Appendix A: Results of Statistical Analysis of Pretest

O*NET REPORT

RESULTS OF STATISTICAL ANALYSIS OF PRETEST

Prepared By
Research Triangle Institute
Statistics Research Division

October 2000

Table of Contents

Page

I. Executive Summary 6

A. Recommended Best Design 7

B. List of Key Findings 8

II. Major Results 10

A. Eligibility Analysis 10

1. Background 10

2. Objective 10

3. Methods 10

4. Results 11

5. Conclusions 13

B. Response Rate Analysis 13

1. Background 13

2. Objective 14

3. Methods 14

4. Results 17

a. Employer Level 17

b. Employee Level 23

c. Response Rates by Occupation 29

5. Conclusions 31

a. Employer Level 31

Toolkit Incentive 31

Number of Occupations Requested 32

Sampling Strategy 32

b. Employee Level 33

Non-Monetary Incentives 33

Monetary Incentives 33

Other Considerations 34

Table of Contents (continued)

Page

III. Data Issues 34

A. Non-Respondents 34

1. Background 34

2. Objective 34

3. Methods 34

4. Results 35

a. Employer Level 35

b. Employee Level 40

5. Conclusions 42

B. Outliers 42

1. Background 42

2. Objective 42

3. Methods 43

4. Results 43

5. Conclusions 45

IV. Sampling Issues 46

A. Sample Size 46

1. Background 46

2. Objective 46

3. Methods 46

4. Results 47

5. Conclusions 49

B. Design Effects 51

1. Background 51

2. Objective 51

3. Methods 52

4. Results 52

5. Conclusions 53

Table of Contents (continued)

Page

V. Descriptive Statistics 53

A. Demographics 53

1. Background 53

2. Objective 53

3. Methods 54

4. Results 54

5. Conclusions 55

B. Item Estimates 55

1. Background 55

2. Objective 55

3. Methods 56

4. Results 56

5. Conclusions 57

VI. Concluding Statement 57

VII. Appendices

Appendix V-1 Skills Questionnaire Estimates

Appendix V-2 Generalized Work Activities Questionnaire Estimates

Appendix V-3 Abilities Questionnaire Estimates

Appendix V-4 Work Context Questionnaire Estimates

Appendix V-5 Knowledge Questionnaire Estimates

List of Exhibits

Page

I-1 Methodological Recommendations 7

II-1 Establishment Eligibility Rates by Occupation 12

II-2 Logistic Regression on Employer Response 19

II-3 Interaction between Toolkit and Sampling Approach 20

II-4 Estimated Employer Response Rates for Each Factor 22

II-5 Raw Employee Response Rates by Monetary Incentives 24

II-6 Logistic Regression on Employee Response 25

II-7 Estimated Employee Response Rates for Each Factor Level 27

II-8 Estimated Response Rates for Each Occupation 29

III-1 Number of Non-Responding Employers by Reason and Survey Stage 36

III-2 Employer Non-Response by Industry Groups 37

III-3 Employer Non-Response by Establishment Size 38

III-4 Employer Non-Response by Geography (Zip Code) 39

III-5 Map of Zip Code Areas 39

III-6 Demographics of Employees, by Response Status, by Occupation (SOC Code) 40

III-7 Percent Missing Data for Individual Questionnaires 45

IV-1 SE_{15} Estimates for the Five-Point Scale by Domain 48

IV-2 SE_{15} Estimates for the Seven-Point Scale by Domain 48

IV-3 SE_{10} Estimates and SE_5 Estimates for the Five-Point and Seven-Point Scales, by Domain 49

IV-4 Design Effect Estimates for the Five-Point Scale by Domain 52

IV-5 Design Effect Estimates for the Seven-Point Scale by Domain 52

I. Executive Summary

The *Occupational Information Network (O*NET) Data Collection Program* successfully implemented a pretest survey intended to optimize the O*NET survey methodology and to maximize the efficiency of the data collection program. The primary objective of the pretest was to test variations of several factors for their impact on employer and employee response rates. This report is one in a series describing the O*NET pretest and subsequent analyses of the data collected. It presents the initial data analyses conducted as prescribed in the *OMB Clearance Package Supporting Statement and Data Collection Instruments (March 1999).* Refer to the 1999 OMB Clearance Package for a detailed description of O*NET and the data collection procedures implemented during the pretest.

The primary analysis reports on the results of the pretest experiment to determine factors that improve response rates. Following from this analysis, recommended best design features are presented for future data collection efforts. An additional analysis of establishment eligibility for sampling is also provided. General descriptive statistics, demographic analyses, sampling error specification and non-respondent analyses are included as planned.

Results of the major analyses are presented first. The post hoc eligibility analysis describes differences by occupation and assesses the effect of different sampling strategies. The two-level response rate analysis examines the effectiveness of various incentives for employers and employees. The number of occupations requested in the Data Collection Program, which represents a burden on the employers, was varied in the pretest and analyzed for its effect. Other factors considered for their impact on response rates are the sampling approach used and the size of establishment surveyed.

Issues that can affect data quality are presented next. Reported establishment reasons for non-participation are summarized. Non-response is examined by industry, establishment size and geographic area. At the employee level, age, gender, and race of respondents was compared to that of non-respondents.

Sampling issues are re-evaluated. Pretest data are used to project the sample size needed in future data collection to accurately describe each occupation. The design of the pretest involved unequal weighting and clustering of employees within establishments at the second stage of sampling. The effects of the pretest sampling design on data collection efficiency are reported.

Finally, descriptive statistics are reported for all of the data collected in the pretest. Demographic characteristics of the pretest respondents are summarized by age, gender, race and disability status. Means and standard errors are provided for each questionnaire item for each subject matter domain for each of the 50 occupations that were included in the pretest.

I.A Recommended Best Design

The O*NET pretest informs the next stage of the data collection program with the following methodological recommendations presented in Exhibit I-1.

Exhibit I-1 Methodological Recommendations

Employer Level	Employee Level
Provide toolkit incentive	Do not include video incentive
Request multiple occupations from each establishment	Provide $10 pre-incentive
Use multiple sampling strategies	Use stamped return envelope

I.B List of Key Findings

Eligibility Rates:

1. *The difficulty in locating an establishment that employs the occupation of interest (i.e., eligible for sampling) varied widely across occupations and may be predicted using characteristics of the occupation.*

2. *Eligibility for sampling can be improved in some cases with a targeted approach, and use of multiple sampling approaches in the future may improve data collection efficiency.*

3. *Adequacy of the coverage of an occupation using the targeted approach has not been precisely determined, but the target sampling frames appear to cover roughly 40% to 50% of the workforce for most of the pretest occupations.*

Establishment Response Rates:

4. *Contrary to expectations, employer response rates were higher from smaller establishments.*

5. *Response rates were better from establishments that received the Job Analysis Toolkit, and this incentive is recommended for future use.*

6. *The number of occupations an establishment was asked about (controlled from one to four) did not appear to be related to the employer response rates.*

Employee Response Rates:

7. *At the employee level, inclusion of an O*NET informational video did not affect the response rate and may be eliminated from future use as an unnecessary expense.*

8. *The type of return postage was not found to affect response at a statistically significant level.*

9. *Monetary incentives had a significant impact on employee response rates, and, as expected, pre-incentives were more successful in encouraging response than were post-incentives.*

10. Employee response rates greater than 50% were generated using the $10 pre-incentive only, and the more expensive plans produced indistinguishable improvements in the response rate.

