Appendix F: Analysis of Response Rates, Response Bias, Effectiveness of Incentives, and the Use of First-Class Stamps

Analysis of Response Rates, Response Bias, Effectiveness of Incentives, and the Use of First-Class Stamps

The terms of clearance provided by OMB to the O*NET® *Data Collection Program*, which was submitted in December 2000 and approved on April 4, 2001, mandated that the resubmission include "an analysis of response rates and response bias, and a discussion of the effectiveness of the incentives used, including the use of first-class postage stamps." This appendix includes those analyses and discussions. First, the response rates obtained in Wave 1.1 data collection are examined. Second, response bias is analyzed by examining the primary potential sources of response bias: the source of information and mode effects. Third, the effectiveness of O*NET® data collection incentives is discussed by comparing experiences in Wave 1.1 to what one would expect, based on a review of the professional literature. Finally, the effects of first-class postage stamps on both questionnaire response rates and return time are discussed.

1. Analysis of Response Rates

In this section, O*NET® response rates are compared to response rates that other similar studies have yielded. The O*NET® response rates for establishments and employees exceed past experience for federal surveys that use a similar data collection approach and meet the needs of the O*NET program, but we will continue to work to improve response rates.

1.1 Establishment and Employee Response Rates

Wave 1.1 data collection ran from June 2001 through May 2002. In Wave 1.1, 12,274 establishments were selected for the sample. These businesses were contacted by an O*NET® interviewer (referred to as a Business Liaison, or BL) who explained the O*NET® data collection effort and attempted to elicit participation from a suitable point of contact (POC) at the business. Results indicated that 15% of the businesses contacted were ineligible for the study because they were out of business, a frame duplicate, or otherwise out of scope for the survey for some other reason (e.g., industry change). Among the eligible businesses, 64% agreed to participate in the data collection effort.

After an establishment was screened and deemed eligible to participate in the study, it was sent relevant information about the occupations of interest. Subsequent calls were made to the establishment to gather data on its total number of employees in up to 10 occupations of interest. These occupations were randomly chosen for each business from the 201 target occupations, while giving greater probability of selection to the occupations with the most employees within the organization. The employment information obtained from the POC on the 10 occupations was used to select up to 15 employees within each establishment. In Wave 1.1, a total of 19,623 employees were selected and sent an incentive and a questionnaire (usually through the POC, to preserve confidentiality). Sixty-three percent of the selected employees responded.

The O*NET® *Data Collection Program* has a unique feature that makes analyzing response rates difficult. The O*NET® survey requires participation at two stages of response—the POC (establishment) level and the employee level—whereas the typical establishment survey requires participation at only one level, the establishment level. Because there are very few surveys that incorporate such a design, the survey methods literature is essentially devoid of examples upon which to base a reasonable response rate expectation for the O*NET® *Data Collection Program*. Therefore, a comparison of O*NET® response rates with other establishment surveys must be done separately for each stage of participation. First, the O*NET® establishment-level response rate is compared with other mail establishment surveys having only one response stage at the establishment level. Then, the O*NET® employee-level response rate is compared with the response rate of other establishments' self-conducted employee surveys.

At the establishment level, Paxson, Dillman, and Tarnai (1995) analyzed the response rates for 46 surveys conducted by both government and nongovernment 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 that methodology. Further, the 20 Census Bureau surveys include 12 well-established and 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 the O*NET® establishment-level response rate of 64% exceeds expectations for this type of survey.

The second response stage of the O*NET® survey can be compared to establishment surveys in which the ultimate sampling units are the employees of the 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. However, following are two examples of government-sponsored surveys:

1. One well-documented 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 final overall response rate for the survey was 55%.

2. 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 within numerous federal

agencies provides some evidence of response rates for employee surveys conducted by the U.S. federal government. The OAS design is very similar to 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 rate across all agencies is approximately 57%. This result is similar to the Canadian experience. It is also supported by Roth and BeVier (1998), who conducted a meta-analysis of 173 surveys in the field of Human Resource Management and Organization Behavior (HRM/OB). They found that surveys implementing many of the response rate enhancing features of the O*NET® survey had response rates in the range of 23–78%, with a median of 51%. These results suggest the O*NET® employee response rate of 63% is favorable and exceeds expectations for mail surveys of employees within their organizations.

Despite these positive findings, the impact of nonrespondents is still a concern. The effect of nonrespondents is usually evaluated in terms of potential differences on such factors as demographics (e.g., age, race, sex), job experience level, job performance level, and education. Morgeson and Campion (1997) reviewed 20 job analysis studies that examined the influence of these factors. They concluded that the results from this body of research have been mixed, with a number of studies showing some differences in job analysis information and others showing no differences. However, even when these differences are found, they tend to be small and difficult to interpret, suggesting they are not substantively meaningful. This leads to the conclusion that job analysis data are very robust with regard to these potential influences, and nonrespondents at the level of the O*NET® Program would be unlikely to bias the results in any systematic fashion (M. Campion, personal communication, March 22, 2002).

In summary, results from the survey methods literature and from other federal surveys in the U.S. and Canada suggest that the current 64% establishment response rate and the 63% employee response rate for O*NET's Wave 1.1 data collection effort exceed expectations for federal surveys based on similar data collection efforts. Furthermore, the job analysis literature suggests that nonresponse is not likely to lead to nonresponse bias. However, in the interest of continuous improvement of the O*NET® *Data Collection Program*, methods for further improving response rates continue to be explored.

