

O*NET Survey and Sampling Questions (4-01-09)*

Section I: Survey Pretesting

Question 1: Focus groups, cognitive interviews, and expert review were conducted to evaluate the suite of O*NET questionnaires. Results are documented in a 2000 report by Hubbard et al. While the report lists various changes made to the questionnaires as a result of the pretesting, it is not always apparent what changes were made as a function of which methodology employed.

In particular, it is not clear how the cognitive interviewing results from Lesgold et al. were utilized. Further, the process for conducting the cognitive interviews is not detailed. For example, what are the respondent demographics, how many respondents were interviewed, and how were respondents recruited? Can you elaborate on this process, including the particular methods used (concurrent vs. retrospective probing, mode of administration, etc.) to evaluate the survey items.

Response 1: Prior to the grant to North Carolina to further develop and implement O*NET, potential problems with the O*NET survey instruments had been identified by their developers, American Institutes for Research (AIR), in their tryout of the survey questionnaires. Some changes were made to the O*NET survey instruments as a result of the potential problems identified by AIR. In addition, the Learning Research and Development Center at the University of Pittsburgh had conducted a review and evaluation of the survey instruments. It is our understanding that at the Learning Research and Development Center, Lesgold et al. utilized the expanded interview method, using survey respondents to provide information about the how they responded to the survey questions; however we do not have detailed information about respondent demographics or the particular methods used. Prior to implementation of data collection, a further review of the questionnaires was conducted by Research Triangle Institute (RTI), HumRRO, and the National Center for O*NET Development to build on the work of the questionnaire developers. This review identified and addressed content and format issues with the O*NET questionnaires.

Question 2: In reviewing the revisions made following the pretesting, many changes were implemented, including formatting adjustments, item and instruction reduction, and anchor revision. However, it appears that most of the input for change came from expert reviewers and/or researchers from RTI. Can you elaborate on the job incumbent feedback? Was that input incorporated into any of the questionnaire revisions? If not, why?

Response 2: Questionnaire items underwent extensive expert review by researchers at RTI, HumRRO, and the National Center for O*NET Development. Additionally, the focus group of incumbents provided valuable feedback on how difficult it was to understand the instructions and use the questionnaire. Three different versions of the questionnaire were tested: (1) the originals, (2) ones with moderate alterations, and (3) ones with extensive alterations. As a result of the focus group of incumbents, questionnaire instructions were further modified and the instruments were reformatted.

Question 3: Based upon the pretesting findings, is there an overall summary that shows the types of changes recommended by each qualitative method employed?

Response 3: The Hubbard et al. documentation provided summarizes information about the methods used, the issues identified, and the specific format and content changes made. Although this summary does not directly link sources of feedback to specific changes made to the questionnaires, there was considerable overlap in responses from incumbents, experts, and group interviews. The feedback from the focus group of employees (incumbents) conducted at RTI was consistent with the expert appraisal. For example, both the experts and the focus group participants found the instructions to be too long, complicated, and abstract. They found it frustrating to complete the level scale when a variable was rated not important to an occupation. The O*NET team's Instrument Revision Working Group, which consisted of members for the National Center for O*NET Development, HumRRO, RTI, and the Department of Labor (DOL), considered all sources of feedback and evaluation, along with past work that had been conducted by AIR, when making questionnaire revisions. Additionally the review of the instruments by Hubbard et al. confirmed suggestions that job incumbents can more easily respond to questions about what they do on the job, and that the most abstract of the survey questionnaires, Abilities, and possibly Skills, might be best rated by trained analysts.

In short, the job incumbent feedback collected from focus group members was incorporated into final questionnaire revisions. These modifications included new skip patterns, shortened instructions, clarification of specific variable definitions, reformatting of questions, and improved page layout. These modifications were in addition to changes made to the questionnaires by AIR and their partners following their pre-test of the O*NET questionnaires.

Question 4: Why were behavioral anchored rating scales used to assess level of work activity/skill instead of traditional survey response categories? Did the pretesting step result in any substantive changes to the content of the anchor descriptors since respondents appeared to find them difficult to understand?

