because often little information is known about the number of incumbents the establishments employ. Consequently, if n and N are the number of establishments selected and the number on the frame, respectively, within a stratum, then the selection probability for the ith selected establishment is

$$p_{5i} = n/N, \tag{12}$$

and the associated weight is

$$w_{5i} = 1/p_{5i}. \tag{13}$$

The selected establishments are then included in Step 6 of the sampling process, with the single target occupation associated with each establishment.

On the other hand, when the supplemental frame lists incumbents working in an occupation, then a simple random sample of incumbents is selected, usually stratified by geographic location and, if available, subspecialty of the occupation. If n and N are the number of incumbents selected and the number on the frame, respectively, within a stratum, then the selection probability for the kth selected incumbent working in the jth occupation from the ith establishment is[24]

$$p_{6ijk} = n/N, \tag{14}$$

and the associated weight is

$$w_{6ijk} = 1/p_{6i}. \tag{15}$$

The selected incumbents are then directly contacted, and their participation is solicited. Those who agree to participate are entered into sampling process Step 7, by which a question type is assigned to each incumbent.

---

[24] Subscripts corresponding to occupation and establishment are added here for ease of notation when the supplemental samples are combined with the Establishment Method samples in the subsection on weighting.

The supplemental frame weights are adjusted for nonresponse and combined with the Establishment Method weights, as will be described in the subsection on weighting.

Employee Sample Size

A key issue in sample design is the level of precision required in the resulting data and the cost of producing a particular level of precision, in terms of both dollars and respondent burden. The O*NET sample design has been developed to provide results with a level of precision that should be adequate to meet the needs of general-purpose users (those seeking information at the occupational level). Consistent with the procedures used by the O*NET Program since 2001, an occupation is considered complete and ready for inclusion in the final O*NET database when a minimum of 15 valid completed questionnaires (after data cleaning) are obtained for each of the three questionnaire domains.

The current sample size goal is based on the final technical report of Peterson et al. (1997), which presents means and standard deviations for both 5- and 7-point Likert scales, with consecutive interger scores, for the descriptors within Skills, Knowledge, Generalized Work Activities, Abilities, and Work Styles. Statistics were computed separately with the reported data for each of six occupations. The data in these tables indicate that when 15 responses per descriptor are obtained, the mean values for virtually all of the 5-point and the 7-point descriptors have 95% confidence intervals that are no wider than plus or minus one to 1.5 scale points for all occupations.

Exhibit 11 displays the half-width of 95% confidence intervals (CIs) for means of 5- and 7-point scales, by sample size, from Analysis Cycles 4 through 8 of the O*NET Data Collection Program. The items are summarized in Exhibit 2 as those with a data source of job incumbents and are presented as part of the questionnaires in Appendix A. The scales were given consecutive integer scores, and estimates were produced as described in the "Estimation" subsection of B.1.1. Across all sample sizes, nearly all of the scale means have 95% CIs that are no wider than plus or minus 1.5 scale points. For those scale means based on sample sizes of between 15 and 25 respondents, more than 95% of the 5-point scales and more than 75% of the 7-point scales have 95% CIs no wider than plus or minus 1.5 scale points. In addition, 90% of the 7-point scales have 95% CIs no wider than plus or minus 1.6 scale points.

Furthermore, Mumford, Peterson, and Childs (1997, pp. 3-8) have cited Fleishman and Mumford (1991) as support that variation of one to 1.5 scale points on a 7-point scale "is typical of that found for well-developed level scales." Setting a minimum employee sample size of 15 (with many occupations achieving a larger sample size) therefore will generally satisfy this requirement. Additionally, Peterson et al. (1997) state that 15 to 30 incumbents typically provide

Exhibit 11.   Half-Width of 95% Confidence Intervals

Percentile	5-Point Scales		7-Point Scales	
	Sample Sizes of 15 to 25	All Sample Sizes	Sample Sizes of 15 to 25	All Sample Sizes
95th	+/− 1.2	+/− 1.1	+/− 1.8	+/− 1.7
90th	+/− 1.0	+/− 0.9	+/− 1.6	+/− 1.4
75th	+/− 0.8	+/− 0.7	+/− 1.3	+/− 1.1
50th	+/− 0.6	+/− 0.5	+/− 0.9	+/− 0.8

Note: Data are taken from 220 5-point scales and 109 7-point scales measured on each of 456 occupations.

sufficient interrater reliability for describing occupations, given the types of measures the O*NET Program uses to describe occupations.

Weighting

After the raw data are edited and cleaned, weights are constructed for each establishment and employee respondent to reduce estimate bias and variance due to factors such as nonresponse, undercoverage, and the complex sample design. The weighting process for the basic Establishment Method is described first. Subsequently, weighting for the supplemental-frame samples is described, together with weighting methods for combining the Establishment Method and supplemental-frame samples.

Estimates generated from O*NET survey data are computed with analysis weights to reflect the combined effects of the following:

- probabilities of establishment selection;

- probabilities of occupation selection;

- early termination of employee sampling activities for particular occupations, because of higher-than-expected yields;

- probabilities of employee selection;

- multiple-sample adjustments;

- nonresponse at both the establishment and the employee levels; and

- under- and overcoverage of the population, caused by frame omissions and undetected duplications.

The final employee-level analysis weights are computed as the product of a number of weight factors. These factors reflect the probabilities of selection from the multistage sample design, as already discussed, as well as appropriate adjustments for nonresponse, early sampling termination, population proportions, and multiple samples.

The starting point for each of these stages is the inverse of the probabilities of selection at each stage (establishment, occupation, and employee)—called the *base sampling weight* for the stage. The base sampling weight accounts for the unequal probabilities with which establishments, occupations, and employees are selected at each stage and are presented in

- Equation (2), w_{1i} associated with the initial simple random sample of establishments from the D&B frame;

- Equation (5), w_{2i} associated with the probability-proportional-to-size sample of establishments from the initial simple random sample;

- Equation (8), w_{3ij} associated with the selection of an occupation at an establishment; and

- Equation (11), w_{4ijk} associated with the selection of an employee from an occupation at an establishment.[25]

The product of these four weights would be the appropriate analysis weight if effects due to such issues as nonresponse and under- or over-coverage were negligible; however, weight adjustments likely will improve the accuracy of the estimates. The weight adjustments are implemented in three weighting steps corresponding to the three main steps of Establishment Method sampling:

- Weighting Step 1, by which establishment weights are computed to account for the probabilities of selecting establishments and to account for adjustments for establishment nonresponse;

- Weighting Step 2, by which occupation weights are computed to account for the probabilities of selecting specific occupations from each establishment and to account for adjustments for the early termination of sampling for some occupations under MAS; and

- Weighting Step 3, by which employee analysis weights are computed to account for the probabilities of selecting employees within each occupation and to account for adjustments for employee nonresponse and for overlapping frames across the subwaves.

The weights are calculated separately for each subwave in Weighting Steps 1 and 2, and then they are combined into an overall analysis weight in Weighting Step 3. The specific methods used in each of these weighting steps are described here after the unit-nonresponse adjustment method is described.

---

[25] As noted in conjunction with Establishment Method sampling Step 7, analysis weights are computed for each occupation, not separately for each of the three O*NET questionnaires used within each occupation. Many of the questions are the same across the three questionnaires; the estimates for these questions are produced with use of all the respondents for the occupation. Selected employees are assigned randomly, with equal probabilities, to each of the three questionnaires, with approximately the same number of employees responding to each of the questionnaires. Consequently, bias is not introduced for estimates of means and proportions specific to a single questionnaire. Producing the analysis weights for the entire occupation effects greater stability because the entire occupation offers a larger available sample size than each of the three individual questionnaires offers.

Nonresponse adjustment. The sampling weights are adjusted for nonresponse with use of a generalized exponential model (GEM). RTI has used the GEM method to create sampling weight adjustments for the 1999 through 2007 annual National Household Survey on Drug Use and Health conducted for the Substance Abuse and Mental Health Services Administration and for several other surveys conducted by RTI, including the 2000 National Postsecondary Student Aid Study and the 2000 Beginning Postsecondary Student Longitudinal Study, both sponsored by the U.S. Department of Education.

The GEM calibration is a generalization of the well-known weighting class approach, the iterative proportional fitting algorithm that is generally used for poststratification adjustments, Deville and Särndal's (1992) logit method, and Folsom and Witt's (1994) constrained logistic and exponential modeling approach. The GEM calibration process causes the weighted distribution of the respondents to match specified distributions simultaneously for all of the variables included in the model. One advantage of the GEM method over simpler weighting class or poststratification adjustments is that the adjustment model can use a larger and more diverse set of control variables because main-effects and lower-order interactions can be used in the model, rather than complete cross-classifications. For Folsom and Singh's (2000) description of the GEM method, see Appendix I.

To summarize, a set of predictor, or adjustment, variables are specified, together with the control total for each variable to which the weighted sample is expected to match. The GEM method is designed to determine a weight adjustment factor for each respondent, such that

$$\sum_k x_k w_k a_k = T_x ,$$

where the summation is over the respondents, x_k is an adjustment variable in the model, w_k is the base sampling (or unadjusted) weight, a_k is the adjustment factor, and T_x is the control total for the variable x. T_x may be either a nonresponse adjustment control total estimated by the sum of base sampling weights for both respondents and nonrespondents, or an external control total to adjust for under- or overcoverage of the frame. The adjustment factors, a_k, are determined to match the control totals for all of the variables in the model simultaneously. Furthermore, upper and lower bounds on the weight adjustment factors can be set to reduce the influence of observations that otherwise might have received a very large weight adjustment. The upper and lower bounds also reduce the effect of unequal weighting that may result from uncontrolled weight adjustments.