Sampling Issues:

11. Differences were noted between responding and non-responding establishments when examined by industry, geographic region, and number of employees.

12. The sample size analysis confirms the adequacy of 15 responses per descriptor and shows that samples of 10 responses may also be adequate, but shows that samples of five responses would be inadequate.

13. The complex sampling design employed in the pretest generally appears to have improved the efficiency of data collection compared to a simple random sample.

Data Issues:

*14. Unbiased estimates of the item means and standard errors are now available for researchers and other data users for the 50 occupations studied in the O*NET pretest.*

II. Major Results

II.A Eligibility Analysis

II.A.1 Background

The pretest was originally designed to collect data for all occupations from a single sample of establishments. This general approach used composite size measures to simultaneously select a national sample of establishments that contained employees in 50 occupations (referred to in this report as the general sample). After this general sampling had begun in the pretest, low eligibility rates reduced the efficiency of collection for some occupations. A more targeted sampling approach was devised in an effort to improve the "hit rate" and reduce the number of calls required to find establishments with the occupations of interest. The targeted samples were used in a dual frame survey design in combination with the general sample. Multiple targeted samples were taken from specific industry groups expected to contain a higher percentage of establishments with employees in a given occupation. Targeted samples were drawn first for the 16 occupations with the lowest eligibility rates (Target 1 sample) and subsequently for another 16 occupations (Target 2 sample). Future data collection will also use a combination of the general and targeted sampling approaches.

II.A.2 Objective

The purpose of this task was to describe differences in eligibility rates between the two sampling approaches and between occupations.

II.A.3 Methods

An establishment was eligible to participate in the study if it was sampled and contained employees in the occupation(s) requested. An establishment was

considered ineligible if a representative of the establishment was contacted and declared that the establishment did not have any employees in the occupations sampled. If an establishment was sampled and eligibility status could not be determined, the establishment was assumed to be eligible and was classified as a non-responder. This occurred for 7% of the establishments contacted. Once employees were selected, no subjects were disqualified from the study. Consequently, we had 100% eligibility at the employee level.

Although the primary sampling unit was the establishment, eligibility records were noted for all establishment-by-occupation combinations. For a particular establishment, one to four occupations were sampled. For example, suppose two occupations were sampled for a particular establishment, and the establishment contained employees in one occupation but not the other. The establishment was considered eligible for one establishment-by-occupation combination and ineligible for the other.

II.A.4 Results

Of the employers contacted, 30% (2,055/6,881) had employees in at least one of the occupations requested. When establishment eligibility was counted separately for each occupation requested, the rate was 31% (3,071/10,017), which varied by occupation from 7% to 72% (see Exhibit II-1). Among the 18 occupations that were only sampled using the general probability-based sampling approach, eligibility was 37%. Among the 32 occupations that were selected for targeted sampling (due to low eligibility with the general approach), better overall rates were obtained in the targeted versus the general sample (39% vs. 15%, chi-square p=0.001). Overall eligibility rates for each of the ONET pretest occupations are shown in Exhibit II-1.

Exhibit II-1 Establishment Eligibility Rates by Occupation

Occupation	Eligibility Rate(%)
Aerospace Engineers	24
Architects, Except Landscape and Marine	39
Bailiffs	16
Bartenders	32
Bicycle Repairers	22
Brokerage Clerks	26
Bus Drivers	39
Camera and Photographic Equipment Repairers	26
Chemical Engineers	21
Chemists, Except Biochemists	21
Child Care Workers	27
Civil Engineering Technicians	20
Civil Engineers, Including Traffic	24
Correction Officers and Jailers	47
Dental Assistants	45
Dental Hygienists	37
Drivers/Sales Workers	29
Food-Service Managers	72
Geological and Petroleum Technicians	15
Guards and Watch Guards	37
Hotel Desk Clerks	66
Human Services Workers	30
Industrial Engineering Technicians	15
Industrial Engineers, Except Safety	32
Industrial Truck and Tractor Operators	42
Insurance Adjusters/Examiners/Investigators	37
Insurance Appraisers, Auto Damage	14
Interior Designers	24
Landscape Architects	30
Legal Secretaries	70
Loan and Credit Clerks	46
Locksmiths and Safe Repairers	45
Lodging Managers	40
Management Analysts	19
Marine Architects	22
Marine Engineers	9
Mechanical Engineering Technicians	19
Millwrights	29
Motorcycle Mechanics and Repairers	40
Opticians, Dispensing and Measuring	53
Paralegals and Legal Assistants	62
Pest Controllers and Assistants	46
Petroleum Engineers	16
Sales Agents, Securities and Commodities	38
Salespersons, Parts	35
Shipping, Receiving, and Traffic Clerks	48
Technical Writers	21
Travel Agents	63
Underwriters	46
Water Treatment Plant and System Operators	7

II.A.5 Conclusions

The difficulty in locating an establishment that employs the occupation of interest (i.e., eligible for sampling) varied widely across occupations. Since the 50 occupations sampled in the pretest were not selected randomly, these rates may or may not be representative of other occupations.

In the O*NET general sample, efforts to improve survey efficiency using stratification by industry groups may have resulted in lower eligibility rates than would be expected if this restriction were not applied. Eligibility for sampling can be improved in some cases with a targeted approach, and the use of multiple sampling approaches in the future may improve data collection efficiency. Availability in the near future of more detailed employment/industry information will also likely lead to increased eligibility rates. The targeted sampling frames appeared to cover roughly 40% to 50% of the workforce for most of the pretest occupations.

II.B Response Rate Analysis

II.B.1 Background

The *O*NET Data Collection Program* recognizes the importance of obtaining high response rates, in order to control non-sampling errors and minimize any biases due to non-response. The O*NET Pretest was designed and conducted to provide information on the impact of several proposed survey procedures on response rates. Specifically, participation incentives for the employers and employees were tested to determine which interventions had the greatest effect on response rates. In the analysis below, the pretest data were used to model the effects of the experimental factors in order to demonstrate which tested levels

of the experimental factors affected the employer and employee response rates, and how much the rates changed when the factors were varied. The results presented identify which procedural improvements can be expected to lead to improved response rates in the upcoming *O*NET Data Collection Program*. (Refer to the *O*NET Data Collection Program Survey Pretest OMB Clearance Package Supporting Statement and Data Collection Instruments, March 1999,* for more background details.)

The experimental nature of the pretest should be considered when reviewing response rates from this study. Factors were intentionally varied to ascertain their effect on response, and it was expected that some combinations of factors would be more successful in generating high response rates. In addition, data were collected in a limited amount of time in order to proceed with analysis and planning for full-scale application of the best design.

II.B.2 Objective

The purpose of this task was to determine which combination of experimental factors produced the highest rate of employer and employee response.

II.B.3 Methods

An establishment was considered a responder if it provided a list of employees in at least one of the requested occupations. An employee was considered a responder if a questionnaire was returned.

Sampling design factors to be considered for the employer response rate were sampling approach (general vs. targeted) and size of establishment. Establishment level experimental variables were 1) the employer incentive (i.e., Job Analysis Toolkit) and 2) the number of occupations/employees to be sampled from each establishment. Experimental variables relating to the

employee level response rate were 1) outreach to the employees (O*NET video), 2) pre- and post-incentives (cash), and 3) return envelope postage. For more details, consult the OMB Clearance Package for the *O*NET Data Collection Program* Survey Pretest.