Response 4: Based on job analysis procedures recommended by Fleishman and Mumford (1988), behavioral anchors were developed and tested on 7-point level scales for the O*NET Content Model via standard judgmental scaling procedures (Childs & Whetzel, 1995). These behavioral anchors were then used on the 7-point level scales within the O*NET Content Model

field test by AIR and Westat in 1995. Interrater agreement coefficients on Skills level ratings were high, ranging from .75 to .92. Considering the cognitive complexity of Skills descriptors, relative to that of the domains that incumbents continue to rate (Knowledges, Work Activities, Work Context), these results support the use of behavioral anchors in rating their jobs on O*NET descriptors. Following the AIR/Westat field test, the O*NET team made some additional formatting changes to the anchors following the work of Hubbard et al. and a pretest of the revised O*NET questionnaires. Anchor placement was changed from "floating," or between whole-number scale points, to placement at whole-number scale points. Questionnaire instructions were modified to reflect this change in anchor scale points. Within the current data collection effort, results of ongoing data quality checks make no indication of substandard scale anchor functioning.

Childs, R.A., & Whetzel, D. L. (1995). Scaling studies. In Technical memorandum: Tryout of O*NET questionnaires and anchor scaling. Washington, DC: American Institutes for Research.

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

Section II: Data Collection

Question 1: Provide response rates at each stage of data collection, including cases that were ineligible, possibly ineligible, and refusals (Establishment Screening, POC Negotiation, Employee).

Response 1:

- Business eligibility rate (79%)
 - Proportion of sampled establishments that were currently in operation at the sampling frame address.
- Screening rate (93%)
 - Proportion of business-eligible establishments that told us whether or not they had any target occupations.
- Occupation eligibility rate (43%)
 - Proportion of screened establishments that had one or more target occupations.
- Consent rate (62%)
 - Proportion of occupation-eligible establishments that participated in data collection.
- Overall cooperation rate (75%)
 - Proportion of business-eligible establishments that had one or more target occupations and participated, or that did not have any target occupations and communicated this to us.

- Establishment Method response rate (65%)
 - Proportion of sampled employees who completed their questionnaire.
- Occupation Expert (OE) Method response rate (81%)
 - Proportion of sampled and eligible OEs who completed the questionnaires.
- Establishment Method item-level response rate (98%)
- Occupation Expert (OE) Method item-level response rate (99%)
- Response rate definition for both of above: Proportion of required questionnaire items that have a valid response

Question 2: Has the O*NET team ever tested a household RDD survey methodology, particularly for collecting data from populous occupations? If not, why?

Response 2: A RDD approach has been considered but not tested because there appears to be a compelling case for the establishment-based approach. The establishment method offers good coverage for a large majority of occupations and minimizes burden and data collection costs because workers are located and sampled in clusters at the establishments where they work, instead of in individual households. Efficient sample design is possible because we are able to use employment data from the Bureau of Labor Statistics' (BLS) Occupation Employment Statistics (OES) program to target our samples to establishments where the occupations of interest are most likely to be found. In addition, cost savings are realized from the use of self-administered mail and Web questionnaires, obviating the need for questionnaire administration by an interviewer.

Alternatively, a RDD approach would be challenging for a number of reasons:

- RDD response rates have decreased dramatically in recent years, with reports of 20 to 30% not uncommon.
- RDD surveys exclude no-phone households, which can be significant and disproportionate across occupations. In addition, the growing number of cell-phone-only households is adding to the coverage problem.
- The screening costs for many occupations would be huge because the prevalence of the occupation in the general household population is quite small. A RDD approach would only be feasible for the most prevalent occupations (e.g., secretaries, waiters, etc.). Even then, the screening would be expensive because the calls would have to continue until the least prevalent cell is filled.
- Because a RDD approach would be limited to only the most prevalent occupations, we would still need to continue the establishment method. The need to staff and maintain both approaches simultaneously would have adverse cost effects.
- It is difficult to provide prepaid incentives over the telephone. It would be necessary to obtain the respondents name and address, mail the incentive, and then follow up with a phone call. Without prepaid incentives and the offer of respondent anonymity, already low response rates could be expected to suffer further.
- It would be very difficult to match household respondents with the correct O*NET occupation. This would add complexity and error to the screening process.

- The O*NET questionnaires would need to be adapted for interviewer-assisted administration. In addition, the introduction of a second data collection mode would introduce potential mode effects that would confound comparisons among occupations.

Given these considerations, we do not believe that a formal test of the RDD approach is warranted.

Question 3: What are the procedures employed to ensure small establishments are not underrepresented?