Weighting Step 1: Establishment weights. The base sampling weight, $w_i^{(1)}$, for the selected establishments in a subwave is the product of the weights in Equations (2) and (5):

$$w_i^{(1)} = w_{1i} \times w_{2i} \, . \tag{16}$$

The establishment sampling weights are adjusted for nonresponse, by subwave, with use of the GEM method with a model that contains different combinations of the following variables:

- industry division
- U.S. Census division
- establishment size
- headquarters/branch type
- number of occupations asked about in an establishment
- urban versus rural location
- time zone
- zip code information from the 2000 U.S. Census (quartile distribution of owner-occupied housing)

Variable selection proceeds by first fitting a model containing only main effects and tightening the upper and lower bounds so that all upper bounds are less than 8 and a minimal increase in the unequal weighting effect (UWE) is achieved. Two-way interactions among the variables are then added to the model. Cells that do not contain any respondents or that are collinear with other cells are removed from the model. If a convergent model cannot be obtained, some covariate levels are collapsed together; for example, U.S. Census divisions are collapsed to regions. Other variables or interactions may be removed from the model until a convergent model is obtained that maintains as many of the covariates and their two-way interactions as possible.

Variable selection and testing are conducted for each sampling subwave to determine the final model for a subwave. Finally, extremely large weights are trimmed back to smaller values. Even though the GEM method provides control over the size of the adjustment factors, it is still possible for large weights to result, though at a rate lower than that from an uncontrolled adjustment process. The total amount trimmed within a subwave is proportionally reallocated across the responding establishments to maintain the same estimated total number of establishments for each subwave. The adjusted establishment weights are denoted as $w_i^{(1a)}$.

Weighting Step 2: Occupation weights. The base occupation weight, $w_{ij}^{(2)}$, for ith occupation selected from the jth establishment is

$$w_{ij}^{(2)} = w_{3ij} \times w_i^{(1a)} \, , \tag{17}$$

which is the product of w_{3ij}, defined by Equation (8), times the adjusted establishment weight for the subwave defined in Weighting Step 2. For most occupations, no further nonresponse adjustments are necessary, because once an establishment agrees to participate, all of its selected occupations are available. However, MAS is used to terminate incumbent sampling early for some occupations with higher-than-expected numbers of sampled incumbents. For such occupations, the rate at which an occupation is encountered is estimated from the establishments contacted before the early termination; the estimated rate is then used to predict the number of additional establishments that would have reported employing the occupation. The occupation weights for the establishments that complete employee sampling are then adjusted to account for the predicted additional establishments. To understand the adjustment for early termination of sampling for some occupations, consider the classification of establishments shown in Exhibit 12.

Exhibit 12.   Classification of Establishments by Occupation Model-Aided Sampling Status

Group	Description
A	Inquired about occupation, and it is present at establishment
B	Inquired about occupation, but it is not present at establishment
C	Did not inquire about occupation because of early termination of incumbent sampling for occupation

Groups A and B are those establishments where the presence or absence of an occupation is known and can be used to estimate the rate at which the jth occupation is present, or the *presence rate,* by

$$ p_j = \frac{\sum_A w_{ij}^{(2)}}{\sum_A w_{ij}^{(2)} + \sum_B w_{ij}^{(2)}} , \tag{18} $$

where the summations are over the establishments in Group A or Group B. Next, the additional number of establishments where the jth occupation would have been found if sampling had not been terminated early is estimated by applying the presence rate to number of establishments in Group C. Thus, the estimated total number of establishments where the jth occupation would have been found is given by

$$ T_j = \sum_A w_{ij}^{(2)} + p_j \sum_C w_{ij}^{(2)} , \tag{19} $$

where the summations are over the establishments in Group A or Group C. It is tacitly assumed in Equation (18) that the establishments where occupations are inquired about approximate a random sample from all three groups listed in Exhibit 12. This assumption is consistent with the

random assignment of establishments to the BLs and the nondeterministic order in which establishments are contacted. The base occupation weights for the establishments in Group A are then adjusted to sum to T_j for the jth occupation.

To make this adjustment more sensitive, the process for estimating the number of establishments where the jth occupation would have been found is completed separately by Census regions, by the business size groups, and by industry divisions—as with the process for defining the MAS target cells. This process yields three sets of estimated marginal totals corresponding with the three variables. The GEM method is then used with a model containing the marginal, or main, effects of Census regions, business size groups, and industry divisions to adjust the base occupation weights from Equation (17) for those establishments in Group A. The adjusted weight for the jth occupation from the ith establishment is denoted by $w_{ij}^{(2a)}$.

Weighting Step 3: Employee analysis weights. The base weights for the responding employees in a subwave are

$$w_{ijk}^{(3)} = w_{4ijk} \times w_{ij}^{(2a)}, \tag{20}$$

which is the product of w_{4ijk}, defined by Equation (11), and the adjusted occupation weight for the subwave, defined in Weighting Step 2. At this point the responding employees from all subwaves are combined, and a multiple-frame adjustment is made. The overlap of target populations among the subwaves for each occupation is determined, and a multiple-frame adjustment is made, as described by Korn and Graubard (1999, sec. 8.2), using the sample sizes in the overlaps. For example, if two subwaves overlap for an occupation within a set of industries, then the adjustment factors for the subwaves are

$$a_l = t_l / (t_1 + t_2), \tag{21}$$

where t_l is the sample size from the lth subwave in the overlap between the subwaves. Then, the multiple-frame adjusted employee weights are

$$w_{ijk}^{(3^*)} = a_l \times w_{ijk}^{(3)}, \tag{22}$$

where the adjustment factor, a_l, is selected to correspond with the industry overlap portion associated with the ijk^{th} employee. This adjustment process is completed separately for each combination of overlapping subwaves.

Next, the employee weights in Equation (22) are further adjusted for nonresponse, using the GEM method with a model that contains different combinations of the following variables:

- industry division

- U.S. Census division

- establishment size

- headquarters/branch type

- number of occupations asked about in an establishment

- urban versus rural location

- time zone

- zip code information from the 2000 U.S. Census (quartile distribution of owner-occupied housing)

As before, variable selection and testing are conducted to determine the final model. Indicator variables for the occupations are included in the final model so that the adjustment maintains the correct sum of weights for each occupation; at the same time, to improve the adjustment, the data across occupations are used for the other variables.

At this point, an examination for extreme weights is conducted for each occupation by domain questionnaire. To prevent a few respondents from being too influential, weight trimming is performed according to two rules. Weights are deemed too extreme for a particular occupation by domain group if

- any weight exceeds the mean weight of the group plus 1.5 standard deviations of the weights, or

- a single weight accounts for more than 50% of the weight sum of the group.

Extreme weights are trimmed to smaller values; the total amount trimmed for an occupation by domain group is proportionally allocated to all respondents in the group.

The employee weights are then ratio-adjusted to match the distribution of workers in an occupation by industry sector as estimated from BLS's OES data. Occasionally, this final ratio adjustment will induce a large unequal weighting effect (UWE).[26] A final check is conducted to determine whether an occupation has a UWE greater than 10.0. If the UWE is greater than 10.0 for an occupation, then this final ratio adjustment is repeated; however, to bring the UWE below 10.0, the distribution of workers for this occupation is allowed to deviate from the OES estimated distribution. The resulting weights are denoted by $w_{ijk}^{(3a)}$. If a supplemental frame was not used to complete the sample of an occupation, then these weights are the final employee

---

[26] The UWE measures the increase in the variance of an estimate—an increase due to unequal weighting above the variance that a sample of the same size would yield if the weights were equal. The UWE is estimated by

$$n \sum_i w_i^2 \Big/ \left(\sum_i w_i \right)^2 .$$

analysis weights used to produce the estimates for the O*NET Data Collection Program. If a supplemental frame was used for an occupation, then an additional, supplemental-frame weighting step completes the calculation of the analysis weights.

Supplemental-frame weighting. As noted in the discussion of sampling methods, locating a sufficient number of respondents under the Establishment Method occasionally requires sampling from a supplemental frame. This situation usually occurs for rare occupations or occupations difficult to locate through Establishment Method sampling. When it does occur, additional weighting steps are necessary to account for the dual frames used to select the samples. Two situations must be considered: one in which the supplemental frame consists of establishments likely to employ the occupation of interest, and another in which the supplemental frame consists of employees likely to be working in the occupation. Described here are the steps taken in each situation.

First, consider the situation in which a supplemental frame of establishments is used, as illustrated in Exhibit 13. In this figure the dashed rectangle represents the full frame; the oval, the supplemental frame. The full frame, denoted by F, is the D&B establishment listing used for the usual Establishment Method sampling, which includes N establishments. The supplemental frame, denoted by F_S, includes N_S establishments.