1.2 Methods to Improve Response Rates

Even though the O*NET® *Data Collection Program* exceeds response rate expectations generated from the response rate literature and other federal surveys, it is believed that steps can be taken to further improve response rates. In September 2001, the Response Rate Improvement Team (RRIT) was formed. The team was charged with recommending protocol enhancements to

maximize O*NET® response rates for the current and future waves of data collection. Activities undertaken by the RRIT are described below.

Development and implementation of STAR training for all Business Liaisons. The survey contractor adapted materials it used in conjunction with the U.S. Bureau of the Census for training Business Liaisons (BLs) in Strategies and Tactics for Averting Refusals (STAR). The process used to develop the training program was as follows:

- Step 1. Focus groups were conducted with the participation of all the BLs in order to identify specific arguments and verbatim statements used by POCs that demonstrate reluctance or refusal to cooperate with the O*NET® survey requests.
- Step 2. These verbatim statements were grouped into themes. An example of a verbatim statement that a POC might say is, "I'm too busy for this." The theme corresponding to this statement would be labeled "Timing of the Call."
- Step 3. Meetings were held with the better-performing BLs to identify the optimal strategies for addressing the arguments under each theme.
- Step 4. Materials were developed to train the BLs to (a) listen to the POCs (or other contacts) in order to precisely understand their arguments for not participating in the survey, (b) categorize their statements into the appropriate theme, (c) quickly recall the appropriate rebuttal for the corresponding theme, and (d) deliver the rebuttal as quickly and smoothly as possible.
- Step 5. Each BL participated in a full-day training session on using the refusal aversion techniques. The training session included a variety of teaching techniques, including lecture, small group work, peer evaluation, and a rapid-fire quiz, in which trainees were challenged to rebut example POC statements on the spot. Each BL was required to pass a certification examination prior to returning to interviewing.
- Step 6. After training, an intensive monitoring system was implemented to ensure that the BLs were able to put the STAR tools to good practice. Supervisors monitored the BLs for appropriate use of persuasive techniques, and the BLs recorded portions of their own phone calls to monitor themselves.

This training will be incorporated as a standard module in the training program for new BL staff members.

Ongoing focus groups for long-term improvements. Survey methods staff continue to conduct focus groups with the BLs and project staff in order to identify successful and problematic features of the data collection design. For example, a feature identified as problematic was the method of relying on information from the screening contact to determine the appropriate POC, without first consulting the person who had been nominated as POC. Because the POC must have enough knowledge of the establishment to determine whether the selected occupations on the sample list are present in the business, it is critical that the appropriate person within the business be assigned to this position. Discovery of this issue in focus group sessions led to proposal of a new design for Wave 1.2 and beyond, to include a

"verification" call. On this call, POCs are contacted prior to Advance Package shipment to verify that the selected POC is appropriate and to brief him or her on package contents.

BL training to deal specifically with refusals due to establishment reductions in force. The survey contractor and the National O*NET® Center developed special scripts to help BLs when they speak with establishments undergoing staff reductions. The new scripts highlight the specific benefits that O*NET® can provide for businesses in this circumstance.

Initiative to address establishments with gatekeepers. In many cases, the POCs agreed to participate, but the employees at the establishment did not return any questionnaires within 60 days. A "gatekeeper" effect was suspected to be the cause. That is, it was surmised that in many of these situations, the POC had not distributed the questionnaires to the sample employees, and hence the employees could not respond. A team of experienced, successful BLs was trained as Gatekeeper Specialists to work cases falling into this category.

"Follow-up call past due" report. To aid in the tracking of cases for which the 7-, 21-, or 31-day follow-up call is overdue, a "follow-up call past due" report was developed. This report, which can be run by either a BL (for his or her own assignment) or the facility manager (for any BL), reports on those cases in the system for which any follow-up call is 14 days or more overdue. The BL can use this information to prioritize cases that need the appropriate follow-up call. The data collection facility manager uses the report to monitor cases that are overdue for a follow-up contact.

Evaluation of the effectiveness of POC, establishment, and employee incentives. The examination of incentives that was conducted is described in ***Section 3*** of this appendix.

Streamlining the advance package to reduce establishment nonresponse. BLs repeatedly reported that the POCs were often intimidated by the size of the advance package, stating that it was confusing and appeared to be excessively long. The advance package was streamlined to reduce the amount of paper received by the POC as well as the POC's perception of burden.

Additional communication options for the BLs. Enhancements were made to the Case Management System (CMS) during Wave 1.1 to provide the BLs with greater flexibility and effectiveness in communicating with POCs. These included the capability to send both electronic faxes (e-faxes) and electronic mail (e-mail). These new capabilities provide the BLs with more communication options and the ability to tailor their communication style to fit the needs and preferences of the POCs with whom they are working.

2. Analysis of Response Bias

In this section, several potential sources of response bias are reviewed, then several features of Wave 1.1 are examined for evidence of response bias.