Frequencies for each effect by outcome are presented with the results of univariate chi-square tests. Separate logistic regression models were fit at the employer and employee levels to examine response rates adjusted for all factors and to quantify the relationship between the design and experimental factors and response rates. Models were stratified to control for the effect of the sampling group and for the differential allocation by establishment size within the general sampling group. The establishment-size factor was collapsed into four groups based on the distribution: 1-9, 10-49, 50-249 and 250+ employees. SUDAAN software was used to account for the correlation of employee responses within an establishment.

The design of the pretest led to a one-to-one correspondence between the number of occupations and the number of subjects to be sampled from the establishment. However, occupations were selected for sampling with replacement, so some establishments had the same occupation selected multiple times. To determine the effects of establishment burden, initial models contained both a variable for the number of employees to be selected and the number of unique occupations requested. Because the number of different employees requested per occupation was not a significant factor in response rate, it was dropped from subsequent modeling, and establishment burden was represented by the number of unique occupations requested only. Also, the targeted samples typically requested only one occupation from an establishment. An additional employer-level model was run on the general sample only due to concerns about this group imbalance with regard to the effect of asking for multiple occupations.

For both the employer- and employee-level analyses, all two-way interaction terms were tested individually with the main effects. In the employer-level analysis, the group-by-toolkit interaction term was retained. It is important to note that the toolkit was not available during the early stage of the pretest, and only 86/865 (10%) of the establishments in the general sample received toolkits. Due to concerns about the effect of this imbalance of toolkit allocation across sampling groups, the model was rerun with only the targeted samples to clarify the toolkit effect.

Additional frequency tables were generated and models were rerun separately by toolkit to examine the group effect further. The targeted samples were a non-random subset of the population represented by the general sample, since only selected occupations were included. Establishments containing these occupations may have responded differently to inclusion of the toolkit from establishments in the general sample.

Beta estimates were generated from the primary model for each factor level. Estimated response rates and 95% confidence intervals are given for every combination of experimental factors adjusted for other variables in the model. Predefined contrast statements to test specific questions of interest regarding number of occupations asked for at the establishment level and types of incentives offered at the employee level are also presented. No adjustments were made for multiple comparisons.

Logistic models were refit for both employer and employee observations using their respective sampling weights. Weights were applied as the inverse of the probability of selection for each sampling unit at the employer and employee level. Results from weighted analyses are typically presented for survey studies and are used for inference to the entire sampling frame. For analysis of the pretest experimental design, the unweighted analysis was preferred because it provided estimates of how the sampling design factors and experimental

variables actually affected employer and employee response rates in the pretest population. Comparisons of the various incentive plans should be made from the unweighted analyses.

The logistic regression models were used to obtain estimates of what the response rates would be if the recommended factor levels were applied to the entire sample (toolkit incentive provided and multiple occupations requested for each employer, business reply envelope [BRE] and $10 pre-incentive provided to employees). Estimated employer and employee response rates are presented separately for every occupation in the pretest to show the wide variability across occupations. Because the pretest data were obtained via a two-stage sampling process, overall response rates were then obtained by multiplying employer and employee response rates. Consequently, overall response rates from this study can be expected to be lower than in situations where a one-stage sampling strategy is feasible.

II.B.4 Results

II.B.4.a Employer Level

6,881 employers were screened. At the screening, 2,421 of the employers claimed that they had employees in at least one the occupations requested for sampling and were therefore classified as eligible. However, at recruitment and sampling, 366 of the employers claimed that they had no employees in the occupations. Consequently, they were reclassified as ineligible. Of the 2,055 remaining eligible employers, 1,149 (56%) responded by providing lists of employees.

Unadjusted chi-square tests on the raw data indicated significant response rate variation across each design and experimental variable. The highest rate of establishment response was found in the second targeted sample taken (70% vs.

62% in first targeted sample and 40% in general sample, p=0.001). Better rates occurred in smaller establishments relative to larger ones (1-9 employees, 70%; 10-49, 57%; 50-249, 43%; 250+,: 34%; p=0.001). Only 10% (86/865) of employers in the general sample received a toolkit versus 36% (423/1,190) in the targeted samples. Employers that received toolkits responded better than those that did not (68% vs. 52%, p=0.001).

Overall, response rates were as follows: 65% for employers that were asked for one occupation, 42% for employers that were asked for two occupations, 43% for employers that were asked for three occupations, and 48% for employers that were asked for four occupations. Note that none of the employers in the targeted samples were asked for more than two occupations, and less than 15 percent of employers in the pretest were asked about three or four occupations. In the general sample only, 39% of employers that were asked for one occupation responded versus 37% of those that were asked for two occupations, 43% of those that were asked for three occupations, and 48% of those that were asked for four occupations (p=0.2).

The initial logistic modeling included the design and experimental variables plus all two-way interaction terms. When each interaction term was tested individually with the main effects, only the group by toolkit term was significant (p=0.006). This interaction term was retained in further modeling. The primary logistic model included sampling group, establishment size, toolkit, number of occupations requested, plus the group-by-toolkit interaction term. Results are shown in Exhibit II-2.

Exhibit II-2 Logistic Regression on Employer Response

Independent Variables	Beta Coefficient	SE Beta	t-test B=0	p-value t-test
Intercept	0.2625	0.1824	1.4389	0.1503
Sampling Group				
General	0.0000	0.0000	.	.
Target 1	0.5732	0.1797	3.1899	0.0014
Target 2	0.5380	0.2156	2.4956	0.0127
Establishment Size				
1-9	0.0000	0.0000	.	.
10-49	-0.3714	0.1328	-2.7964	0.0052
50-249	-0.7384	0.1545	-4.7787	0.0000
250+	-1.0021	0.1645	-6.0919	0.0000
Toolkit				
No	0.0000	0.0000	.	.
Yes	0.6819	0.2366	2.8824	0.0040
Number Occupations Requested				
1	0.0000	0.0000	.	.
2	-0.2252	0.1602	-1.4058	0.1599
3	0.1273	0.2033	0.6265	0.5311
4	0.3149	0.2531	1.2444	0.2135
Group by Toolkit Interaction				
General, No	0.0000	0.0000	.	.
General, Yes	0.0000	0.0000	.	.
Target 1, No	0.0000	0.0000	.	.
Target 1, Yes	0.1870	0.3210	0.5825	0.5603
Target 2, No	0.0000	0.0000	.	.
Target 2, Yes	-0.6973	0.3086	-2.2594	0.0240

After adjusting for all other variables, establishment response rates varied by establishment size (p=0.0001), with better rates occurring in smaller establishments.