Exhibit 13.   Overlap of Full and Supplemental Frames
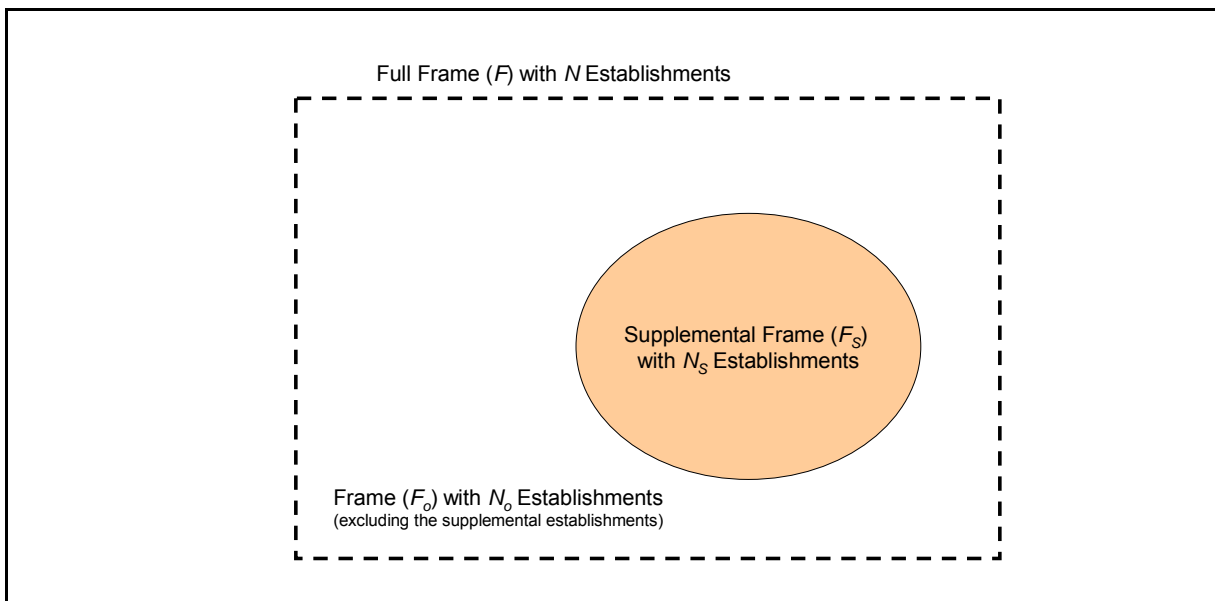
For a supplemental-frame sample of establishments, Equation (16) of Weighting Step 1 is modified to start with a base sampling weight, $w_i^{(1s)}$, given by

$$w_i^{(1s)} = w_{5i} \, , \tag{23}$$

where w_{5i} is defined in Equation (13). The supplemental-frame establishment sampling weights are then adjusted for nonresponse with use of the GEM method and under a model containing the stratification variables used to select the supplemental-frame sample. To determine the final model with the GEM software, variable selection and testing are conducted for each supplemental-frame sample. The nonresponse-adjusted supplemental-frame establishment weights are denoted as $w_i^{(1sa)}$.

A supplemental frame is developed to target a small number of related occupations—and often only one occupation—because of the difficulty of locating them. For a supplemental-frame sample, occupations are not randomly assigned to establishments from a large set of occupations as is done for the subwaves in Establishment Method sampling. Consequently, the base sampling weight for the jth occupation selected from the ith establishment for Weighting Step 2, $w_{ij}^{(2)}$ in Equation (17), is modified to be

$$w_{ij}^{(2s)} = w_i^{(1sa)} \, , \tag{24}$$

which excludes the weighting factor related to the random assignment of an occupation to an establishment. The subscript j is added, however, to recognize that a specific occupation is associated with each selected establishment in the supplemental-frame sample. The supplemental-frame sample then proceeds through the rest of Weighting Step 2. The adjusted weight for the jth occupation from the ith establishment is denoted by $w_{ij}^{(2sa)}$.

As part of Weighting Step 3, the base sampling weights for the responding employees in a supplemental-frame sample are defined as in Equation (20),

$$w_{ijk}^{(3s)} = w_{4ijk} \times w_{ij}^{(2sa)} \, , \tag{25}$$

which is the product of w_{4ijk}, defined in Equation (11), and the adjusted occupation weight for the supplemental-frame sample, defined in Weighting Step 2 as $w_{ij}^{(2sa)}$.

Next the occupation's employee weights from the supplemental-frame sample must be combined with the same occupation's weights from the Establishment Method sample. The employee weights from each of the Establishment Method subwaves are first combined as shown in Equation (22) into a single set of weights for all employees from an occupation selected by the Establishment Method. An extra step is added to Weighting Step 3 to combine the Establishment

Method employee weights from Equation (22) with the supplemental-frame employee weights from Equation (25).

At this point it must be assumed that the establishments listed on the supplemental frame, F_S in Exhibit 13, are equivalent to a random sample of the establishments on the full frame, F. This assumption is made because of the inability to determine which establishment on the full D&B-derived frame, F, links to which establishment on the supplemental frame. Under this assumption, a multiple-frame situation again emerges because data come from two samples: the Establishment Method sample and the supplemental-frame sample. A multiple-frame adjustment is made, as described by Korn and Graubard (1999, sec. 8.2), using the sample sizes from the two samples. Similarly, as was described for Equation (21), let t_F be the sample size for an occupation from the Establishment Method full sample, and let t_S be the sample size for the occupation from the supplemental sample. The multiple-frame adjustment factors are then given by

$$a_F = t_F / (t_F + t_S)$$

and

$$a_S = t_S / (t_F + t_S), \tag{26}$$

which correspond with the full frame and the supplemental frames, respectively. Finally, the weight for the kth employee from the jth occupation in the ith establishment is given by

$$w_{ijk}^{(3**)} = a_F w_{ijk}^{(3*)}$$

for the full sample and

$$w_{ijk}^{(3**)} = a_S w_{ijk}^{(3s)} \tag{27}$$

for the supplemental sample, where $w_{ijk}^{(3*)}$ and $w_{ijk}^{(3s)}$ are defined in Equations (22) and (25), respectively, and a_F and a_S correspond with the jth occupation. The weights $w_{ijk}^{(3**)}$ are then used in the adjustment steps after Equation (22) to complete Weighting Step 3, which yields the final analysis weights $w_{ijk}^{(3a)}$.

Next, consider the situation in which a supplemental frame consists not of establishments but of employees working in the occupation. In this situation, no actions equivalent to those in Weighting Steps 1 and 2 exist for the supplemental-frame sample. For Weighting Step 3, the supplemental-frame sample weighting starts with the base sampling weight

$$w_{ijk}^{(3s)} = w_{6ijk} \, , \tag{28}$$

where w_{6ijk} was defined in Equation (15). These weights are then combined with the Establishment Method employee weights for an occupation, as was described in connection with Equations (26) and (27), to yield $w_{ijk}^{(3**)}$. Again, the key assumption is made that the employees on the supplemental frame are equivalent to a random sample of all employees working in the occupation who might have been selected through the Establishment Method. This assumption is necessary because it is not possible to link the employees on the supplemental frame back to the D&B frame of establishments used for the Establishment Method. The weights $w_{ijk}^{(3**)}$ are then used in the adjustment steps after Equation (22) to complete Weighting Step 3, which yields the final analysis weights $w_{ijk}^{(3a)}$.

Estimation

The estimates produced for each occupation consist of scale means and percentage estimates. The number and type of scales are those listed in Exhibit 2, with a data source of job incumbents. Each of these scales is a 3-, 5-, or 7-point Likert scale. The final estimates are the means of the scales for each occupation. No subgroups or domain estimates are produced or released. The standard deviation will be available for each item mean as a measure of response variation among an occupation's respondents. Finally, there are several percentage estimates produced for items concerning work context, education and training, background items, and occupation tasks. Again, the final estimates are the percentages for each occupation, and no subgroup or domain estimates are produced or released.

For each item, if respondents do not provide an answer to a particular question, they are excluded from both the numerator and the denominator of the estimated mean. Because item nonresponse tends to be very low for this study (see Appendix H), no item imputation is conducted, and no value for missing items is assumed for estimation.

Variances are estimated with the first-order Taylor series approximation of deviations of estimates from their expected values. These design-based variance estimates are computed with SUDAAN software (RTI International, 2004). These estimates properly account for the combined effects of clustering, stratification, and unequal weighting—all of which are present in the O*NET data. The variance estimation clusters are the establishments; the stratification is by industry grouping and establishment size as used in selection of the establishment samples. These estimated variances are used to estimate both the standard errors associated with the mean or percentage and the confidence intervals. Standard error estimates and 95% confidence intervals are included with all estimates of means and proportions.

The estimate of a mean or a proportion is given by the formula

$$\hat{y} = \frac{\sum_{hik} w_{hik}^{(3a)} y_{hik}}{\sum_{hik} w_{hik}^{(3a)}}, \tag{29}$$

where $w_{hik}^{(3a)}$ is the final analysis weight for the kth respondent from the ith establishment in the hth stratum, and y_{hik} is the response variable. For a scale mean, the response variable, y_{hik}, is the scale value reported by the respondent; for a proportion, the response is one for a positive response and zero for a negative response. The Taylor series linearized values for the estimated mean or proportion are given by

$$z_{hik} = w_{hik}^{(3a)} \left(\frac{y_{hik} - \hat{y}}{\sum_{hik} w_{hik}^{(3a)}} \right). \tag{30}$$

The variance of $\hat{y}$ is estimated by

$$\text{var}(\hat{y}) = \sum_h \frac{n_h}{n_h - 1} \sum_i (z_{hi} - \bar{z}_h)^2, \tag{31}$$

where n_h is the number of variance estimation clusters from the hth stratum, $z_{hi} = \sum_k z_{hik}$ and

$\bar{z}_h = \sum_i z_{hi} / n_h$.