2.1 Potential Sources of Response Bias

There are essentially two categories of response bias that could influence the O*NET® results: the source of information (incumbent respondent or Subject Matter Expert [SME] or analyst rater) and the mode of data collection (whether by Web or paper and pencil). Regarding the source of information, the O*NET® data are collected from job incumbents, occupational analysts, and SMEs. While job incumbents are considered to be the most accurate source of data about the activities, knowledge, and skill required to perform their jobs, occupational analysts are used to collect data from the more abstract Abilities domain, and SMEs are used for occupations whose incumbents are not readily accessible through the General Employer approach. In addition, two survey modes are used to collect these data: a mailed paper questionnaire and a Web survey. Respondents have the option of using either mode.

With regard to the source of information, evidence of bias can be analyzed by comparing the responses from the different sources. Significant differences in response means between two information sources suggest the presence of bias in one or both sources. Regarding the mode of administration, we expect the only difference in response quality to be the use of the electronic medium for the Web, rather than paper and pencil for the mailed version because both models use self-administration. Further, since the questionnaire format for the two modes is essentially the same, a response bias in one mode resulting from the questionnaire is likely to affect the other mode as well.

This section discusses potential response bias in the O*NET® data collection design from the use of job incumbents versus occupational analysts, job incumbents versus subject matter experts (SMEs), and paper versus Web questionnaire administration. ***Sections 2.1.1*** and ***2.1.2*** compare the responses from job incumbents with occupational analysts and with SMEs, respectively, in order to test for the bias in all three sources. ***Section 2.1.3*** considers the bias in self-administered mode as well as any evidence from the literature regarding the differential bias for mailed versus Web questionnaires.

2.1.1 Collecting Data from Job Incumbents vs. Occupational Analysts

In the O*NET *Data Collection Program*, job incumbents are the source of data for four of the five O*NET® questionnaire packets for the majority of the occupations. Occupational analysts are used to collect data for the Abilities domain because it is considered to be the most abstract and incumbents may have more difficulty applying the items to their own jobs. Both incumbents and analysts complete the Skills Questionnaire, providing the O*NET® team with the opportunity to evaluate both the extent to which rating sources agree and the quality of the data from each source. Two of the most compelling advantages of the incumbent ratings are the opportunity to collect current data on occupations and the inherent legitimacy of the source as perceived by many of the potential users of O*NET® information.

Both the incumbent and analyst sources have advantages. For example, incumbents are able to provide first-hand, job-level information, whereas analysts are able to discern relative levels of a descriptor across occupations and can receive specific training on completing the questionnaires. The O*NET® team is also aware of potential response biases for each source. These biases are a concern not only for O*NET® data collection but also for occupational and

job analysis research, in general (Morgeson & Campion, 1997). Researchers must be aware of, and take precautions to guard against, these potential sources of inaccuracy. However, given the breadth of many O*NET® occupational classifications, interrater differences may correspond to actual job-related differences (Harvey, 1991), in contrast to typical job analyses conducted within organizations.

Early O*NET® research compared ratings made by incumbents in a preliminary data collection effort to ratings made by analysts on each of the domains. This research found that while incumbent ratings had somewhat lower reliabilities compared to analyst ratings, they provided a more varied and complex description of the occupations (Peterson, Mumford, Levin, Green & Waksberg, 1997). Despite the fact that interrater reliability coefficients for analysts were slightly higher, incumbent interrater reliabilities were exceptionally high. Peterson et al. reported the following median incumbent interrater reliabilities for each of the O*NET® domains:

- Skill median incumbent interrater reliability was 0.84 for Level scales and 0.83 for Importance scales.
- Generalized Work Activity median incumbent interrater reliability was 0.82 for Level scales and 0.78 for Importance scales.
- Knowledge median incumbent interrater reliability was 0.76 for Importance scales.
- Work Context median incumbent interrater reliability was 0.83 for the range of scales in that instrument.

The current and forthcoming O*NET® data collection waves will have more respondents completing the ratings for each occupation than Peterson et al. (1997) reported. Based on increased sample sizes, higher interrater reliabilities might be observed.

While incumbents and analysts likely provide reliable and accurate data, potential response biases have been associated with both incumbent and analyst raters. These biases do not necessarily exist in the O*NET® data but are possible, and the O*NET® team is alert to them. Two potential types of response bias for incumbent raters are impression management and socially desirable responding (Peterson et al., 2001). Results from preliminary O*NET® data collection efforts indicated that incumbents' ratings were somewhat higher than analysts' on some descriptors, whereas analysts' ratings were higher on others (Peterson et al., 1997), possibly due to differing frames of reference for the two rating sources. Socially desirable responding is most likely to occur on descriptors generally considered to reflect positive attributes, skills, knowledge areas, and abilities. This tendency could cause incumbents to inflate their ratings of the items that would portray their jobs in a more socially desirable light (Morgeson & Campion, 1997).

Although incumbents are considered to be more susceptible to inaccuracies resulting from impression management and social desirability than are analysts, these inaccuracies are probably less likely to occur in the O*NET® data collection than in an organization's job analysis. These tendencies are maximized when the outcomes are of greater consequence to the incumbent, such as when a job analysis is completed for compensation purposes (Morgeson &

Campion, 1997). The anonymity of respondents in the O*NET® incumbent sample may reduce socially desirable responding and impression management. The multiple source methodology for the ratings on the Skills Questionnaire allows these potential self-presentation response biases to be further evaluated for all occupations in the O*NET® database. This not only benefits the O*NET® Program, but also contributes to the job analysis research literature.