Response 3: OMB is concerned about the burden placed on small establishments. The sample design ensures adequate representation of small establishments, but it does not overly burden these establishments. O*NET sampling selects small establishments at a lower rate than that at which they occur in the population. In addition data collection procedures place lower burden on small establishments than on large establishments. Therefore, the sample design for the Establishment Method attempts to strike a balance between a design that proportionally represents all establishments and a design that efficiently selects incumbents in target occupations. Our frame, provided by Dun and Bradstreet, contains approximately 15 million establishments, including the self-employed. Of the establishments on the frame, 72% have fewer than 10 employees and 84% have fewer than 50 employees. Our sample design stratifies by the number of employees in an establishment and undersamples smaller establishments (fewer than 50 employees) and oversamples large establishments (250 or more employees) to (1) ensure that large establishments are represented, and (2) maximize the number of potential incumbents we can select while still representing all establishment sizes. Under this design, 31% of selected establishments have fewer than 10 employees and 55% have fewer than 50 employees. Although these percentages are smaller than their actual percentages in the population, we believe that they are large enough to provide adequate representation. Further, as described in the OMB package (pages 64-65 and Appendix G), the sample design uses model-aided sampling (MAS) to ensure that our establishment samples are representative. Since one of the factors for which MAS defines response targets is employee size, the use of MAS ensures that employees in small establishments are represented in occupations that are found in establishments of this size.

Question 4: Discuss the employee selection process. Has the O*NET team ever considered providing the names of employees to the data collection team for nonresponse follow-up purposes?

Response 4: During the recruiting call to the establishment's point-of-contact (POC), the RTI Business Liaison (BL) requests that the POC compile a numbered list of employees' names for each selected occupation. During the subsequent sampling call, the BL obtains from the POC the number of names on each roster and enters the counts into the case management system which utilizes a preprogrammed sampling algorithm to select the sample. The BL then tells the POC the

roster line numbers of the employees selected for each occupation and explains that a set of questionnaire packages will be sent to the POC for distribution to these sampled employees.

To minimize POC burden, the following constraints are imposed: (1) a maximum of 5 occupations can be sampled at an establishment, (2) the number of employees selected from a single sampled occupation may not exceed 8, and (3) no more than 20 employees can be selected within any single establishment.

The BL obtains the roster line numbers of the selected employees but does not normally obtain their names – only the POC has this information. The POC (not the BL) is responsible for coordinating data collection within the establishment and for following up with nonresponding employees. The BL assists in this process by maintaining contact with the POC to keep him/her informed of the roster line numbers of those employees who have responded and to encourage the POC to follow up with those who have not yet responded.

In rare cases, a POC will request that we contact the employees directly. In this event, we will proceed accordingly. However, our experience has been that almost all establishments prefer to coordinate data collection themselves, and our sense is that many establishments would refuse to participate if we were to require them to provide us with the names of their employees and allow us to survey them directly.

Question 5: How does the O*NET team determine which occupations are surveyed using the supplemental method (i.e., the occupation expert method, the non-probability based approach)?

Response 5 : The O*NET Center Sampling Team and Occupation Expert Team determine the data collection methodology to be used for each occupation. The Occupation Expert (OE) Method is used to collect data on occupations less suited to the Establishment method and for which one or more relevant professional associations or other organizations are available to assist by identifying expert respondents. Occupations without Occupational Employment Statistics (OES) employment-by-industry data at the specific-occupation level, those with a small number of incumbents, and those in which employees work in remote locations are considered for OE. Occupations for which OES data are not found at the specific-occupation level are New and Emerging (N&E) O*NET-SOC occupations, such as 17-2141.01 Fuel Cell Engineers, and O*NET-SOCs at a more detailed level than the SOC, such as 19-3091.02 Archeologists. If preliminary contact with relevant professional groups indicates a high likelihood of their cooperation, the occupation is assigned to the OE Method.

In conjunction, the Establishment team examines the likely success of being able to collect the occupation within Establishment data collection. The team examines occupation employment, industries likely to have the occupations, and the projected success of finding the occupations within these industries. The O*NET team conducts an ongoing review of these factors to determine whether low to medium projected OE success occupations are surveyed using OE or Establishment methodology.

Question 6: How many employees complete the survey in Spanish? Describe assurances that the Spanish version is culturally equivalent to the English version.