Expected Response Rates

Data collection had been completed for 76 subwaves as of September 30, 2007. These subwaves consisted of 153,981 sampled establishments and 200,942 selected employees. The overall response rate was 75% for establishments and 64% for employees. Although these response rates compare favorably with those of similar studies (see Section A.1.5), methods to further enhance response rates are continually being evaluated and implemented (see Section B.3).

B.1.2 Occupation Expert Method

The OE Method is used for occupations as necessary to improve sampling efficiency and avoid excessive burden, as when it is difficult to locate industries or establishments with occupation incumbents, employment is low, or employment data are not available, as is the case for many new and emerging occupations. To determine which sampling method should be used for an occupation, a comparison is made of the advantages and disadvantages of the Establishment and OE Methods. For each occupation, information on the predicted establishment eligibility rate and the predicted establishment and employee response rates is used to quantify the efficiency of sampling the occupation by means of the Establishment Method. The OE

Method is used for an occupation when the Establishment Method of data collection is not feasible and an appropriate source of occupation experts is available, as when a membership list of a professional or trade association exists and provides good coverage of the occupation. A random sample is selected from provided lists to prevent investigator bias in the final selection of occupation experts. Sample sizes are designed to ensure that at least 20 completed questionnaires are available for analysis after data cleaning. A goal of 20 questionnaires was set as a reasonable number to enable adequate coverage of experts, occupation subspecialties, and regional distribution.

Through September 30, 2007, the OE Method was used to collect data for 103 occupations. Of these, 82 were completed and 21 were still in process as of that date. A total of 3,845 occupation experts were sampled, of which 2,903 were found to be eligible. Of these, 2,375 participated, for an overall response rate of 82%.

For the OE method, unweighted estimates of the same means and percentages are reported as for the Establishment Method, together with the estimated standard deviation of the mean estimates.

B.2   Procedures for the Collection of Information

Data collection operations are conducted at the contractor's Operations Center in Raleigh, North Carolina, and its Survey Support Department in Research Triangle Park, North Carolina. For the Establishment Method, the Operations Center's BLs contact sample business establishments, secure the participation of a POC, and work with the POC to carry out data collection in the target occupations. The data are provided by randomly selected employees within the occupations of interest. All within-establishment data collection is coordinated by the POC; the BLs do not contact employees directly. After the POC agrees to participate, informational materials and questionnaires are mailed to the POC, who distributes the questionnaires to the sampled employees. Completed questionnaires are returned directly to the survey contractor for processing. The BLs contact occupation experts (the OE Method) directly, as well as job incumbents when sampling from a professional membership list in a dual frame approach; no POC is involved.

Survey Support Department staff mail materials to POCs, job incumbents, and occupation experts, and receive and process completed questionnaires that are returned by respondents. Both the telephone operations of the BLs and the mailing and questionnaire-receipt operations of the support staff are supported by the CMS. Data-entry software supports the keying and verification of incoming survey data.

B.2.1 Establishment Method

As described in Section B.1.1, the Establishment method uses a two-stage design involving a statistical sample of establishments expected to employ workers in the target occupations, followed by a sample of the workers in the target occupations within the sampled establishments. The sampled workers are asked to complete the survey questionnaires.

The Establishment Method works well for most occupations. Occasionally, however, the standard protocol is supplemented with a special frame, such as a professional association membership list, when additional observations are required to complete data collection for an occupation. The primary difference with this approach is that the supplemental respondents are sampled directly from the special frame and contacted directly by the BLs, without involvement of a sampled establishment or a POC.

*O*NET Operations Center*

Data collection activities are housed in the O*NET Operations Center, located in Raleigh, North Carolina, and covering 3,581 square feet. The Operations Center staff includes BLs, Team Leaders, a Monitoring Coordinator, and the Operations Center Manager, who reports to the Data Collection Task Leader. Usual operating hours for the Operations Center are Monday through Friday, 8:45 a.m. to 5:15 p.m., Eastern Time. Operating hours are extended during periods of unusually high workloads or when necessary to contact a high concentration of Pacific time zone businesses.

The BLs form the nucleus of the Operations Center staff. The number of BLs fluctuates somewhat, ranging from 9 to 13, depending on workload. New BLs are recruited and hired at various intervals in order to compensate for attrition and increases in workload. BL job candidates are carefully screened and evaluated by Operations Center management, who use a job description and a set of criteria that include a minimum of 2 years of work experience in a call center or related work experience in a human resources department.

Case Management System

The O*NET CMS is a Web-based control system that supports and monitors the data collection activities of the BLs, the mailing of informational materials and questionnaires, and the receipt of completed paper and Web questionnaires.

Questionnaires and Information Materials

The Establishment data collection protocol calls for each sampled worker to receive one of three randomly assigned domain questionnaires—Knowledge (which includes Education and Training as well as Work Styles), Generalized Work Activities, and Work Context. Each domain questionnaire also includes a Background section that asks a standard set of 11 demographic

questions about the respondent. In addition, each worker receives a Task Questionnaire specific to his or her occupation.

Task Questionnaires are developed initially through the extraction of task information from multiple sources located on the Internet. This questionnaire includes a definition of the occupation, a list of tasks, and space for the respondent to write in additional tasks. The respondent is instructed to indicate whether or not each task is relevant to his or her occupation and to rate each relevant task's frequency and importance. In subsequent updating efforts, task inventories are revised to reflect the new and most current information from respondents, including write-in tasks.

For all occupations, sampled workers also receive an occupation-specific Association Membership Questionnaire. The questionnaire provides a list of associations related to the worker's occupation and asks the respondent to indicate whether he or she belongs to any of them. The respondent is also asked to write in any other associations to which he or she belongs. This information is collected in case it becomes necessary to complete the occupation with use of the dual-frame approach.

Each sampled employee receives an integrated questionnaire consisting of the randomly assigned domain questionnaire and the Task and Association Membership Questionnaires applicable to the employee's occupation. Questionnaires are custom-printed on demand for each sampled worker. In addition, workers are encouraged to complete their questionnaire online at the project's Web site instead of completing and returning the paper questionnaire.

Spanish versions of the questionnaires are available for occupations with high proportions of Hispanic workers. The Spanish questionnaires are sent to workers who are identified as Spanish-speaking by their POC. In addition, an employee who has been sent an English questionnaire can request a Spanish version by calling the survey contractor with a toll-free number.

Examples of the English questionnaires are included in Appendix A. The Spanish versions are available on request.

In addition to the questionnaires, the Establishment Method data collection protocol includes a variety of letters, brochures, and other informational materials mailed to POCs and sampled workers. Spanish versions of the materials addressed to workers are available for occupations with high proportions of Hispanic workers. Appendix F contains examples of the English versions of these materials. The Spanish versions are available on request.

Data Collection Procedures: Establishment Method

Described here are the steps of the Establishment Method standard data collection protocol. A summary of this protocol is shown in Exhibit 14.

Exhibit 14.   Establishment Method Data Collection Flowchart

Step 1:
Make Verification Call to Receptionist

↓

Step 2:
Make Screening Call to the Point of Contact (POC)

↓

Step 3:
Send Advance Package

↓

Step 4:
Make Recruiting Call to POC

↓

Step 5:
Make Sampling Call to POC

↓

Step 6:
Send Questionnaire Package

↓

Step 7:
Send Toolkit

Step 8:
Make 7-Day Follow-up Call to POC

↓

Step 9:
Send Thank You/ Reminder Postcards

↓

Step 10:
Make 21-Day Follow-up Call to POC

↓

Step 11:
Make 31-Day Follow-up Call to POC

↓

Step 12:
Send Replacement Questionnaires

↓

Step 13:
Make 45-Day Follow-up Call to POC

Step 1: Verification call to the receptionist. The BLs call each sampled business to determine whether the business is eligible (i.e., whether it is still in operation at the sampled address). The other component of the verification call is to identify the anticipated POC, who must be knowledgeable about the types of jobs present in the establishment and to whom the screening call is placed.

Step 2: Screening call to the point of contact. The BLs next call (or are transferred to) the anticipated POC to ascertain whether the business has at least one employee in at least one of the occupations targeted for that establishment. If so, the following POC information is obtained:

- name and title of the POC,
- U.S. Postal Service delivery address,
- telephone number,
- e-mail address (if available), and
- fax number.

None of the BLs' conversations with the POC is scripted in advance. Instead, "talking points" are provided to guide the BLs' interactions with POCs. BLs are trained to listen and interact effectively and in a comfortable style, rather than to read from a prepared script; therefore, reading off a computer screen is discouraged. The BLs enter all information gathered during each conversation with a POC into the CMS.

Step 3: Send information package. The information package is sent to the POC after the completion of the screening call and contains more detailed information about the O*NET Program. The following information is included in the information package:

- DOL lead letter;
- O*NET brochure;
- "Who, What, and How" brochure;
- Selected Occupations List, providing title and descriptions of target occupations;
- list of endorsing professional associations;
- brochure describing the business-, POC-, and employee-level incentives; and
- POC incentive (i.e., the O*NET desk clock).