A third type of biased responding that could affect both analysts' and incumbents' ratings is caused by information-processing-related inaccuracies. These inaccuracies range from careless responding to responses based on inadequate information. The effects of careless responding, or "satisficing," are probably more likely to be found for incumbents than for analysts. Satisficing refers to the tendency for raters to use response strategies that minimize the effort required of them (Krosnick, 1991). For example, they may choose randomly from the response options, or consistently choose the first option that makes sense. Fortunately, the O*NET® instruments are not very lengthy, which may help limit careless responding and satisficing. Careless errors are thought more common for occupations to which many items do not apply or when raters do not understand what they are being asked to rate (Morgeson & Campion, 1997). The applicability of the various questionnaire items to occupations vary, but for each of the domain questionnaires, it is expected that the majority of the items are at least somewhat applicable to most occupations. Analysts may be less likely to respond carelessly; however, it is understood that they are more likely to have lower validity and reliability if they have inadequate information about the job. Analysts with inadequate information may, in addition, tend to rely on their overall impressions of an occupation when making ratings. The O*NET® analyst approach strives to provide analysts with updated and complete information based on recently collected incumbent data.

There are clearly many potential sources for rater inaccuracy from various data sources, and no job analysis is able to prevent them all. The O*NET® data collection's multiple source approach provides an opportunity to use previous research and theory to guide analyses that can detect and possibly correct inaccuracies. Each source of data is considered to be indispensable for its strengths and purpose in the O*NET® data collection effort. With awareness of each source's imperfections, data collection can successfully proceed, informed and proactive.

2.1.2 Collecting Data from Job Incumbents vs. Subject Matter Experts

Whereas the General Employer Sample approach is used to collect data for most of the occupations in the O*NET® database, data are collected from SMEs for a very small number of occupations. Data from this source are collected on occupations that are relatively rare or difficult to sample efficiently under the General Employer Sample or the Association List Sample approaches. For example, the SME approach is used for occupations with very small employment in the national economy, for new occupations with inadequate industry employment data, and for occupations with inaccessible incumbents. SME samples are convenience samples of respondents identified through professional associations, academic organizations, and other professional groups; however, emphasis is placed on representativeness of the occupation. For example, care is taken to represent any factors that could influence the ratings, such as region, company size, and industry type. In addition to the sampling strategy, a primary difference between the SME and General Employer Sample approaches is that each SME is given all four O*NET® questionnaires (the goal is to collect data from 21 SMEs for each occupation).

While incumbent raters are considered to be SMEs of their particular jobs, respondents used in the SME approach are distinguished from the incumbents based on their level of experience and knowledge of the relevant occupation. For example, SMEs are commonly supervisors or trainers of the occupation of interest. In the O*NET® SME approach, the SMEs must meet experience level criteria to be qualified to make these ratings. Because of the close similarity between types of individuals completing the SME and incumbent ratings, the ratings from these sources are expected to be very similar.

Recent research comparing SME and incumbent ratings found high agreement on ratings of tasks, knowledge, skills, and abilities (Maurer & Tross, 2000). That study's method was similar to the O*NET® SME approach in that it used a small number of SMEs and a larger number of field incumbents. Both groups rated tasks on both importance and relative time spent and rated knowledge, skills, and abilities on importance and level. All respondents in that study made their ratings independently. The authors argued that their findings are part of a growing body of research demonstrating that job ratings obtained from a small group of SMEs are likely to agree with those obtained from a larger sample of incumbents. For example, earlier research found high agreement between knowledge and ability importance ratings made by a small SME committee and a large sample of incumbents (Tannenbaum & Wesley, 1993).

Maurer and Tross (2000) cautioned that some conditions could influence or bias SME ratings. Most notably, the group dynamics of intact SME committees could lead to response bias in their ratings. Group dynamics could, for instance, result in the whole group's making higher or lower ratings than the raters would have made independently. This type of bias is not highly relevant for the O*NET® SME approach because the SME respondents are unlikely to come from intact groups. Incumbents and SMEs making independent ratings are similar in most respects, aside from the means by which they are sampled to participate in the O*NET® *Data Collection Program.*

In 2000, the O*NET® Program pilot-tested the SME approach for a single occupation. The pilot test was considered a methodological success. The SME ratings were compared to Wave 1.1 ratings for that occupation (Water and Liquid Waste Treatment Plant and System Operators). ***Section 2.2.1*** below provides more details on this analysis.

2.1.3 Mode Effects Between Mailed and Web Questionnaires

A second category of response bias in surveys is mode effects. Mode effects are a constant concern in surveys where a mixed mode design is used to collect data. In the case of the O*NET® *Data Collection Program*, respondents are offered a choice of two modes in providing their responses: a Web version of the survey and a self-administered paper version.