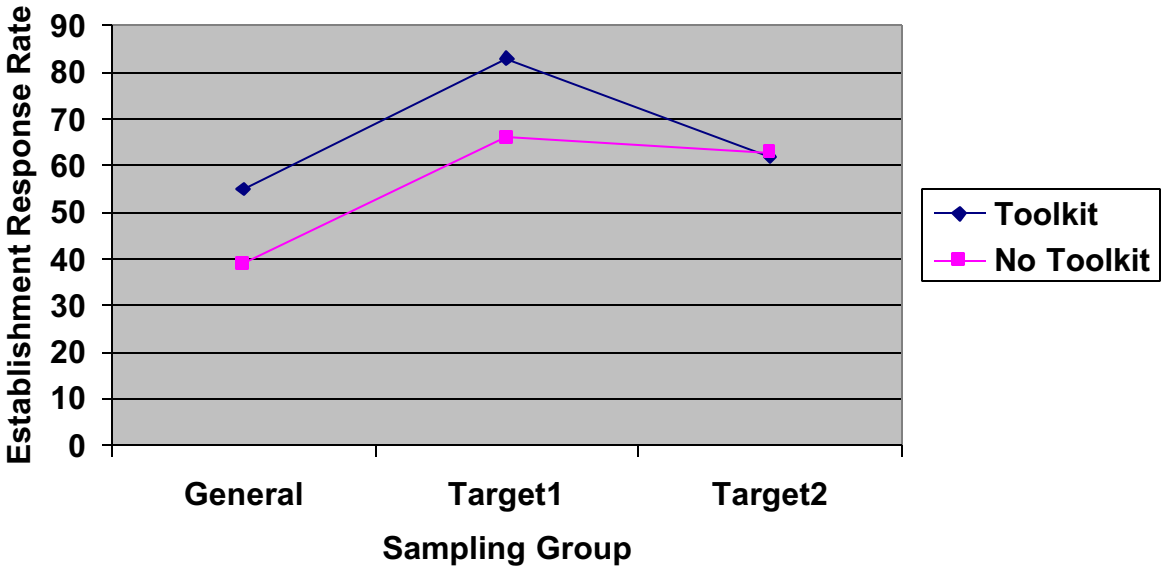
The number of unique occupations had a significant effect on the establishment response rate (p=0.04), although individual comparisons of asking for two, three or four occupations were not different from asking for one. Additional contrasts showed that the combined effect of asking for more than one occupation was not statistically different from asking for just one (p=0.7). Perhaps because the effect of asking for two occupations was in the negative direction even though not

statistically significant, asking for three or four occupations was significantly better than asking for one or two (odds ratio=1.95, p=0.03).

To corroborate these results, the model was rerun on the general sample only, since the targeted samples typically requested one or two occupations only. The effect of establishment size was unchanged from the original model (p=0.0001). The overall effect of the number of different occupations requested was no longer statistically significant (p=0.08), but the combined effect of asking for three or four was still better than asking for one or two (p=0.01).

In the primary model, the effects of the toolkit and sampling group on response rates were complicated by a significant interaction (p=0.006), as shown in Exhibit II-3.

Exhibit II-3 Interaction Between Toolkit and Sampling Approach



It appears that providing the toolkit compared to not providing the toolkit resulted in improved establishment response rate. The small number of establishments in the general sample who received the toolkit responded to it in a positive manner (odds ratio=1.97, 95% confidence interval : 1.2, 3.1). However, in the two targeted samples where appropriate toolkit allocation occurred, the effect was

inconsistent. In the first targeted sample, the odds ratio for toolkit was 2.3 (confidence interval: 1.5, 3.6); in the second targeted sample, the odds ratio was 0.98 (confidence interval: 0.6, 1.4).

Establishment response rates varied by sampling group, with generally better rates occurring in the targeted samples. However, when models were rerun separately by toolkit, we found that the better response rates in the Target 1 sample compared to the general sample were significant only when the toolkit was not included (with toolkit odds ratio: 1.3, 95% confidence interval: 0.5, 3.5; without toolkit odds ratio: 1.9, 95% confidence interval: 1.3, 2.7). The Target 2 sample also had better response rates than the general sample when the toolkit was not included; the response rate for this group was lower than the general sample when the toolkit was included, though the difference was not significant (with toolkit odds ratio: 0.5 95% confidence interval: 0.2, 1.3; without toolkit odds ratio: 1.8, 95% confidence interval: 1.2, 2.8). The model that only includes employers who received the toolkit may have lacked power to detect the group differences in response rates. These results suggest that the apparent difference in establishment response rates attributed to the sampling approaches can be explained at least in part by the lack of toolkit distribution in the general sample.

Following are estimated employer response rates and 95% confidence intervals from the primary model for all possible combinations of toolkit and number of unique occupations requested given the combination of sampling approaches and establishment sizes from the pretest. Inclusion of the toolkit resulted in the highest response rates. Asking for multiple occupations did not necessarily improve response rates, but asking for three or four occupations was not a disadvantage compared to asking for just one occupation. These factor levels are shaded in Exhibit II-4 and are recommended for future O*NET planning.

Exhibit II-4 Estimated Employer Response Rates for Each Factor

Combination of Factors Toolkit / Number of Occupations	Estimated Response Rate (%)	95% CI
No / 1	53	53, 54
No / 2	48	47, 48
No / 3	56	55, 58
No / 4	61	58, 63
Yes / 1	69	67, 71
Yes / 2	64	62, 66
Yes / 3	72	69, 74
Yes / 4	75	73, 78

Undue emphasis should not be placed on the last two lines of this exhibit due to the small sample size in the cells where toolkit was given when three (n=19) or four (n=14) occupations were requested. The estimate of the overall establishment response rate when toolkit is given weighted across all levels of occupations requested is 69% (95% confidence interval: 67, 70).

When this analysis was repeated on weighted data, design effects across the factors ranged from 2 to 7, indicating that clustering led to a loss of efficiency. Consequently, none of the main effects nor the interaction term were significant. The direction of effect was consistent with the unweighted analysis for sampling group, establishment size and toolkit. The effect of number of occupations requested reversed direction with better response rates seen when employers were asked for only one occupation. These results are presented for completeness; inferences from the pretest experiment are most appropriately made from the unweighted analysis.

II.B.4.b Employee Level

Overall, 1,662/3,766 (44%) of the employees who were sent questionnaires responded. No difference in raw response rates was detected for type of questionnaire the employee was asked to complete (A, 45%; B, 46%; C, 41%; D, 45%; E, 43%; p=0.3). Unadjusted employee response rates ranged across occupations from 20% to 79% (median=45%). Since the pretest occupations were selected to represent high-growth industries rather than as a random sample, it is not known whether this variation in employee response rates is representative.

Unadjusted chi-square tests on the raw data indicated significant employee response rate variation across each design and experimental variable except for the video. The highest rate of employee response came from the general sample (47% vs. 44% in first targeted sample and 40% in second targeted sample, p=0.002). Better rates occurred in larger establishments (1-9 employees, 44%; 10-49, 39%; 50-249, 50%; 250+, 47%; p=0.001). No difference in response rates was detected between employees who received the video and those who did not (43% vs. 45%, p=0.4). A slightly larger difference in response rates was detected for those receiving stamped (46%) versus business reply (43%) return envelopes was statistically significant at p=0.04. Raw response rate data for the monetary incentives are shown in Exhibit II-5 (p=0.001).

Exhibit II-5 Raw Employee Response Rates by Monetary Incentives

Incentive	Raw Response Rate
Pre $0, Post $0	125/396 (32%)
Pre $0, Post $10	174/473 (37%)
Pre $0, Post $20	209/566 (37%)
Pre $5, Post $0	303/684 (44%)
Pre $5, Post $10	241/477 (51%)
Pre $10, Post $0	317/621 (51%)
Pre $10, Post $10	293/549 (53%)

None of the two-way interactions were significant when tested individually with the main effects. The primary logistic model included terms for sampling group, establishment size, video, postage, and monetary incentive. Results from this model are shown in Exhibit II-6.