Response 6 : Spanish versions of the questionnaires are available for occupations with high proportions (>20%) of Hispanic workers. To date, a total of 1,033 Spanish questionnaires have been completed and returned. The total number of Spanish questionnaires requested is small compared to the approximately 142,000 overall number of questionnaires completed and returned. The Spanish questionnaires were initially translated from the English versions by a DOL contractor, Aguirre International. RTI language specialists checked the translations and made revisions as deemed appropriate. The RTI process involved an initial check of the translation by a language specialist, a second check by a different language specialist, and a final review by a third language specialist/supervisor, who resolved any discrepancies between the first two checks. All revisions were documented and reviewed by the O*NET Center. The three RTI language specialists were all experienced translators and native Spanish speakers from three different Spanish-speaking countries.

Section III: Sampling Design Questions

Question 1a: Two approaches are used in the data collection steps. One is referred to as the establishment method and the second is referred to as the occupation expert method.

Questions on the establishment method: It is our understanding that this is a stratified two-stage sampling design. Can you describe the details of the sampling design? For example:

a. What is the frame? Discuss the dual frame that is also used.

Response 1a: The two-stage sampling design involves eight sample selection steps. Establishments are selected during the first steps of selection and employees are selected during the later steps. A detailed description of these eight sample selection steps is provided on pages 65-72 of the 2008 OMB Clearance Package and Supporting Statement.

The frame used in the Establishment Method is the Dun and Bradstreet (D&B) database of establishments. This frame contains over 15 million establishments, including self-employed businesses. For a small group of occupations, there are three scenarios where a non-D&B frame may be used either to supplement the D&B frame or to replace it. In some cases, the O*NET Center Sampling Team determines that other frame options should be employed.

1. *Supplemental Frame Incumbent.* Under this approach the D&B frame is supplemented with a sample of job incumbents selected from the membership list of a professional association or other professional organization. We examine the use of the supplemental frame methodology when experience from previous data collection subwaves indicates we are likely to have difficulty completing data collection without a supplemental frame.

The supplemental frame is compiled through contacts with knowledgeable informants, such as professional associations and experts in the target occupations. Frame coverage, availability, and likelihood of completing data collection influence the choice of whether to use an incumbent or establishment supplemental frame.

2. *Supplemental Frame Establishment.* Under this approach the D&B frame is supplemented with an employer sample selected from a targeted listing of establishments that employ the occupation. We use the same processes for both supplemental frame incumbent and establishment to evaluate the need for a supplemental frame, to compile the frame, and to determine whether to use a frame based on incumbents or establishments.
3. *Special Frame Establishment.* Under this approach a special frame of establishments is compiled that provides very good coverage of the occupations of interest. We use a special frame when it is projected to be more efficient in completing data collection compared with using a D&B frame alone or by using a supplemental frame combined with a D&B frame. In this scenario the special frame is used in lieu of the D&B frame. For example, this approach was used to survey the occupation Nuclear Power Plant Operators. This occupation is only present in nuclear power plants. Since there is a known finite number of nuclear power plants in the U.S. that can easily be identified and listed, the special frame is more efficient than the D&B frame. We use this approach when occupation eligibility at selected establishments within a special frame is projected to be high and the probability of completing the occupation within a series of D&B subwaves is very low.

Under scenarios one and two, the D&B frame is being augmented and, thus, a dual frame situation is created. Under the Supplemental Frame Establishment method, special care is taken to remove any establishments in the supplemental frame that are also in the D&B frame. However, since the Supplemental Frame Incumbent method is a list of job incumbents, we cannot compare the supplemental frame with the D&B frame to remove duplicates. Therefore, if during data collection, we identify an incumbent in the supplemental frame sample who was previously selected in the D&B sample, that person is deemed ineligible from the supplemental incumbent frame. The Special Frame Establishment method has no dual frame issues since only one frame is used.

Question 1b: What criteria are used to select the number of establishments and employees?

- 1) **Discuss sample size for occupations. Were fixed levels of precision required?**
- 2) **Was the funding adequate to meet these precision requirements?**

Response 1b: 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 burden to the public.

The number of establishments sampled in a particular sampling wave is dependent on the number of occupations being targeted in that wave. For a wave containing 50-55 occupations our experience has shown that 7,000 to 8,000 establishments are needed to complete all 50-55 occupations. For efficiency, we allocate that sample across 3 subwaves with 3,000 establishments being assigned to the initial subwave and 2,000 establishments being allocated to the subsequent two subwaves. An additional 1,000 establishments is held in reserve for a completion wave if some of the occupations do not meet the minimum threshold of completed responses to be deemed complete. This wave design is further described in part (e) of this question.