According to the literature on mode effects in surveys, two critical variables may cause bias in response due to mode effects: whether the survey is self-administered or interviewer-administered and whether the survey task is communicated to the respondent orally or visually (Dillman et al., 2000; Couper, Traugott, & Lamias, 2001). Self-administered surveys tend to result in more honest answers to sensitive questions (such as income level) and questions with answers that vary in social desirability (such as how often people wash their hands before preparing food). On the other hand, interviewer-administered surveys generally produce better

quality data because the interviewer provides on-the-spot edit checks and is adept at navigating complex skip patterns. Furthermore, when the method of communication varies between oral and visual in surveys, answers are subject to the effects of the ordering of response categories. Neither of these variables causes concern for mode effects in the O*NET® survey, however, because both the type of administration (self) and type of communication (visual) are held constant. For this reason, the most likely sources of response bias due to mode effects are avoided.

The literature on mode effects suggests that there are at least three other, relatively minor conditions of response bias due to mode effects in Web as compared to paper questionnaires. First, using a mixed mode approach can affect the overall response rate. Second, the quality of data varies between Web and paper self-administered surveys. Finally, differences in formatting between Web and paper questionnaires can result in response bias.

Griffin, Fischer, and Morgan (2001) found that when comparing an experimental mail-plus-Web response option to the traditional mail-only response option for the American Community Survey (ACS), the mixed-mode approach generated lower response rates. About 43.6% of people responded in the mail-only group, whereas 37.7% of people responded in the mail and Web group. However, Griffin et al.'s findings contradict the more common observation in survey research that allowing multiple modes of response improves response rates because it allows the survey to appeal to the preferences of a wider range of respondents (Dillman, 2000). While there is no way to replicate the ACS test with the O*NET® data because every respondent is offered the Web option, the percentage of the total questionnaires mailed that are returned via the Web is about 9%, whereas the ACS Web response rate was only about 2%. Because O*NET® respondents have a greater tendency to respond via paper, it appears that option is more appealing to O*NET® respondents. There is not likely to be a negative effect on employee participation from offering the Web response option.

Response bias between modes can also be a result of variation in data quality. Griffin et al. (2001) found that on the ACS, data quality was much higher for the Web data than the paper questionnaire data. The Web version of their survey resulted in an 11% failure rate of editing checks as opposed to a 33% failure rate for the paper version. An edit check was considered a failure when an item was missing a response or the response was inconsistent with a known population parameter.

Finally, variations in instrument formatting between Web and paper questionnaires are likely to affect response bias. Couper et al. (2001) argue that instrument design is critical to minimizing response bias in self-administered surveys. When there is no interviewer present to lead a respondent through an instrument, respondents look for cues within the visual display of the questionnaire (Ware, 2000; Couper et al.). Thus, the physical layout of an instrument can have a powerful effect on the answers that the respondent chooses to provide. The O*NET® Web surveys were designed with this concern in mind. The Web questionnaires were formatted to replicate the layout of the paper questionnaires to reduce the likelihood of bias due to questionnaire design.

In summary, while mode effects on response bias are always a concern in mixed mode surveys, in the O*NET® *Data Collection Program* these effects are minimized. Since the only two modes are both self-administered, and since both instruments are formatted similarly to reduce mode differences, every precaution possible has been taken to reduce the impact of mode differences on response bias.

2.2 Examining Wave 1.1 Data for Response Bias

It is difficult to evaluate these potential sources of response bias because no data are yet available to compare job incumbents to occupational analysts. However, there is a small dataset that will allow simple comparison of the SME data to the job incumbent data, and also some sources of data are available to evaluate the overall quality of data being collected in the O*NET® Program.

In this section, analysis of the SME data is presented first. Second, levels of item nonresponse in the O*NET® data are examined, as these are indicators of problems, such as misunderstanding questions, that may lead to response bias. Finally, the percentage of responses that were submitted via the Web across types of occupations is examined.

2.2.1 Comparing Job Incumbent to Subject Matter Expert Data

In 2000, the National Center for O*NET® Development pilot-tested the SME approach for Water and Liquid Waste Treatment Plant and System Operators. The raw data collected from SMEs in the pilot test are compared to the data collected in Wave 1.1 (through February 25, 2002) for this occupation only. The data from both sources are raw, and the data from Wave 1.1 have not been assigned any weights. ***Tables 1*** and ***2*** below display the means and standard deviations from the Wave 1.1 incumbent ratings and the SME ratings. The particular items shown below were the five that had the largest difference between incumbent and SME mean ratings and the five that had the smallest difference between incumbent and SME mean ratings. These items are shown for the Skills and Generalized Work Activities questionnaires. The Knowledge and Work Context questionnaires had similar differences between incumbent and SME mean ratings. Although this is a very small-scale test, both of these tables confirm that the differences between job incumbent and SME ratings are relatively small. For both the Skills and the Generalized Work Activities questionnaires, no difference in mean rating was greater than 1.5 points.