Step 4: Recruiting call to the point of contact. The next call to the POC is made approximately 7 days after the information package is shipped, in order to give the POC adequate time to receive, read, and process the information. During the recruiting call, the BL

- verifies that the information package was received;

- confirms that the POC is qualified to serve in the POC role;

- reviews with the POC the titles and descriptions from the Selected Occupations List for the target occupations, to determine whether the establishment has any employees in those occupations;

- (if one or more target occupations are present) explains the O*NET Program in greater detail, answers questions, and attempts to secure the POC's commitment to participate;

- for participating establishments explains the need for the POC to prepare a numbered list of employees' names for each selected occupation, for use in selecting a sample of employees; and

- sets an appointment for the sampling call, allowing sufficient time for the POC to compile the occupation rosters (in smaller businesses, the sampling call is sometimes combined with the recruiting call).

Step 5: Sampling call to the point of contact. During this call, the BL obtains from the POC the number of names on each roster and enters the counts into the CMS, which selects the sample according to preprogrammed random sampling algorithms. The BL then informs the POC which employees are selected for each occupation. The POC is asked to note for later reference the line numbers of the selected employees on his or her list (or lists) when the questionnaires are distributed. For designated O*NET-SOC occupations with a high percentage of Hispanic employees, the BL also asks the POC if any of the selected employees should receive a Spanish version of the questionnaire instead of the English version. The language preference of each employee is then indicated in the CMS.

Step 6: Send questionnaire package. After completion of the sampling call, the employee packets are shipped to the POC for subsequent distribution to the sampled employees. As part of the same mailing, the POC receives a thank you letter and a framed Certificate of Appreciation from DOL, personalized with the name of the POC and signed by a high-ranking DOL official. Each questionnaire packet contains a letter from the contractor's project director, the assigned questionnaire (including the domain questionnaire and the Task and Association Questionnaires integrated into a single booklet), a return envelope, an information sheet for completing the questionnaire, and a $10 cash incentive. A Spanish questionnaire is sent to any Hispanic employees who the POC indicated during the sampling call should receive this version. In addition, all employees in these O*NET-SOC occupations are informed through a bilingual notice included in the mailing that they have a choice between English and Spanish versions, and they are provided with a toll-free number to call if they would like to receive the alternate version.

Step 7: Send toolkit. Approximately 3 days after mailing the Questionnaire Package, the contractor also mails the POC the O*NET Toolkit for Business—a packet of information about

the O*NET Program, which managers can use for human resource planning and preparation of job descriptions.

Step 8: 7-day follow-up call to the point of contact. Approximately 7 days after the shipment of the original questionnaire package to the POC, the BL calls to verify receipt of the mailing and to review the process for distributing the questionnaires to the selected employees. The BL also informs the POC of a forthcoming shipment of thank you/reminder postcards and asks him or her to distribute them to all sampled employees.

Step 9: Send thank you/reminder postcards. After the 7-day follow-up call, the BL places an order for thank you/reminder postcards to be sent to the POC for distribution to all sampled employees.

Step 10: 21-day follow-up call to the point of contact. Approximately 21 days after the shipment of the original questionnaire package, the BL calls to thank the POC for his or her ongoing participation and to provide an update on any employee questionnaires received to date.

Step 11: 31-day follow-up call to the point of contact. Approximately 31 days after the shipment of the original questionnaire package to the POC, the BL calls to again thank the POC for his or her ongoing participation and to provide an update on any employee questionnaires received to date. At this time, the BL informs the POC of a forthcoming shipment of replacement questionnaires, which are to be distributed to any employees who have not yet returned the original questionnaire.

Step 12: Send replacement questionnaires. After the 31-day follow-up call, the BL places an order for the shipment of replacement questionnaires. These packages are ordered for any employees who have not yet responded. The replacement questionnaire package is like the original one, except for a slightly different cover letter and the absence of the $10 cash incentive. Using roster line information or employee initials provided by the BL during the 31-day follow-up call, the POC then distributes the appropriate replacement questionnaire package to each nonresponding employee.

Step 13: 45-day follow-up call to the point of contact. Approximately 45 days after the shipment of the original questionnaire package to the POC, the BL places one final follow-up call to the POC to thank the POC for his or her assistance and to provide one final status report regarding employee questionnaires. If all questionnaires have been received at this point, the BL thanks the POC for his or her organization's participation. If questionnaires are still outstanding,

the BL confirms receipt and distribution of the replacement questionnaire packets. This step is usually the final one in the data collection protocol.[27]

Mailout Operations, Questionnaire Receipt, and Processing

Orders for mailings of questionnaires and informational materials to support data collection are placed by the BLs and processed by data preparation staff. The CMS supports and monitors the entire process, including placing the order, printing on-demand questionnaires and other order-specific materials, shipping the order, and interacting with the U.S. Postal Service to track delivery of the order. Staff members follow written procedures in fulfilling orders, including prescribed quality control checks. They are also responsible for maintaining an adequate inventory of mailing materials and for inventory control.

Completed questionnaires returned by mail are delivered to the contractor, where they are opened and batched and the barcodes are scanned to update the CMS for receipt. The batches are then delivered to data-entry staff, where the survey data are keyed and 100% key verified. The questionnaire batches are then stored in a secure storage area. Data from the paper questionnaires are merged with the Web questionnaire data and readied for data cleaning routines.

Establishment Method Data Collection Results

Establishment data collection, which began in June 2001, is still under way. As of September 30, 2007, the contractor had completed data collection for 76 waves. The data collection results for these waves are summarized in Exhibit 15.

Exhibit 15.   Establishment Method Data Collection Results

Sampled establishments	153,981
Eligible establishments	127,083
Participating establishments	95,058
Establishment response rate	75%
Sampled employees	200,942
Participating employees	128,401
Employee response rate	64%

B.2.2 Occupation Expert Method

The Occupation Expert (OE) Method is an alternate method of collecting information on occupational characteristics and worker attributes that is used to improve sampling efficiency

---

[27] If no employee questionnaires have been received at the time of the last scheduled follow-up call, the case is referred to a Team Leader, who reviews the history notes for the case to try to determine if the POC actually distributed the questionnaires; if necessary and appropriate, the Team Leader will make an additional follow-up call to the POC.

and avoid excessive use of burden for problematic occupations. This situation occurs when occupations have low employment scattered among many industries or when employment data do not yet exist (e.g., for new and emerging occupations).With this method, persons considered experts in the target occupation, rather than job incumbents, are surveyed. Occupation experts are sampled from lists provided by source organizations (usually, professional associations) and are contacted directly by the BLs, without involvement of a sampled establishment or a POC. Unlike the standard Establishment Method, under which workers complete only one questionnaire, the OE Method requires that occupation experts complete all three domain questionnaires, as well as a Background Questionnaire and a Task Questionnaire. Because of the increased burden, occupation experts receive a $40 cash incentive instead of the $10 incentive offered to Establishment Method respondents.

The same facility used for Establishment Method data collection—the Operations Center in Raleigh, North Carolina—is also used for the OE Method work and uses a CMS, a Web-based control system that supports and monitors the data collection activities of the BLs, the mailing of informational materials and questionnaires, and the receipt of completed paper questionnaires.

Questionnaires and Information Materials

With the exception of a few additional items in the Background Questionnaire, the OE Method questionnaires are the same as those used for Establishment Method data collection. Occupation experts are asked to complete all three domain questionnaires (as well as a Background Questionnaire and Task Questionnaire), whereas Establishment Method respondents complete only one domain questionnaire (as well as Background, Task, and Association Questionnaires bound together with the domain questionnaire). Paper questionnaires are bundled prior to shipping, with the order of the domain questionnaires randomized at the respondent level. As with the Establishment Method, occupation experts are given the option of completing their questionnaires online at the project Web site.

OE Method information materials resemble the Establishment Method materials but are modified to reflect how the OE Method differs from the Establishment Method (direct contact with the respondent, identification through a named source organization, reference to only one occupation, multiple questionnaires, and a higher incentive).
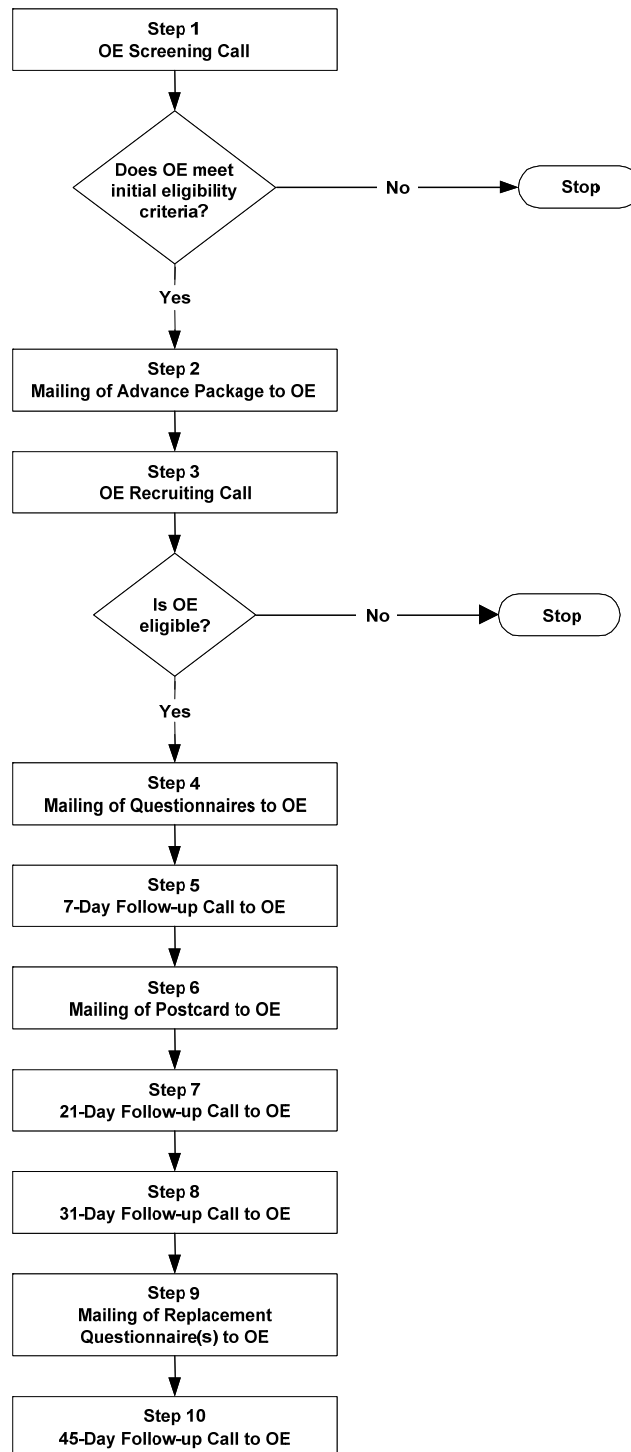
Examples of OE Method questionnaires are presented in Appendix A; information materials are presented in Appendix F.