Exhibit II-6 Logistic Regression on Employee Response

Independent Variables	Beta Coefficient	SE Beta	t-test B=0	p-value t-test
Intercept	-0.6918	0.2380	-2.9072	0.0037
Sampling Group				
General	0.0000	0.0000	.	.
Target 1	-0.0234	0.1683	-0.1392	0.8893
Target 2	-0.1914	0.1726	-1.1090	0.2677
Establishment Size				
1-9	0.0000	0.0000	.	.
10-49	-0.2482	0.1546	-1.6056	0.1087
50-249	0.1933	0.1976	0.9783	0.3281
250+	0.0196	0.2002	0.0980	0.9219
Video				
No	0.0000	0.0000	.	.
Yes	-0.0782	0.1174	-0.6659	0.5056
Postage				
BRE	0.0000	0.0000	.	.
Stamped	0.1006	0.1181	0.8516	0.3946
Monetary Incentive				
$ 0/0	0.0000	0.0000	.	.
$ 0/10	0.2228	0.2297	0.9701	0.3322
$ 0/20	0.2501	0.2298	1.0886	0.2766
$ 5/0	0.5383	0.2285	2.3559	0.0187
$ 5/10	0.8049	0.2340	3.4391	0.0006
$10/0	0.8547	0.2204	3.8788	0.0001
$10/10	0.9022	0.2272	3.9703	0.0001

After adjustment for other variables in the model, employee response rates were not statistically different in the targeted samples compared to the general sample (p=0.5). Establishment size did not have a significant effect on employee response rates (p=0.12).

Neither the video (p=0.5) nor the type of postage (p=0.4) on the reply envelope had statistically significant effects on the employee response rate. Note that the small effect that was produced was in the negative direction for inclusion of the

video and in the positive direction for the stamped envelope versus the business reply envelope (BRE).

The cash incentive had a significant effect on the employee response rate ($p<0.0001$). While neither the post$10 nor post$20 provided better response rates than no incentive at all, each of the other incentive plans did. Pairwise comparisons were not made for each of the total dollar amounts, since total compensation was confounded with the pre and post provisions of each plan.

A comparison of the two plans that provided pre-incentives only ($5 or $10) to the two plans that provided post-incentives only ($10 or $20) showed better response rates from the pre-incentives ($p=0.002$) even though the total dollar amount was higher in the post-incentive plans. Twenty dollars split into a pre and post-incentive was better than $20 given as a post-incentive ($p=0.002$).

The pre$10/post$10 plan had the highest odds ratio compared to no incentive (odds ratio= 2.5, 95% confidence interval: 1.6, 3.9). No difference in employee response rate was detected between this plan and the cheaper alternatives (pre$5/post$10 [$p=0.7$], pre$10 [$p=0.8$], pre$5 [$p=0.08$]). The least expensive of these other plans (pre$5) was 1.7 times more likely to get a response than no incentive at all (odds ratio=1.7, 95% confidence interval: 1.1, 2.7). No statistical difference in response rates was detected for the pre$10 compared to the pre$5 ($p=0.12$). However, the lack of power to detect a difference between these better-performing plans may have been an artifact of the study design and sample size. An estimated 8% improvement in response rate occurred by providing a $10 pre-incentive compared to a $5 pre-incentive. Additional improvement was gained by adding a post-incentive (compare pre$5 to pre$5/post$10 and pre$10 to pre$10/post$10 in Exhibit II-7), though the amount of improvement was dependent on the amount of pre-incentive included.

Following are estimated employee response rates and 95% confidence intervals for all possible combinations of video, postage type, and cash incentive, given the combination of sampling approaches and establishment sizes from the pretest. The four incentive plans that included a pre-incentive resulted in the highest employee response rates. These plans (using stamped mail and no video) are shaded in the Exhibit II-7 and are recommended for consideration in future O*NET planning.

Exhibit II-7 Estimated Employee Response Rates for Each Factor Level

Combination of Factors Video / Postage / Incentive	**Estimated Response Rate (%)**	**95% CI**
No / Stamped / $0, $0	33	32, 35
No / Stamped / $0, $10	39	37, 40
No / Stamped / $0, $20	39	38, 41
No / Stamped / $5, $0	46	45, 48
No / Stamped / $5, $10	53	51, 54
No / Stamped / $10, $0	54	53, 55
No / Stamped / $10, $10	55	54, 57
No / BRE / $0, $0	31	30, 33
No / BRE / $0, $10	36	35, 38
No / BRE / $0, $20	37	35, 38
No / BRE / $5, $0	44	42, 45
No / BRE / $5, $10	50	49, 52
No / BRE / $10, $0	52	50, 53
No / BRE / $10, $10	53	51, 54
Yes / Stamped / $0, $0	32	30, 33
Yes / Stamped / $0, $10	37	35, 38

Exhibit 11-7 (continued)

Combination of Factors Video / Postage / Incentive	Estimated Response Rate (%)	95% CI
Yes / Stamped / $0, $20	37	36, 39
Yes / Stamped / $5, $0	44	43, 46
Yes / Stamped / $5, $10	51	49, 52
Yes / Stamped / $10, $0	52	51, 53
Yes / Stamped / $10, $10	53	52, 55
Yes / BRE / $0, $0	30	28, 31
Yes / BRE / $0, $10	34	33, 36
Yes / BRE / $0, $20	35	34, 36
Yes / BRE / $5, $0	42	40, 43
Yes / BRE / $5, $10	48	47, 50
Yes / BRE / $10, $0	50	48, 51
Yes / BRE / $10, $10	51	49, 52

The analysis on weighted data was primarily driven by the general sample. The general sample was designed to emphasize larger establishments, and employees in larger establishments were weighted more heavily than those in smaller establishments. Employees in the general sample represented 40% of the pretest but accounted for 86% of the weighted analysis.

As in the establishment analysis, design effects across the factors ranged from 2 to 7, indicating a loss of efficiency due to clustering. In a main effects model on the weighted employee data, the direction of effect for establishment size was consistent with the unweighted analysis. The effects of sampling group and postage reversed direction, but neither was statistically significant. The video effect reversed direction and was significant at p=0.02. The monetary incentive remained a significant factor in employee response rate (p=0.001), but a

difference from no incentive at all was only detected for the pre$5/post$10 plan (odds ratio= 5.9, 95% confidence interval: 1.05, 32.9).

II.B.4.c Response Rates by Occupation

For each of the 50 occupations in the pretest, estimates of response rates were calculated for an establishment of median size receiving a toolkit averaged over any number of occupations requested and for employees who received no video, a business reply envelope, and a $10 pre-incentive. These two estimates were then multiplied together to get an estimated overall response rate for each occupation (see Exhibit II-8).

Overall estimated response rates under the conditions of the best factors in the pretest experiment ranged from 13% to 59% across the 50 occupations. The effect of the two-stage strategy and experimental nature of the pretest should be considered when viewing these results.