Within a given occupation we require a minimum of 15 valid, completed questionnaires for each of the three domain questionnaires to meet precision targets. Task and background information is collected via a minimum of 45 respondents. The current precision targets for O*NET data are that virtually all 5-point scale ratings will have 95% confidence intervals (CI) *no wider* than +/- 1 and that virtually all 7-point scale ratings will have CIs *no wider* than +/- 1.5. Mumford, Peterson, and Childs (1007, pp. 3-8) have cited Fleishman and Mumford (1991) as support that variation of 1 to 1.5 scale points on a 7-point scale “is typical of that found for well-developed level scales.” Findings in the technical report of Peterson et al. (1997) indicate that with 15 responses, the mean values for virtually all 5-point and 7-point descriptors will have confidence intervals within the targeted range.

Experience-to-date indicates that the average number of respondents per domain questionnaire is approximately 33 and approximately 100 for task and background information. Ninety percent of the 5-point and 7-point scale ratings are within the precision targets. Furthermore, 75% of the 5-point ratings have a confidence interval within +/- .7 and 50% have a confidence interval of +/- .5 scale points or less. Seventy-five percent of the 7-point ratings have confidence intervals within +/- 1.1 and 50% have confidence intervals of +/- .8 scale points or less. (See Exhibit 11 on page 75 of the 2008 OMB Supporting Statement for additional information).

The current funding was adequate to meet the current precision goals. We have investigated what would be required to improve precision targets. To cut the maximum confidence interval targets in half, we estimate that the minimum sample size per domain questionnaire would need to be 60. We derived this number as follows: If we assume a standard error (SE) of 0.5 when $n = 15$, then we solve for σ in the formula for $SE = .25$ with the formula ($\sigma = S.E. * \sqrt{n}$). This equals 1.93649 (i.e., $0.5 * \sqrt{15}$). Now solve for n ($n = (\sigma / S.E.)^2$) when $S.E. = 0.25$ and $\sigma = 1.93649$, $n = 60$.

Question 1c: Businesses with specific occupations are selected and surveyed in this establishment survey. Describe how new and emerging occupations might be included with this approach?

Response 1c: For each occupation, a set of industries is targeted by the Center Sampling Team to identify establishments to sample. New and emerging (N&E) occupations are treated no differently. However, if we do not have adequate information from the Occupation Employment Statistics (OES) survey, the Center Sampling Team obtains supplemental information from the

Center Occupational Taxonomy Team and substantive experts on the industries that employ the new and emerging occupation. Once these industries are identified, the standard establishment sampling methodology can be used. For those new and emerging occupations where we project considerable difficulty completing data collection within the standard establishment methodology, we will examine whether a special or supplemental frame will allow us to complete data collection. If we place the occupation within standard establishment methodology and subsequently experience difficulty finding the occupation, we will use a supplemental frame to complete data collection.

Question 1d: Once an establishment is selected, how are the employees selected within the establishment?

Response 1d: During the recruiting call to the establishment's point-of-contact (POC), the RTI Business Liaison (BL) requests that the POC compile a numbered list of employees' names for each selected occupation. During the subsequent sampling call, the BL obtains from the POC the number of names on each roster and enters the counts into the case management system which utilizes a preprogrammed sampling algorithm to select the sample. The BL then tells the POC the roster line numbers of the employees selected for each occupation and explains that a set of questionnaire packages will be sent to the POC for distribution to these sampled employees.

To minimize POC burden, the following constraints are imposed: (1) a maximum of 5 occupations can be sampled at an establishment, (2) the number of employees selected from a single sampled occupation may not exceed 8, and (3) no more than 20 employees can be selected within any single establishment.

Question 1e: 76 waves of data collection occurred between June 2001 and September 2007. Explain the waves and the data collection process.

Response 1e: 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). A team of substantive experts conduct a rational review of the resulting tables created by merging the Dun & Bradstreet (D&B) industry information with BLS employment information. The team reviews the D&B industries and revises the industry targets based on many factors, including differences in the level of the information merged, occupation research and analysis, exclusionary and related occupations, and an analysis of the best industries to target. The team also incorporates empirical information, if available, from earlier D&B subwaves to guide their industry revisions.

Groups of approximately 50 occupations each, called primary waves, are formed, based on occupation clusters, so that the occupations in a primary wave are employed in a similar set of industries. A team of substantive experts reviews the occupation clusters for a wave and revises the clusters and occupations included in the wave to maximize industry overlap. For these revisions, the team uses experience data from completed data collection waves, if available.