Data Collection Procedures

The steps in the OE Method data collection protocol closely follow those for establishments. The primary differences are the absence of the verification and sampling calls.

Verification calls are inapplicable because a specific individual is contacted instead of an establishment. Sampling calls are inapplicable because the individual is not sampled from a larger group of employees. All other steps follow the Establishment Method protocol. The OE Method data collection protocol is shown in Exhibit 16.

Exhibit 16.   Occupation Expert Method Data Collection Flowchart

```
                    ┌──────────────────────────┐
                    │        Step 1            │
                    │    OE Screening Call     │
                    └──────────────────────────┘
                                │
                                ▼
                          ╱──────────╲
                         ╱  Does OE   ╲
                        ╱ meet initial ╲        No
                        ╲ eligibility  ╱──────────────────▶ ( Stop )
                         ╲ criteria?  ╱
                          ╲──────────╱
                                │ Yes
                                ▼
                    ┌──────────────────────────┐
                    │        Step 2            │
                    │ Mailing of Advance       │
                    │ Package to OE            │
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 3            │
                    │   OE Recruiting Call     │
                    └──────────────────────────┘
                                │
                                ▼
                          ╱──────────╲
                         ╱   Is OE     ╲        No
                         ╲  eligible?  ╱──────────────────▶ ( Stop )
                          ╲──────────╱
                                │ Yes
                                ▼
                    ┌──────────────────────────┐
                    │        Step 4            │
                    │ Mailing of Questionnaires │
                    │ to OE                    │
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 5            │
                    │ 7-Day Follow-up Call to OE│
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 6            │
                    │ Mailing of Postcard to OE │
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 7            │
                    │ 21-Day Follow-up Call to OE│
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 8            │
                    │ 31-Day Follow-up Call to OE│
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 9            │
                    │ Mailing of Replacement   │
                    │ Questionnaire(s) to OE   │
                    └──────────────────────────┘
                                │
                                ▼
                    ┌──────────────────────────┐
                    │        Step 10           │
                    │ 45-Day Follow-up Call to OE│
                    └──────────────────────────┘
```

Mailout Operations, Questionnaire Receipt, and Processing

OE Method mailing operations and questionnaire receipt and processing follow the same procedures as those described for the Establishment Method in Section B.2.1.

Occupation Expert Method Data Collection Results

Data collection was undertaken for a total of 103 occupations through September 30, 2007. Of these, 82 were completed, and 21 were still in process as of that date. The data collection results for the completed waves are summarized in Exhibit 17.

Exhibit 17.   Occupation Expert Method Data Collection Results

Sampled occupation experts	3,845
Eligible occupation experts	2,903
Participating occupation experts	2,375
Occupation expert response rate	82%

B.3   Methods to Maximize Response Rates

The O*NET Data Collection Program is committed to achieving the highest possible response rates through continuous improvement. Accordingly, the data collection procedures are continually evaluated in light of the experience gained in the data collection process. The previous OMB Supporting Statement dated September 2, 2005, listed a number of enhancements that had been made since the submission of the OMB Supporting Statement dated July 11, 2002. Continuing this approach, this section summarizes the principal enhancements to maximize response rates introduced since the 2005 Supporting Statement:[28]

- Supervisory review by the Operations Center Team Leaders was expanded to include all cases in which the POC had agreed to participate and workers were sampled but no questionnaires were completed by the time of the last scheduled follow-up call. The goal was to provide supervisor review and intercession, as appropriate, for de facto refusals, as well as for explicit refusals.

- The CMS and protocol were modified to enable BLs to make greater use of e-mail and fax communications with POCs, to supplement and facilitate the standard communication protocol.

- The training program for new BLs was enhanced to include an expanded module for answering POC questions and overcoming objections.

- Several small modifications were made to the OE Method data collection protocol to help encourage increased use of the Web questionnaire option by occupation experts. These adjustments included the addition of BL "talking points" to the CMS to remind

---

[28] For the full description of the data collection procedures, see Section B.2.

occupation experts of the Web option, an e-mail reminder to the occupation expert, revisions to informational materials, and a reminder on the project Web site.

- The project Web site (http://onet.rti.org) was expanded to include additional frequently asked questions and to provide a separate set for each type of survey participant (POCs, job incumbents, and occupation experts).

- A number of enhancements were made to the monitoring system used to monitor the BLs' telephone interactions with survey participants. These improvements enabled Team Leaders to provide better feedback and coaching to the BLs, resulting in enhanced overall performance.

- Skills-enhancement training programs were developed and implemented to help the BLs identify the best-qualified POC at a sampled establishment and address other problematic situations, such as instance when they are referred to a company's corporate headquarters or to an outside vendor that manages human resource matters for the sampled company.

- The CMS was modified to allow the BLs to distinguish between an initiated and a completed Web questionnaire. (Previously, initiated questionnaires appeared in the CMS as though they were completed.) This enhancement allows the BLs to give the POC more complete information about the status of data collection at the establishment, leading to better-informed follow-up activities.

- The Establishment Method protocol was modified to allow some flexibility regarding the requirement that POCs have a copy of the Selected Occupations List in front of them at the time of the recruiting call. (The Selected Occupations List is sent to the POC as part of the information package and provides a brief description of the occupations of interest for the establishment.) Although their having the Selected Occupations List in front of them is still the default requirement, experience indicated that it is likely impractical for some occupations, especially those for which the POC is often contacted away from an office setting (e.g., Rock Splitters, Quarry or Farm Labor Contractors). In these situations, the requirement that the POC have the Selected Occupations List physically in hand may be waived if the BL is confident that the POC has a clear understanding of the occupations of interest.

These and other enhancements have had a positive effect on the O*NET Program's ability to secure the participation of both establishments and employees. For example, the establishment response rate for data collection waves completed through September 30, 2007, was 75%, compared with 64% for the initial data collection wave completed during 2001–2002. Improvement in the employee response rate was less dramatic, increasing from 63% in the initial 2001–2002 wave to 64% for waves completed through September 30, 2007. However, the employee response rate has increased to 75% for waves completed during January–May 2008, reflecting several protocol enhancements implemented during 2007. The O*NET Program will continue to explore ways to enhance response rates still further through its continuous improvement program.

B.4   Tests of Procedures

Continuous improvement of survey quality and efficiency has been a constant focus of the O*NET Data Collection Program. The survey design specified in this document has evolved over years of testing and evaluating alternative procedures. The O*NET Program team believes it reflects the best combination of design parameters and features for maximizing data quality while reducing data collection costs. Summarized here are some of the tests of procedures that have been conducted for the O*NET Data Collection Program.

B.4.1 1999 Pretest Experiments

The initial design of the O*NET Program was based on rigorous testing and evaluation of alternative design protocols. Tests of seven different cash incentive plans were carried out between June 1999 and January 2000 on a sample of about 2,500 eligible businesses and 3,800 employees. In addition, the use of stamped return envelopes and various options for contents of the Toolkit for Business incentive were tested. These tests found that the best design appeared to be the combination of the $10 prepaid incentive to employees, various material incentives to the POC, and the use of first-class stamped return envelopes. These experiments also determined that a videotape describing the O*NET database should not be included in the Toolkit for Business. A report documenting the pretest activity and results was included in the 2002 O*NET OMB submission and can be found at http://www.onetcenter.org/ombclearance.html.

B.4.2 Wave 1.1 Experiment

A split-sample, randomized assignment experiment was conducted in Wave 1.1 to compare the benefits of using business-reply envelopes with the benefits of using envelopes with first-class stamps. Preliminary results from the 1999 O*NET pretest provided mixed results regarding the benefits of using real stamps on the return envelopes. Effects on response rates were insignificant; however, there was some anecdotal evidence from the BLs that the first-class stamps yielded shorter response times than the business-reply envelopes. The Wave 1.1 experiment was undertaken to explore this issue further. The results of the experiment supported the pretest finding that the use of first-class stamps had no effect on either response rates or response time. As a result, the use of stamps was discontinued in favor of the less expensive business-reply envelopes.