Table 1. SME and Job Incumbent Ratings for the O*NET® Skills Questionnaire from Water and Liquid Waste Treatment Plant and System Operators

Five Items with the Largest Difference							
	Job Incumbents			SMEs			
Variable	Total N	Mean	Standard Deviation	Total N	Mean	Standard Deviation	Difference
Level of learning strategies	18	4.39	0.78	14	3.07	0.92	1.32
Importance of managing financial resources	18	2.17	1.25	15	3.47	0.92	-1.30
Level of programming	7	3.57	0.98	6	2.33	1.03	1.24
Importance of equipment selection	18	2.94	1.26	15	3.87	0.64	-0.92
Importance of repairing	18	2.94	1.21	15	3.87	0.99	-0.92
Five Items with the Smallest Difference							
	Job Incumbents			SMEs			
Variable	Total N	Mean	Standard Deviation	Total N	Mean	Standard Deviation	Difference
Importance of installation	17	3.18	1.33	15	3.20	0.77	-0.02
Importance of complex problem solving	18	3.11	1.08	15	3.13	0.83	-0.02
Level of instructing	16	4.19	0.83	14	4.14	0.77	0.04
Level of time management	17	3.82	1.13	15	3.87	1.19	-0.04
Importance of active listening	18	3.72	0.46	15	3.67	0.72	0.06

Table 2. SME and Job Incumbent Ratings for the O*NET® Generalized Work Activities Questionnaire from Water and Liquid Waste Treatment Plant and System Operators

Five Items with the Largest Difference							
	Job Incumbents			SMEs			
Variable	Total N	Mean	Standard Deviation	Total N	Mean	Standard Deviation	Difference
Level of analyzing data or information	12	4.83	1.53	15	3.80	1.21	1.03
Level of updating/ using relevant knowledge	14	5.43	1.16	15	4.40	0.99	1.03
Importance of performing administrative tasks	16	2.44	1.36	15	3.47	0.83	-1.03
Level of thinking creatively	13	4.08	1.44	15	3.07	1.16	1.01
Level of drafting/ laying-out/ specifying	6	4.17	1.47	10	3.20	0.63	0.97
Five Items with the Smallest Difference							
	Job Incumbents			SMEs			
Variable	Total N	Mean	Standard Deviation	Total N	Mean	Standard Deviation	Difference
Importance of identifying objects/ actions/events	16	3.81	1.05	15	3.80	1.01	0.01
Level of guiding/ directing subordinates	13	3.46	1.39	15	3.47	1.46	-0.01
Importance of updating/using relevant knowledge	16	3.75	1.18	15	3.80	0.68	-0.05
Importance of evaluating information	16	3.94	1.18	15	4.00	1.00	-0.06
Level of developing and building teams	11	3.27	1.27	14	3.21	1.05	0.06

The mean absolute differences between these incumbents and SMEs across all items in a domain are shown in ***Table 3*** for all four questionnaire types. Again, these results suggest that differences in raters are minimal. These findings lead us to make the preliminary conclusion that the differences between job incumbents and SMEs do not contribute substantially to any response bias.

Table 3. Mean Absolute Differences by Questionnaire Type

Questionnaire	Mean Absolute Difference Across All Items
Skills	.39
Generalized Work Activities	.42
Work Context	.39
Knowledge	.49

2.2.2 Item Nonresponse

In the data collected in Wave 1.1, there was no substantial item nonresponse on the individual questions. Across the four questionnaire types, item completion rates ranged from 96.1% to 99.7%. ***Table 4*** presents the range of item nonresponse for each type of question.

Table 4. Range of Item Nonresponse for Each Question Type

Question Type	Range of Item Nonresponse*
Skills—Importance	0.27 – 1.85%
Skills—Level	0.97 – 2.36%
Generalized Work Activities—Importance	0.71 – 2.00%
Generalized Work Activities—Level	1.73 – 3.89%
Work Context	0.30 – 1.70%
Knowledge—Importance	1.18 – 2.28%
Knowledge—Level	1.79 – 3.58%

*These rates are based on raw, unweighted data that was received as of March 11, 2002.

Item nonresponse is not expected to be a source of any meaningful response bias. The questionnaire types that include both Level and Importance scales (Skills, Generalized Work Activities, and Knowledge) present slightly higher rates of nonresponse to the Level scale items. The Level items require respondents to make a somewhat more difficult judgement, utilizing descriptive anchors to determine the level of an item that is required to perform their jobs. This may account for the small differences in nonresponse. Overall, none of the items had a nonresponse rate of more than 3.89%, and there were no patterns based on an item's position within the questionnaire. Furthermore, there did not appear to be any pattern in item nonresponse between the two different modes of administration.

2.2.3 Percentage of Respondents Who Used the Web

Table 5 displays the percent of responding employees who used the Web, by occupational category. *Of those who responded to the questionnaire*, the overall rate of use of the Web to complete the questionnaire was 15%. This percentage is the number of responses received via the Web divided by the total number of questionnaires received (Web or paper). (The 9% return rate reported in ***Section 2.1.3*** is the rate at which questionnaires were returned via the Web out of *all questionnaires mailed.*) Not surprisingly, Computer and Mathematical Occupations employees responded by Web much more frequently than employees in other

occupations. They responded at a rate of 43%, compared to 25% for Life, Physical and Social Science Occupations, the next most frequent Web responders. The occupational category with the lowest level of response via the Web was Personal Care and Service Occupations.