Exhibit II-8 Estimated Response Rates for Each Occupation

Occupation	Title	Estimated Employer Response Rate (%)	Estimated Employee Response Rate (%)	Overall Estimated Response Rate (%)
63023	Baliffs	84	70	59
66002	Dental Assistants	90	59	53
22308	Landscape Architects	80	65	52
53305	Insurance Appraisers, Auto Damage	62	81	51
28305	Paralegals and Legal Assistants	68	73	50
32908	Dental Hygienists	71	68	48
55102	Legal Secretaries	61	77	47
95002A	Water Treatment Plant and System Operators	71	64	45
32514	Opticians	79	57	45
34005	Technical Writers	72	60	43
22128	Industrial Engineers, Except Safety	68	62	42
22121	Civil Engineers, Including Traffic	63	66	42
21102	Underwriters	49	85	42
49014	Salespersons, Parts	75	56	42
85914	Camera and Photographic Equipment Repairers	81	51	41

Exhibit II-8 (continued)

Occupation	Title	Estimated Employer Response Rate (%)	Estimated Employee Response Rate (%)	Overall Estimated Response Rate (%)
24105	Chemists, Except Biochemists	73	56	41
21905	Management Analysts	49	83	41
58028	Shipping, Receiving, and Traffic Clerks	60	67	40
53302	Insurance Adjusters, Examiners, and Investigators	57	71	40
65005	Bartenders	75	53	40
63017	Correction Officers and Jailers	58	68	39
85923	Locksmiths and Safe Repairers	65	59	38
22305	Marine Architects	62	62	38
22111	Petroleum Engineers	74	50	37
22102	Aerospace Engineers	63	58	37
85951	Bicycle Repairers	76	47	36
27308	Human Services Workers	61	57	35
22302	Architects, Except Landscape and Marine	68	51	35
22138	Marine Engineers	65	51	33
85123B	Millwrights	58	57	33
85308	Motorcycle Repairers	85	39	33
53808	Hotel Desk Clerks	70	46	32
97947	Industrial Truck/Tractor Operators	60	52	32
15026A	Lodging Managers	71	44	31
68038	Child Care Workers	62	49	30
97108	Bus Drivers	63	46	29
15026B	Food-Service Managers	67	42	28
43021	Travel Agents	56	48	27
53121	Loan/Credit Clerks	41	65	27
24511	Geological and Petroleum Technicians	59	45	26
22114	Chemical Engineers	58	45	26
22508	Industrial Engineering Technicians and Technologists	71	36	25
67008	Pest Controllers and Assistants	62	40	25
53128	Brokerage Clerks	53	46	24
22502	Civil Engineering Technicians	79	30	24
97117	Drivers/Sales Workers	43	52	22
34041	Interior Designers	69	29	20
63047	Guards and Watchguards	46	38	18
43014	Sales Agents, Securities and Commodities	26	63	16
22511	Mechanical Engineering Technicians and Technologists	41	31	13

A small negative correlation was noted between employer and employee response rates when compared by occupation (r= -.36, p=0.01). This result led to a further examination of the relationship between responding establishments and their non-responding employees.

The probability of observing employers with zero responding employees (stratified by the number of employees sampled at each establishment) is extremely small. For the purposes of this analysis, the assumption was made that if zero employees of an employer returned the survey, then the point-of-contact (POC) for that employer might have failed to distribute the survey forms. This occurrence is referred to as POC non-response.

The observed rates of zero employee responses were compared to the expected rates of response based on number of employees sampled at each establishment and the overall response rate. This calculation gave the probability that the forms were not distributed as 32%. The probability that an employee responded to the survey contingent on the condition that the POC distributed the forms is 69%.

II.B.5 Conclusions

II.B.5.a Employer Level

Toolkit Incentive

Providing the toolkit generally improved response rates from employers compared to not providing the toolkit. It is difficult to quantify the effect precisely due to significant differences in how the three pretest samples reacted to this incentive. Further examination of the occupations sampled in Target 1 versus Target 2 may be helpful in discovering inherent differences that would explain why the toolkit was more successful in one group than in the other.

Number of Occupations Requested

Response rates were not adversely affected by requesting multiple occupations from a single establishment. This information will be used to devise a sampling strategy that increases establishment eligibility by requesting multiple occupations. Analysis results were not adjusted if an establishment responded with fewer than the number of occupations requested. For example, an establishment may have been asked about four different occupations but, only one of the occupations was available within the establishment. This situation would tend to inflate the estimates obtained from the modeling when requesting multiple occupations from an establishment.

Sampling Strategy

The multiple sampling strategy that was implemented in the pretest after the design phase in an effort to improve eligibility rates also had an effect on establishment response rates. Reasons for improved employer response rates in the targeted samples are not immediately obvious. Three differences among the sampling groups should be noted:

1) The targeted samples had a much higher percentage of small sized establishments, which tended to have better response rates than the larger establishments.

2) The employers in the targeted samples were typically asked only for one occupation.

3) Very few employers in the general sample received a toolkit. This resulted in a design bias against the general sample, since the toolkit improved establishment response rates

In conclusion, asking for several occupations did not appear to have a negative effect on response rates. The toolkit is a recommended incentive at the establishment level, although it may be useful to examine further conditions

under which it is most valuable. Contrary to what was expected, smaller establishments had a higher response rate than larger ones.

II.B.5.b Employee Level

Non-Monetary Incentives

Inclusion of a video was not a statistically significant factor in response, but the small effect was in the negative direction. For planning purposes, this result suggests that the video is an unnecessary expense. Type of postage on the reply envelope had a positive but statistically insignificant effect. Since type of postage was not significant for response rate, a decision regarding whether to use stamps or BRE in the future may be based on other factors such as ease of handling and expense. In the unweighted analysis, employee response rate did not vary by sampling group or by establishment size.

Monetary Incentives

As was expected, pre-incentives were more successful in encouraging response than were post-incentives. From the unweighted analysis, the pre$5 plan is the cheapest monetary incentive that resulted in better response rates than no incentive at all. After adjusting for the other factors, this plan was statistically indistinguishable from the more expensive plans. The $5 pre-incentive plan (with no video and BRE) can be expected to generate a 44% employee response rate, comparable to the overall pretest rate. However, employee response rates greater than 50% were generated by increasing the pre-incentive to $10, and this plan is expected to be the most cost-effective approach.

Weighting in the employee analysis led to different conclusions regarding the effectiveness of the video and a more complex scenario for how the monetary incentive might play out in a larger population. This result acts as a reminder that the pretest was performed on a non-random sample of occupations. Estimates of

response rates from the pretest experiment at both the employer and employee level are dependent on the conditions under which this sample was worked.

Other Considerations

Estimates of employer and employee response rates were generated based on the recommended combination of incentives from the pretest. These estimates varied widely across the pretest occupations. Analysis of employee response by employer suggests that as many as 32% of the responding establishments may have failed to distribute the survey forms. Further efforts to encourage POC compliance with the survey protocol may result in higher employee response rates.

III. Data Issues

III.A Non-Respondents

III.A.1 Background

It is important to know the reasons why some of the sampled employers and employees did not participate in the O*NET pretest, to improve cooperation in future survey replications.

III.A.2 Objective

Understanding the reasons for non-cooperation at the establishment and employee levels was the objective of the research described in this section.

III.A.3 Methods

The reasons employers gave for not cooperating with the survey are presented in this section. The incidence of non-cooperation is compared by industry,

establishment size and zip code. At the employee level, respondents are compared to non-respondents by age, gender and race.