B.4.3 Point-of-Contact Incentive Experiment

The POC incentive experiment considered the effects on establishment and employee response rates of offering the POC a $20 incentive in addition to the other incentives that the POC receives for O*NET participation (i.e., the desk clock and framed Certificate of Appreciation). About 80% of the approximately 10,500 establishments and 30,000 employees

involved in the experiment were assigned to the treatment group, and the remaining 20% were assigned to the control group. With this large sample size, statistical power of the experiment was very high.

The results provided no evidence that the incentive had any effect on establishment cooperation rates: the POC appeared equally as likely to initially agree to participate with the $20 incentive as without it. Nor was there evidence of any benefit for employee response rates. Given these results and the considerable cost of providing monetary incentives to the POC, in December 2004 it was decided that the $20 POC incentive should be discontinued for all newly recruited POCs. Detailed results can be found in Biemer, Ellis, Robbins and Pitts (2006).

B.4.4 Experiments in Weight Trimming

A substantial component of the sampling error for O*NET estimates is due to the survey weights. Known as the *unequal weighting effect* (UWE), it can be quite large because of the disproportionate sampling methods that must be employed to achieve data collection efficiency. The UWE can be reduced through weight trimming but only at the risk of increasing selection bias. Alternative methods for weight trimming were investigated from 2005 to 2007. This investigation involved assessing the effect of successively aggressive weight-trimming plans for a wide range of estimates and population domains. The weight-trimming analysis was comprehensive, including:

- Comparison of UWEs

- Graphical and tabular displays of current weight estimates versus aggressively-trimmed weight estimates

- Evaluation of weights on suppression of estimates

- Evaluation of statistical differences between current weight estimates and aggressively trimmed weight estimates

- Evaluation of substantive differences between current weight estimates and aggressively trimmed weight estimates.

The method and results of the evaluation are described in an internal report (Penne & Williams, 2007, July 30). The evaluation resulted in the implementation of a more aggressive weight-trimming plan that provides an optimal balance of sampling variance and bias.

B.4.5 Experiments in Model-Aided Sampling

The use of MAS to enhance the efficiency of the O*NET sample is considered one of the most important sample design innovations in the history of the program. This approach dramatically reduced data collection costs with minimal effects on the accuracy of the estimators. Work on the development of the MAS methodology began in early 2004 and continued through

the end of 2007. As part of this research effort, various sample cutoff or stopping rules were investigated by means of Monte Carlo simulation. A cutoff rule determines the point at which efforts to interview certain types of establishments and employees are discontinued because the prespecified quota cell criteria have been satisfied. These studies showed that, under a fairly wide range of cutoff rules, MAS does not bias the O*NET estimates or their standard errors. Furthermore, MAS substantially reduced the number of establishment contacts required in order to satisfy the random sample allocation for an occupation. This innovation resulted in substantial reductions in respondent burden, data collection costs, and time required to complete an occupation (Berzofsky et al., 2006). Details of MAS work are given in Appendix G.

B.4.6 Alternative Measures of Uncertainty

The standard error of an estimate is a measure of statistical precision that is inversely proportional to the sample size. However, for many O*NET data users, the interrater agreement for an average scale rating is very important for their applications. Therefore, beginning in 2005, alternative measures of uncertainty of scale estimates were investigated to supplement the current use of standard errors. Three alternatives—the standard deviation, kappa, and weighted kappa—were analyzed and contrasted, using actual O*NET data as well as simulation.[29] This work led to the decision to make available the standard deviation as a second measure of uncertainty.

B.4.7 Suppression of Estimates with Poor Precision

Before the O*NET data are published, a set of suppression rules is applied to identify estimates that may have extremely poor precision due to very large variances or inadequate sample sizes. Estimates that fail these rules are flagged to caution data users that the estimates are unreliable and should be interpreted as such. Ideally, estimates that are sufficiently reliable are not flagged (i.e., suppressed). An optimal set of suppression rules balances the need to provide as much data as possible to the data users with the need to duly warn users about

---

[29] To define kappa and the weight kappa, consider the cross-classification of ratings from two raters, *A* and *B*. For a standard 5-point Likert scale, the expected proportion of entries in cell (k, k') of the AB table is $n_k n_{k'} / n^2$ for $k, k' = 1, \ldots 5$, where n_k is the number of raters in the sample that select category *k*. The kappa statistic is defined as

$$\kappa = \frac{P_e - P_0}{1 - P_0},$$

where P_e is the agreement rate (sum of the diagonal elements of the AB table) and P_0 is the expected agreement rate, assuming completely random ratings (i.e., all categories are equally likely). The weighted kappa is similar, except the agreement rates, P_e and P_0, include some fraction, f_d, of disagreements, where *d* is the distance of a disagreement from the diagonal (see Johnson, Jones, Butler, & Main, 1981). In the O*NET application, the f_d weights proposed by Cicchetti and Allison (1971) were used, which have also been implemented in SAS Proc Freq.

estimates that are extremely unreliable. In 2004 and 2005 alternative methods for suppression were investigated. The study also evaluated the current methodology for estimating the standard errors of the estimates—the generalized variance function (GVF) method. Using Monte Carlo simulation techniques, as well as actual O*NET data, the study found that the GVF method produced standard error estimates that were severely positively biased. As a result, many estimates that were sufficiently reliable were erroneously suppressed because their standard errors were overstated. Use of the GVF method was discontinued in favor of a direct method of variance estimation. Consequently, estimates published before this change were republished under the direct method of variance estimation. No other changes in the method of estimate suppression were required.

B.4.8 Dual-Frame Sampling for Hard-to-Find Occupations

Some occupations are rare and dispersed across many SICs, making them difficult to find. Identifying an employee in one of these occupations can require calling scores of establishments over many weeks of data collection. To address this inefficiency, a dual-frame approach for sampling employees in hard-to-find occupations was tested. As the term implies, the dual-frame sampling approach involves two sampling frames: (1) the usual SIC frame, which has very good coverage of all occupations, regardless of their size, and (2) a smaller, more targeted frame (such as a professional association list), which may have much lower coverage of all employees in the occupation but contains a very high concentration of them, making them more efficient to sample and contact than a multitude of establishments in many industries. Using these two frames in combination provides the benefits of good population coverage with reduced data collection costs and reduced establishment burden. The testing and evaluation of the dual-frame sampling option was introduced in 2004 for selecting college teachers. This test showed the dual-frame approach to be an effective method of efficiently completing difficult-to-find occupations. Details regarding weighting and estimation for the approach were further developed and refined and, eventually, it was expanded for use in other hard-to-find occupations when suitable second frames of employees in the occupations are available.

B.4.9 Alternative Levels of Population Coverage

In constructing a sampling frame for an occupation, the sampling statistician has a choice: include all possible establishments where the occupation could be found, or only those where the probability of finding the target occupation is high. The former strategy will generate a sampling frame having 100% coverage of the population but with many nonpopulation units. The latter approach will reduce the number of nonpopulation units (and the sampling and data collection costs) but at a lower rate of population coverage and, thus, with greater potential coverage bias. The optimal level of frame coverage is one that maximizes data collection efficiency while maintaining an acceptable level of coverage bias. In 2004 experiments were

conducted to determine the optimum coverage level. Frames were constructed having a range of coverage levels (from 50% to 100%), and sampling from each of these frames was simulated with existing data. Both the estimates and their standard errors and the associated costs of completing the occupations were compared across the frames for many occupations. These studies clearly indicated that frame coverage could be safely reduced to as low as 50% with essentially no risk of coverage bias and that data collection costs could be substantially reduced. These results led to the adoption of a minimum coverage rate of 50% for all occupation sampling frames.

B.4.10   Analysis of Unit and Item Nonresponse

Nonresponse in the O*NET Data Collection Program can occur from the establishment POC at the screening/verification, recruiting, or sampling stages of selection. Within-unit nonresponse occurs at the employee level when a selected employee fails to complete and return a questionnaire. In addition, employees who return their questionnaires may inadvertently or intentionally skip one or more question items on the questionnaire. This type of missing data is known as *item nonresponse.* The effects of all three types of nonresponse on the estimates have been continually investigated since 2003 and reported to OMB; the nonresponse analysis for Analysis Cycles 4–8 appears as Appendix H.[30] Such analyses have shown that nonresponse errors, whether unit, within unit, or item, are not an important source of error in the O*NET Program, because the lost data appear to be missing at random.

B.4.11   Additional Tests of Procedures

This list of the O*NET tests of procedures is far from complete; many other, smaller tests and evaluations have been conducted for the O*NET Program in the name of continuous quality improvement. Still, the list indicates the breadth and depth of the investigations and highlights some of the major results and design innovations. Work continues on quality improvement to this day because optimum quality for an ongoing survey of a dynamic, ever-changing population is a quest, not a destination.

B.5   Statistical Consultants

The DOL/ETA official responsible for the O*NET Data Collection Program is Pam Frugoli (202-693-3643). Through a DOL grant, the National Center for O*NET Development in Raleigh, North Carolina, is responsible for managing O*NET-related projects and contracts and for providing technical support and customer service to users of the O*NET data and related products. Under contract to the Center, RTI International is responsible for providing sampling,

---

[30] The OMB Supporting Statement dated September 2, 2005 (Appendix G), included nonresponse analyses for Analysis Cycles 1–3.

data collection, data processing, and data analysis services. Additional analyses are provided by HumRRO, Inc., in Alexandria, Virginia, and North Carolina State University in Raleigh, North Carolina. The statistical consultants listed in Exhibit 18 reviewed this OMB Supporting Statement.