Industry projections may also be used for the new and emerging occupations. 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. 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. This method minimizes the number of establishments that must be contacted for selection of the required number of employees for an occupation.

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 team gains experience from the previous subwaves 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.

The O*NET team completes data collection for a small group of occupations using waves composed of special or supplemental frames. Special frame occupations are completed without using a standard establishment wave. Supplemental frame occupations are completed with a targeted supplemental wave after fielding the occupation within standard establishment wave(s). The team builds and designs the supplemental and special frames to finish within one targeted wave rather than a series of standard establishment subwaves.

Question 2a: Questions on the Occupation Expert method.

- a. Describe this method and how it is combined with the probability method discussed above.

Response 2a: The Occupation Expert (OE) Method is an alternative approach to collecting information on occupational requirements and worker attributes. It is used for selected occupations less suited to the Establishment Method, such as those for which OES data do not exist at a level specific to the occupation (e.g., New and Emerging occupations), with a small number of incumbents, in which employees work in remote locations, and for which relevant, qualified professional groups are available to assist. With this method, persons considered experts in the target occupation, rather than job incumbents, are surveyed. To qualify as an expert, individuals must have worked in the occupation for at least 5 years in some capacity (performing, supervising, or training), including at least one year actually performing the work of the occupation. In addition, they must have worked with the occupation in the past six months. Depending on the occupation, additional criteria may be used to identify experts, such as additional years of experience, education level, or publication history.

The data from occupations completed using the OE Method are not combined with the data from occupations completed using the Establishment Method. No occupation can be partially completed using the OE Method; we continue to sample experts until we complete the occupation. Occupations in the O*NET database are designated as having been completed by one method or the other.

Question 2b: How many occupation experts are there?

Response 2b: Distinct frames of occupation experts are compiled for each occupation that is designated for the OE Method. The number of experts in the sampling frame for a particular occupation can range from over 100 to thousands. The minimum final sample of OEs for an occupation is 20 respondents, with an average of approximately 27, and roughly 30% of occupations having 30 or more respondents.

Question 2c: How are the occupation experts selected?

Response 2c: A member of the O*NET data collection team is assigned an occupation designated for the OE Method. This person researches the occupation and professional and other types of organizations that could be potential sources of experts. The team member contacts relevant organizations to learn more about the occupation and the requirements the organizations would consider important for considering someone an expert in the occupation (e.g., level of education, years of experience). Frequently, we obtain lists from at least three different organizations, which help ensure broad coverage of the occupation. We then work with the organizations to obtain lists of qualified experts representing all important facets of the

occupation, including specialty area and industry. Once we have obtained representative lists, we randomly sample experts to contact for data collection according to region and specialty area.

Question 2d: How many occupations are rated by an occupation expert?

Response 2d: Currently, within a given database release, from 25% to 35% of occupations have data from occupational experts. Occupation experts are identified for a specific occupation; thus they complete questionnaires for only that occupation.

Question 2e: How do you decide which occupations are rated by occupational experts?

Response 2e: The OE Method is used for occupations, as necessary, to improve sampling efficiency and avoid excessive burden, such 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 and for many occupations at the detailed O*NET-SOC level. 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. Experience data on establishment eligibility assists the O*NET sampling team in estimating an occupation's Establishment Method viability in the current data collection. The OE Method is used for an occupation when the Establishment Method is not feasible and an appropriate source of occupation experts has been identified.

Question 3a: The estimates provided also include a measure of variability.

a. Describe the weights, including weight construction for all adjustments used to the sampling weight and the estimation procedure. Discuss the incorporation of the dual frame and the two methods of data collection.

Response 3a: The sample weights are described in detail in the OMB Supporting Statement (pages 75-85). In summary, the final employee-level analysis weights for Establishment Method occupations are computed as the product of a number of factors. These factors reflect the probabilities of selection from the multi-stage design (establishment, occupation, and employee), as well as, appropriate nonresponse, early sampling termination, ratio, and multiple-sample adjustments.

The starting point for each of these stages is the inverse of the probabilities of selection at each stage—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. At each stage, the weights are adjusted for nonresponse using the GEM approach (described in Appendix I of the OMB package) and trimmed to reduce the impact of very small or large weights. Additionally, occupation-stage weights are adjusted to account for early termination of sampling activities. Next, at the employee selection stage, the weights are adjusted for the multiple waves, including dual frames, by which an occupation was selected into the sample. Finally, a ratio adjustment is made to match external population distributions from BLS's OES data.