III.A.4 Results

III.A.4.a Employer Level

Employers gave a variety of reasons why they were unable or unwilling to participate in the pretest survey. The employers could decline participation at any of three stages during the survey process:

- Stage 1—initial screening to determine whether or not the establishment was eligible to participate in the survey
- Stage 2—recruiting the establishment representative (POC) to prepare lists of the employees in selected occupations
- Stage 3—selecting a sample of employees from the occupation-specific lists that had been constructed.

When employers declined to participate and the survey specialists could not convince them otherwise, the representatives were asked why. The reasons given in response to this question by the stage in the sampling process at which the establishments became non-respondents are shown in Exhibit III-1. Aside from flat refusals and those who said they were just not interested in the survey, many of the reasons given relate to being too busy or not having enough time to complete the survey process. The establishments that could not be contacted after repeated attempts were included as non-respondents in this tabulation. The numbers in Exhibit III-1 are based on hand tabulations of survey specialist comments and differ slightly from the counts presented in other exhibits in this report.

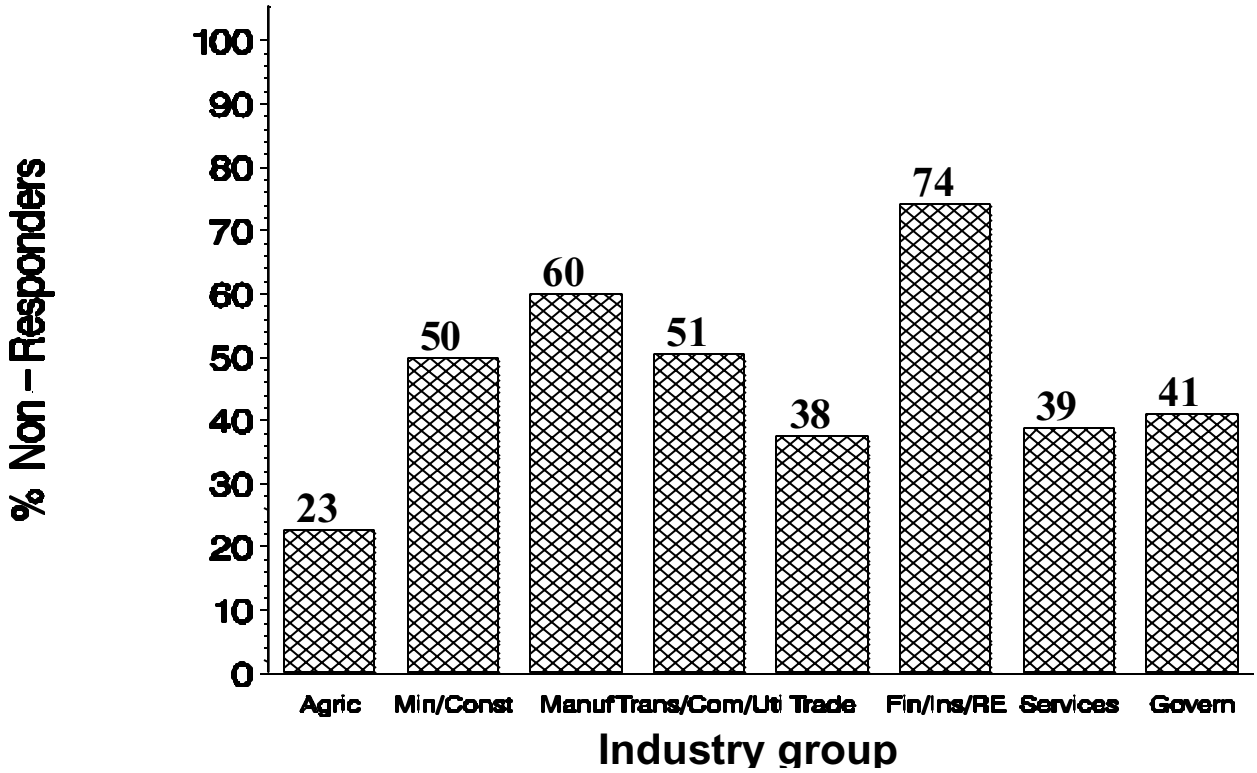
Exhibit III-1 Number of Non-Responding Employers, by Reason and Survey Stage

Reason	Screening	Recruiting	Sampling	Total
Out of Time	39	47	14	100
No Contact	81	169	0	250
Too Busy	74	218	18	310
Not Interested	163	240	10	413
Not Authorized	46	69	6	121
Too Complicatied	2	3	2	7
Flat Refusal	59	100	0	159
Other	16	9	0	25
Total	480	855	50	1385

It is important to know the extent of establishment non-cooperation by characteristics of the establishments. While 56% of all eligible establishments participated in the survey, there were considerable variations in the rate of participation across subgroups of the sample establishments. Statistically, the non-responding establishments differed from those that responded by industry group ($p=0.001$), establishment size in terms of total employees ($p=0.001$), and geographic region as measured by the first digit of zip code ($p=0.047$). This means that the differences between responding and non-responding establishments were not due to chance alone. Survey analysts often use survey weights that have been adjusted for differential non-response among the sampled sub-groups, to correct the estimates.

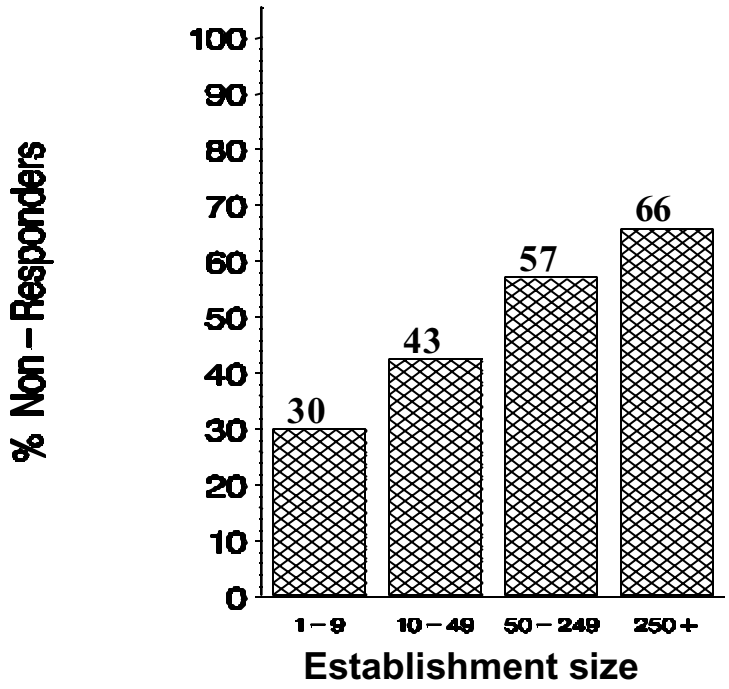
Exhibit III-2 shows that the highest rate of non-participation occurred in the Finance/Insurance/Real Estate industry, while the lowest rate of non-participation occurred in the Agricultural industry.

Exhibit III-2 Employer Non-Response by Industry Groups



The larger establishments sampled in the pretest were considerably less likely to participate in the survey (see Exhibit III-3). Special efforts should be made to increase the cooperation of large establishments in future O*NET data collections. Outreach activities are expected to increase the O*NET awareness. Employer incentives that appeal to larger establishments should be identified in order to increase their participation levels.