Table 5. Rate of Web Response by Occupation Category

Occupation Category	Total # of Questionnaires Mailed to POCs or Employees	% of Total Responding Via the Web
Computer and Mathematical Occupations	732	43%
Life, Physical, and Social Science Occupations	1,133	25%
Arts, Design, Entertainment, Sports, and Media Occupations	1,058	24%
Legal Occupations	190	24%
Architecture and Engineering Occupations	1,543	23%
Education, Training, and Library Occupations	203	21%
Business and Financial Operations Occupations	1,510	18%
Management Occupations	1,427	17%
Office and Administrative Support Occupations	1,826	12%
Protective Service Occupations	1,041	12%
Healthcare Practitioners and Technical Occupations	3,979	11%
Installation, Maintenance, and Repair Occupations	547	10%
Community and Social Services Occupations	670	10%
Sales and Related Occupations	678	9%
Production Occupations	376	8%
Healthcare Support Occupations	974	8%
Construction and Extraction Occupations	1,695	7%
Personal Care and Service Occupations	440	4%
All Occupations	**20,022**	**15%**

The fairly broad range of Web response across occupation category (40 percentage points) might raise concern regarding response bias if there were reason to suggest that people responded differently according to the mode they selected. However, given the minimal mode effects of O*NET® data collection instruments (***Section 2.1.3***) and minimal cross-modal nonresponse differences, response bias by mode does not pose a significant threat.

3. Effectiveness of Incentives

In Wave 1.1, incentives were offered to the POC, the employee, and the business as a whole to encourage their participation. These incentives, the clock and mousepad/calculator for the POC, the $10 for the employee, and the Toolkit for the business, were decided upon after the Pretest, using a "best design" approach to project response rates for Wave 1.1. The best design assumed that the pretest used the incentives of the O*NET® Toolkit for the establishment and the $10 incentive for the employee, as well as an in-kind gift (comparable to the clock and mousepad/calculator in value) to the POC. As indicated in ***Section 1.1*** of this appendix, the Wave 1.1 response rates, 64% for establishments and 63% for employees, are in line with results

from other, similar surveys. These response rates suggest that the incentive plan used has been reasonably effective. However, a review of the literature leads to the conclusion that response rates could be even better with a revised incentive plan.

The literature on incentives in surveys consistently reports that prepaid cash incentives are the most effective incentive tool at improving response rates (Church, 1993; Kulka, 1994; Singer, Van Hoewyk, Gebler, Raghunathan, & McGonagle, 1999). Church's meta-analysis of 38 studies found that monetary incentives, when prepaid, improve response rates over all other types of incentives. When compared to monetary incentives, non-cash gifts, such as the clock, mousepad/calculator, and the Toolkit, are found in some cases to have a positive impact on response rates but less of an impact than money, and in other cases to have no impact at all on response rates (Church; Dommeyer, 1988).

Furthermore, Kulka (1994), citing Dillman's research on mail surveys (1991), reports that "Hundreds of studies have been conducted, and review after review—both qualitative and quantitative—concludes the importance of financial incentives" (Kulka, p. 7). In addition to typical mail surveys, Kulka argues that according to a smaller body of research, "the greatest potential effectiveness of monetary incentives appears to be in surveys that place unusual demands upon the respondent, require continued cooperation over an extended period of time, or when the positive forces on respondents to cooperate are fairly low" (p. 7).

The O*NET® *Data Collection Program* is a survey that places demands upon the POC. The responsibilities of the POC include

- reading the introductory package to become familiar with the purposes of the study and the role of a POC.
- seeking permission within the company, as necessary, to participate in the O*NET® Program.
- making a roster of all employees at the location who work in up to five different occupations.
- participating in a sampling process that selects up to 15 employees from these occupations and maintaining this sample roster for future reference.
- distributing questionnaires to the sample persons within the company and addressing their questions and concerns about the survey.
- cooperating with several attempts to follow up with the employees, including distributing thank you/reminder cards, replacement questionnaires, and e-mail requests for cooperation.

Further, the POC role is particularly important because the POC can unwittingly function as a gatekeeper, effectively reducing or even eliminating employee response.

Because of their minimal perceived value, the clock and mousepad/calculator appear to be limited in their ability to motivate the POC to complete all of the O*NET® functions. This has the potential to translate into poor cooperation at the employee level. For these reasons, a

revised incentive plan that includes a cash incentive of $20 to the POC is recommended, in addition to an in-kind gift, the $10 employee incentive, and the Toolkit for the business.

4. Effectiveness of Stamps

Wave 1.1 data collection included a split-sample, random assignment experiment to compare the effects of using first class stamps on the return questionnaire envelopes to the effects of using Business Reply Mail (BRM) envelopes. Differences between the two approaches were evaluated in terms of response rate as well as response time (see ***Table 6***). The difference in mean response rate for stamps as opposed to BRM is statistically insignificant.[1] These findings are not consistent with the literature on response rates, which suggests that a 2% to 4% increase in response rate should be expected with the use of stamps as compared to BRM (Dillman, 2000). Little research has been published on the impact of stamps and BRM on the speed of questionnaire return, but Elkind, Tryon, and DeVito's (1986) findings concur with the results of no difference presented here.