When a supplemental frame is used to complete data collection, additional weighting steps are necessary to account for the use of dual frames to select the samples. As described in Sampling Design Question 1a above, a supplemental frame may consist of a listing of either establishments or incumbents. In either case, the establishment-level and employee-level weights are constructed in the same manner as when only the D&B frame is used. Thus, when combining a supplemental employee sample with a D&B-based employee sample, the same procedures as those for multiple waves of data collection are followed. That is, employee-level weights are proportionately adjusted to the unweighted percentage of the full sample attributable to that frame (e.g., if the employee sample from the supplemental frame is 20 out of a total of 50 total employees for that occupation, then the adjustment is 0.40).

Since the Occupation Expert (OE) Method does not use a probability-based random sample of all incumbents working in an occupation, the use of weights in the estimation process is inappropriate for these occupations. Additionally, though the frame of occupation experts for any specific occupation may be constructed from multiple sources, the exclusion of weights precludes any necessary adjustments for this aspect. For this reason, the use of the dual-frame concept with regard to occupation experts is not applicable.

Question 3b: It seems that the data used to produce estimates are obtained from the probability sample of establishments and the supplemental non-probability approach using the occupation experts. How are these two sources of data combined to obtain unbiased estimates with measures of variability?

Response 3b: Data collection efforts for any specific occupation are limited to a single methodology (Establishment Method or Occupation Expert Method). Hence, appropriate techniques for the aggregation of probability based and non-probability based samples are unnecessary. Additionally, derived estimates are occupation-specific (i.e., no estimates, including variability measures, are calculated based on combined samples across multiple occupations).

Section IV: Data Processing Errors

Question 1a: What approaches are used to assure data quality? For example:

a. Was double entry used for the mail questionnaires?

Response 1a: Yes, the mail survey questionnaire data are keyed and 100% key verified.

Question 1b: Were any protocols in place in the data editing phase to flag unusual data observations for both data collected on the web and mail?

Response 1b: Paper questionnaires are reviewed and 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.

We then require 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 9% of all returned questionnaires. The cases passing all data cleaning criteria are used to create the estimates for publication.

Question 1c: Were comparisons of data collected from the web vs. mail done to determine if mode differences exist? If any differences were found, how were they addressed?

Response 1c: A statistical analysis of mode differences would need to be performed at the occupation/domain level, and sample sizes at this level would not provide enough statistical power to detect meaningful differences. Additionally, statistical attempts to determine mode differences utilizing combined data across multiple occupations may be confounded by differences in responses by occupations, irrespective of data collection mode.

Substantive reviews have not detected differences between web and paper responses. The O*NET team conducts the same extensive data editing and deviance analysis processes on questionnaires received by both web and mail.

Section V: Other

Question 1: How do you ensure the most “high” high performing occupations are selected for inclusion in O*NET?

Response 1: The O*NET Center Sampling Team ensures the most “high” high performing occupations are selected for inclusion in O*NET by evaluating a range of different occupation criteria. The O*NET classification system (the O*NET-SOC) conforms to the SOC, permitting O*NET data to be aligned to and analyzed with sources of information on current occupational employment and trends, wages, and demographic data. Occupations with high or fast growth, high employment, and high ratings on mathematics, science, computers and electronics, engineering and technology, and job zone all increase an occupation’s likelihood of being selected within the current O*NET data collection. The O*NET team automatically includes those occupations designated as Top 50 by DOL, In-Demand and New and Emerging found in high-growth industries. For the “Update Phase” of data collection, we collected and published data for all 812 O*NET-SOCs contained within the 2006 O*NET-SOC taxonomy.

Question 2: In your opinion, what should be the procedure employed to decide on the appropriate degree of aggregation or within-occupation variability that an occupational title should have? What, if any, should be the procedure employed to disaggregate an occupational title?

Response 2: The O*NET Center Occupational Taxonomy Team determines the level of aggregation for the O*NET-SOC occupations. The O*NET occupational taxonomy has always been a product of empirical analysis, rational review, and customer feedback. A cluster analysis using Ward’s minimum variance was used to develop the original O*NET occupational

taxonomy. The analysis was performed using DOT data variables for DOTs linked to each occupation in the Occupation Employment Statistics (OES) structure. Occupational analysts were used to evaluate the cluster solutions by estimating retraining time within and across clusters. The confirmed clusters were then named according to the membership of DOT occupations included in order to develop O*NET Occupational Units. This early taxonomic work informed the development of the O*NET Analyst Database, and subsequently the 2000 Standard Occupational Classification (SOC).