Exhibit 18.   Statistical Consultants

Name	Organization	Telephone Number
Nonfederal Statisticians and Researchers		
John Campbell	University of Minnesota	612-625-9351
Michael Campion	Purdue University	765-494-5909
Federal Government		
Daniell Toth	U.S. Bureau of Labor Statistics	202-691-7380
Data Collection/Analysis Contractor (RTI International)		
Paul Biemer	RTI International	919-541-6026
Rick Williams	RTI International	919-541-6075
Michael Penne	RTI International	919-541-5988
Marcus Berzofsky	RTI International	919-316-3752

The primary authors of Section B.1 of the Supporting Statement are Drs. Rick Williams and Paul Biemer. Dr. Williams is a senior research statistician and Dr. Biemer is a Distinguished Statistical Fellow at RTI International.

Draft versions of the Supporting Statement were reviewed by BLS, in accordance with DOL protocol and using DOL's presubmission checklist for OMB packages. ETA received one set of six questions about the draft version dated March 17, 2008, and a second set of six questions on the revised draft dated May 30, 2008. ETA's responses to these questions appear in Appendix J and indicate revisions made to the text of the Supporting Statement as a result of BLS's review. ETA was notified July 30, 2008, of BLS's approval of the OMB package.

References

ACT. (2006). *Ready for college and ready for work: Same or different?* Retrieved February 6, 2008, from http://www.act.org/research/policymakers/pdf/ReadinessBrief.pdf

Berzofsky, M. E., Welch, B., Williams, R. L., & Biemer, P. P. (2006). Using a model-assisted sampling paradigm instead of a traditional sampling paradigm in a nationally representative establishment survey. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 2763–2770.

Biemer, P., Ellis, C., Pitts, A., & Robbins, K. (2006). Effects on response rates and costs of a monetary incentive for the point of contact in an establishment survey. *Proceedings of the American Association of Public Opinion Research*.

Blinder, A. S. (2007). *How many U.S. jobs might be offshorable? (CEPS Working Paper No. 142)*. Princeton, NJ: Princeton University.

Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Job analysis: Methods, research, and applications for human resource management* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Cassio, J. (2006). *The Career Pathways Handbook*. Victoria, Canada: Trafford Publishing.

Chromy, J. R. (1979). Sequential sample selection methods. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 401–406.

Cicchetti, D. V., & Allison, T. (1971). A New Procedure for Assessing Reliability of Scoring EEG Sleep Recordings. *American Journal of EEG Technology, 11*, 101–109.

Deville, J. C., & Särndal, C. E. (1992). Calibration estimation in survey sampling. *Journal of the American Statistical Association, 87*, 376–382.

Dillman, D. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.

Dillman, D. (2000). *Mail and Internet surveys: The tailored design method*. New York: Wiley.

Employment and Training Administration, U.S. Department of Labor. (2005). *O*NET data collection program: Office of Management and Budget clearance package supporting statement and data collection instruments.* Washington, DC: Author.

Fleishman, E. A., & Mumford, M. D. (1988). The ability requirements scales. In S. Gael (Ed.), *The job analysis handbook for business, industry, and government*. New York: Wiley.

Fleishman, E. A., & Mumford, M. D. (1991). Evaluating classifications of job behavior: A construct validation of the ability requirements scales. *Personnel Psychology, 44*, 523–575.

Folsom, R. E., Potter, F. J., & Williams, S. R. (1987). Notes on a composite size measure for self-weighting samples in multiple domains. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 792–796.

Folsom, R. E., & Singh, A. C. (2000). A generalized exponential model of sampling weight calibration for extreme values, nonresponse and poststratification. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 598–603.

Folsom, R. E., & Witt, M. B. (1994). Testing a new attrition nonresponse adjustment method for SIPP. In *Proceedings of the American Statistical Association, Social Statistics Section*. Washington, DC: American Statistical Association.

Fuchs, M. (2002). *Paper and pencil versus Web survey: Measurement effect and mode differences*. Paper presented at the annual conference of the American Association for Public Opinion Research.

Johnson, L., Jones, A., Butler, M., & Main, D. (1981). *Assessing Interrater Agreement in Job Analysis Ratings*. San Diego, CA: Naval Health Research Center.

Klein, R. J., Proctor, S. E., Boudreault, M. A., & Tuczyn, K. M. (2002). *Healthy people 2010 criteria for data suppression*. Hyattsville, MD: National Center for Health Statistics.

Korn, E. L., & Graubard, B. I. (1999). *Analysis of Health Surveys*. New York: Wiley.

LaPolice, C. C., Carter, G. W., & Johnson, J. W. (in press). Linking O*NET Descriptors to Occupational Literacy Requirements Using Job Component Validation. *Personnel Psychology*.

Lommel, J. M. (2004). The big "O" in your job search: O*NET. *Gary Johnson's Brave New Work World Web site* Retrieved December 17, 2007, from http://www.newwork.com/Pages/Networking/2004/The%20big%20O.html

Mumford, M. D., Peterson, N. G., & Childs, R. A. (1997). Basic and cross-functional skills: Evidence for the reliability and validity of the measures. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret, E. A. Fleishman & K. Y. Levin (Eds.), *Final technical report*. Salt Lake City, UT: Utah Department of Workforce Services through a contract with American Institutes of Research.

Paxon, M. C., Dillman, D. A., & Tarnai, J. (1995). Improving response to business mail survey. In B. G. Cox (Ed.), *Business survey methods*. New York: Wiley.

Penne, M., & Williams, R. (2007, July 30). *Aggressive Weight Trimming Evaluation*. Research Triangle Park, NC: RTI International.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., & Fleishman, E. A. (1995). *Development of prototype Occupational Information Network (O*NET) Content Model*. Salt Lake City, UT: Utah Department of Workforce Service through a contract with the American Institutes of Research.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., & Levin, K. Y. (Eds.). (1997). *O*NET final technical report*. Salt Lake City, UT: Utah Department of Workforce Services through a contract with American Institutes for Research.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., et al. (2001). Understanding work using the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology, 53*, 451–492.

Peterson, N. G., Mumford, M. D., Levin, K. Y., Green, J., & Waksberg, J. (1997). Research method: Development and field testing of the content model. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret & K. Y. Levin (Eds.), *O*NET final technical report*. Salt Lake City, UT: Utah Department of Workforce Services, through a contract with the American Institutes for Research.

Peterson, N. G., Mumford, M. D., Levin, K. Y., Green, J., & Waksberg, J. (1999). Research method: Development and field testing of the content model. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret & E. A. Fleishman (Eds.), *An occupational information system for the 21st century: The development of O*NET*. Washington, DC: American Psychological Association.

Peterson, N. G., Owens-Kurtz, C., Hoffman, R. G., Arabian, J. M., & Whetzel, D. C. (1990). *Army synthetic validation project*. Alexandria, VA: U.S. Army Research Institute for the Behavioral Sciences.

Roth, P. L., & BeVier, C. (1998). Response rates in HRM/OB survey research: Norms and correlates, 1990–1994. *Journal of Management, 24*(1), 97–117.

RTI International. (2004). *SUDAAN language manual* (Release 9.0). Research Triangle Park, NC: Author.

Schmelter-Davis, H. (2004). O*Net in Action—Expanding Youth Career Horizons. In S. M. Wakefield (Ed.), *Unfocused kids: Helping students to focus on their education and career plans: A resource for educators* (pp. 489-512). Indiana, IN: CAPS Press.

Tsacoumis, S. (2007, May). *The feasibility of using O*NET to study skill changes*. Paper presented at the National Academies Center for Education, Workshop on Research Evidence Related to Future Skill Demands.

Tsacoumis, S., & Van Iddekinge, C. H. (2006). *A comparison of incumbent and analyst ratings of O*NET skills*. Alexandria, VA: Human Resources Research Organization.

Tulp, D. R., Jr., Hoy, E., Kusch, G., & Cole, S. (1991). Nonresponse under mandatory vs. voluntary reporting in the 1989 survey of pollution abatement costs and experiments (PACE). *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 272–277.

U.S. Bureau of Labor Statistics. (2007). Household data not seasonally adjusted quarterly averages: D-13. Employed persons by sex, class of worker, full- or part-time status, and race. *Employment and Earnings*. Retrieved December 19, 2007, from http://www.bls.gov/web/cpseed13.pdf

U.S. Office of Management and Budget. (2000). *The standard occupational classification manual*. Springfield, VA: U.S. Department of Commerce, Technical Administration, National Technical Information Service.

Valentine, B. (2004, February). O*NET Consortium: Occupational Information Network. *C&RL News*. Retrieved November 23, 2007, from http://www.ala.org/ala/acrl/acrlpubs/crlnews/internetreviews/200402/onet.cfm

Wall, J. E. (2006). *Job seeker's online goldmine: A step-by-step guidebook to government and no-cost Web tools*. Indiana, IN: Jist Works.

Willenborg, L., & De Waal, T. (1996). *Statistical Disclosure Control in Practice Series: Lecture Notes in Statistics* (Vol. 111). New York: Springer.

Worden, G., & Hoy, E. (1992). Summary of nonresponse studies conducted by industry division, 1989–1991. U.S. Bureau of the Census.