Table 6. Results of Stamp/Business Reply Mail Experiment

	Response Rate	Mean Response Time
Business Reply Mail	52.2%	48.1 days
Stamped Envelope	50.3%	47.7 days

Furthermore, an examination of the costs associated with BRM and stamped envelopes yields little difference between the two. For 5,000 returned questionnaires, assuming a reasonable 60% return rate for Wave 2.1, in which 8,333 questionnaires are projected to be mailed, BRM will cost about $2,300. This cost includes a $125 annual permit and a $375 annual accounting fee paid to the U.S. Postal Service, as well as a per piece charge of $0.31 for mailing and a per piece fee of $0.05. To put a $0.34 stamp on 8,333 questionnaires costs $2,833. The $2,833 projection for stamp expenses is fixed, because a stamp must be placed on all questionnaires sent, regardless of the number returned. In contrast, the $2,300 estimate for BRM is variable, because mailing charges and per piece fees are incurred only when questionnaires are returned. If a 72% return rate is anticipated, with 6,000 questionnaires returned, the BRM cost would be $2,660, still below the stamp cost of $2,883. For the BRM cost to exceed the stamp cost, the questionnaire return rate would have to exceed 77%. Given these cost assumptions and a reasonable expected return rate between 60% and 70%, the cost differential for BRM versus stamped envelopes is minimal.

Design differences between the O*NET® data collection design and the standard mail survey as designed by Dillman (2000) may account for the lack of improvement in response rates found by using first-class stamps. Dillman's approach calls for a reminder postcard 1 week after the initial questionnaires are sent. Dillman argues that the effect of the stamp is to make the postcard reminder more effective because respondents are more likely to keep a stamped

[1] The standard errors were not available for this analysis. However, it is safe to assume that there is no significant difference between 50.3 and 52.2, even if we assume a conservative design effect of 2.0.

envelope on hand but throw out a BRM envelope. The Wave 1.1 design did not offer a fair test of Dillman's hypothesis, since the postcards were sent more than a week after the questionnaires and were only sent to sampled employees who had not returned the questionnaire.

While the findings of the split sample experiment do not suggest that BRM envelopes are any less effective than stamped return envelopes, the O*NET® survey designers defer to Dillman's extensive research in this field of survey methodology. Dillman finds stamped return envelopes overwhelmingly more effective than BRM envelopes. After considering the parity of cost, it was decided to implement first-class stamps for Wave 1.2 and beyond. To maximize the potential of the stamped return envelope, the postcards will be sent, as Dillman suggests, after the 7-day call in all future waves.

5. References

Church, A.H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly, 57,* 62–79.

Couper, M.P., Traugott, M.W., & Lamias, M.J. (2001). Web survey design and administration. *Public Opinion Quarterly, 65*, 230–253.

Dillman, D. (1978). *Mail and Telephone Surveys: The Total Design Method.* New York: John Wiley & Sons, NY.

Dillman, D.A. (1991). The design and administration of mail surveys. *Annual Review of Sociology, 17*, 225–249.

Dillman, D.A. (2000). *Mail and Internet Surveys: The Tailored Design Method.* New York: John Wiley & Sons, Inc.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., & Berck, J. (2000). Response rate and measurement differences in mixed mode surveys: Using mail, telephone, interactive voice response, and the Internet. Paper presented at the annual meeting of the American Association for Public Opinion Research, Portland, Oregon.

Dommeyer, C.J. (1988). How form of the monetary incentive affects mail survey response. *Journal of the Market Research Society, 30,* 379–85.

Elkind, M., Tryon, G., & De Vito, A. (1986). Effects of type of postage and covering envelope on response rates in a mail survey. *Psychological Reports, 59,* 279-283.

Griffin, D.H., Fischer, D.P., & Morgan, M.T. (2001). Testing an Internet response option for the American Community Survey. Paper presented at the annual meeting of the American Association for Public Opinion Research, Montreal, Quebec.

Harvey, R.J. (1991). Job Analysis. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology.* Palo Alto, CA: Consulting Psychologists Press.

Krosnick, J.A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213–236.

Kulka, R.A. (1994). The use of incentives to survey "hard-to-reach" respondents: A brief review of empirical research and current practice. Paper presented at the Seminar on New Dimensions in Statistical Methodology, Bethesda, Maryland.

Maurer, T.J., & Tross, S.A. (2000). SME committee vs. field job analysis ratings: Convergence, cautions, and a call. *Journal of Business and Psychology, 14,* 489–499.

Morgeson, F.P., & Campion, M.A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology, 82,* 627–655.

Paxson, M.C., Dillman, D.A., & Tarnai, J. (1995). Improving response to business mail surveys. In B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge, & P.S. Kott (Eds.), *Business Survey Methods*. New York: John Wiley & Sons.

Peterson, N.G., Mumford, M.D., Borman, W.C., Jeanneret, P.R., Fleishman, E.A., Levin, K.Y., Campion, M.A., Mayfield, M.S., Morgeson, F.P., Pearlman, K., Gowing, M.K., Lancaster, A.R., Silver, M.B., & Dye, D.M. (2001). Understanding work using the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology, 54*, 451–492.

Peterson, N.G., Mumford, M.D., Borman, W.C., Jeanneret, P.R., Fleishman, E.A., & Levin, K.Y. (1997). *O*NET Final Technical Report*. Salt Lake City, UT: Utah Department of Employment Security.

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.

Singer, E., John Van Hoewyk, J., Gebler, N., Raghunathan, T., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics, 15,* 217–230.

Tannenbaum, R.J., & Wesley, S. (1993). Agreement between committee-based and field-based job analyses: A study in the context of licensure testing. *Journal of Applied Psychology, 78,* 975–980.

Ware, C. (2000). *Information Visualization: Perception for Design.* San Francisco: Morgan Kaufman.