Since then, the O*NET occupational taxonomy transitioned to utilize the SOC organizational structure while keeping much of the additional specificity of the O*NET Occupational Units for selected SOCs. A limited number of detailed occupations were rolled up to the SOC detailed level where warranted, based on task analysis, and correlations of knowledge, skill and work activity ratings, taking into account the employment statistics and number of titles associated with the occupation. In addition, some more specific and distinct occupations have been identified and included in the O*NET-SOC occupational taxonomy as a result of industry studies conducted in researching new and emerging occupations. An attempt has been made to keep the taxonomy as close to the detailed SOC level as possible while yielding homogeneous occupational units, to be able to provide users of O*NET data with a total picture of both occupational and labor market information.

Steps have been considered to further evaluate the degree of aggregation, now that data has been collected on all O*NET-SOC occupations (excluding recently identified new and emerging occupations). For example, in order to inform potential disaggregation of O*NET-SOC occupations, the O*NET team again has considered a combination of statistical and rational processes, involving the use of task importance ratings to identify distinct functional subgroups within an existing O*NET-SOC occupational classification. Once data are collected for an occupation, a statistical cluster analysis procedure might also be used to identify subgroups of respondents with similar importance ratings of the occupation's task statements. The number of clusters was not specified in advance, allowing the analysis to identify as many occupational subgroups as are statistically meaningful. The identified clusters for each occupation would then be reviewed by a team of analysts to examine the functional distinctness of the subgroups, based on differences in endorsement of specific task statements. If subgroups of respondents are found to differ in their ratings of tasks considered functionally core to the original target O*NET-SOC occupation, a case could be made to disaggregate the O*NET-SOC occupation to the number of meaningful subgroups identified.

Question 3: Do you see it as a problem that specific jobs or titles comprising an O*NET/SOC may have different (true) means and variances on rated descriptors (e.g., abilities, skills, work styles, etc.), resulting in differential degrees of accuracy (validity) of O*NET-SOC descriptor ratings when rating data from different jobs are combined?

Response 3: No, we do not see it as a major concern for workforce development customers. The level of aggregation within the taxonomy has been appropriate for many different uses including career development and skills transfer. Concerns voiced from O*NET customers, establishment

points of contact, and the leaders of professional associations are triggers that would lead us to take a closer look at this issue. While the O*NET team sometimes receives requests to modify occupation titles, descriptions, and tasks, the leaders of these professional associations have not requested breaking up an occupation because it was composed of different jobs. Establishment points of contact within the establishment method sampling methodology are typically human resource professionals with knowledge of a range of different jobs, and can easily identify jobs in their organization that correspond to O*NET-SOC occupations. To date, neither establishment points of contact, the leaders of professional associations, nor O*NET customers have expressed specific concerns about combining data from different jobs into occupations. They have, on occasion, asked for additional occupations or for an occupation to be split into two occupations.

Furthermore, measures of the degree of similarity among raters, or interrater reliability, provide support for the aggregation of data from different jobs into occupations. One measure, the intraclass correlation coefficient (ICC), describes how the variance of the ratings within an occupation compares with the variance in that same item's ratings across occupations. Intraclass correlation coefficients (ICCs) are calculated for each item across all occupations. Across all occupations, most ICCs are above .5. For a sample of 75 occupations published in 2007, the median ICC for Work Activities Importance items was .85. For Work Activities Level items, the median ICC was .80. The Work Context items had a median ICC of .89. The Knowledge Importance items had a median ICC of .83. Knowledge Level items had a median ICC of .86 and the Work Styles items had a median ICC of .64.

The O*NET team examines relative standard error (RSE), or the ratio of the estimate's standard error to the estimate, itself, as a measure of item estimate quality. The lower an estimate's RSE, the higher the estimate's precision. Of all weighted means on Work Activities, Knowledge, and Work Context descriptors considered relevant to the occupation (low rate of legitimate skip of Level rating), approximately 98% are within the acceptable range, or less than .50, the level above which the O*NET Program recommends an estimate for suppression. Approximately 80% of estimate RSEs are below .20.