Exhibit III-3 Employer Non-Response by Establishment Size



The establishment response propensity varied less by geographic area than by industry or number of employees. Exhibit III-4 shows the lowest rate of participation occurred in Region 1, which includes the eastern states of New York and Pennsylvania. The highest participation rate among sample establishments occurred in Region 5, which includes the mid-western states of North Dakota, South Dakota, Minnesota, Iowa and Wisconsin. Exhibit III-5 is a map showing all of the areas of the country by first digit of the zip code. Increased outreach efforts could be aimed at the areas with the lowest cooperation rates.

Exhibit III-4 Employer Non-Response by Geography (Zip Code)

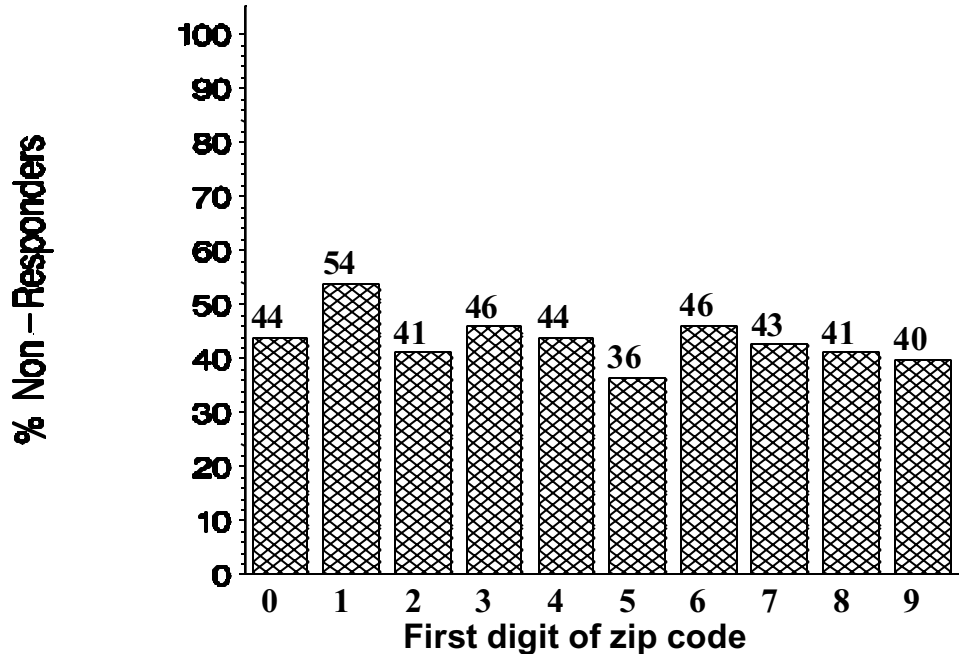
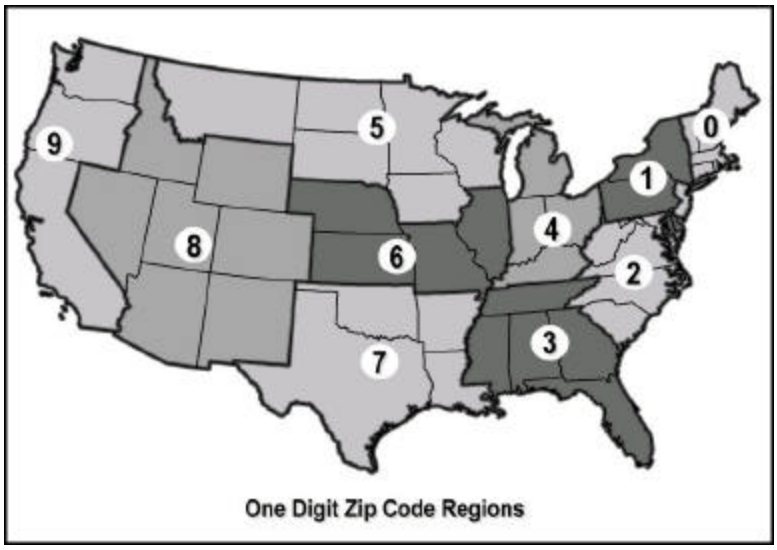


Exhibit III-5 Map of Zip Code Areas



III.A.4.b Employee Level

At the employee level, non-responding employees were compared to the respondents in terms of the available demographic information. Background questionnaire information on gender, age and race were available for virtually all of the 1,662 responding employees. However, it was only possible to obtain the corresponding information from the establishment representatives for approximately 25% of the 2,104 non-responding sample employees. This limits the strength of any conclusions one can make about the differences between the responding and non-responding employees.

Exhibit III-6 shows the results of comparing the demographic characteristics of two groups of sampled employees. The comparisons were made at the occupation level, and the sample size for some occupations was quite small. Occupations for which information was available for fewer than 10 non-respondents were not included in the comparisons.

Exhibit III-6 Demographics of Employees, by Response Status, by Occupation (SOC code)

	Respondents				Non-Respondents			
SOC	N	Median Age	Percent Male	Percent White	N	Median Age	Percent Male	Percent White
15026B	32	38	45	66	15	43	24	77
22102	52	38	62	83	19	56	100 *	100
22111	15	58	100	100	12	46	95	95
22121	42	38	86	95	14	29	99	48 *
22302	63	45	80	98	39	47	79	100
22305	43	51	98	90	14	33	100	93
22308	59	36	80	98	24	34	86	86 *
22502	26	28	99	92	12	32	100	78
24105	37	40	65	54	13	38	53	67
24511	27	42	73	93	12	46	82	100
27308	49	46	25	91	21	46	13	80

Exhibit III-6 (continued)

	Respondents				Non-Respondents			
SOC	N	Median Age	Percent Male	Percent White	N	Median Age	Percent Male	Percent White
32514	70	41	21	95	15	25	19	89
32908	41	41	3	94	10	.	.	90
34005	36	27	5	92	20	35	43 *	73
43021	45	51	9	87	10	39	48 *	65
49014	48	42	99	100	21	33	91	100
53808	35	33	26	87	21	42	22	70
63017	152	41	87	91	26	30	99	100
63047	47	44	66	53	20	.	100 *	27
65005	32	35	44	91	10	25	55	92
66002	46	39	.	85	15	28	.	72
67008	28	38	94	92	17	31 *	92	50 *
68038	33	28	2	100	14	38	7	60 *
85308	29	38	100	95	20	38	97	94
85923	55	43	93	100	21	45	90	97
85951	14	24	100	99	10	27	100	17 *
97108	58	56	81	99	39	56	88	73 *
97117	11	47	55	100	10	36	100 *	66
97947	31	43	99	89	21	33	100	60 *

The Wilcoxon rank sum test was used to test for differences in age. A chi-square test was used to test for differences in gender and race. Statistically significant differences between respondents and non-respondents at the five % significance level ($p<0.05$) are noted in Exhibit III-6 by an asterisk in the non-respondent columns. Few age differences were noted between respondents and non-respondents. All occupations that differed by gender or race involved either a higher percentage of male non-respondents or a lower percentage of white non-respondents. This may indicate a higher propensity to respond by white females than for the other demographic sub-groups, at least for some occupations. As noted, however, conclusions are limited due to the scarcity of information on the non-respondents.

III.A.5 Conclusions

Establishments reported a wide variety of reasons for non-cooperation. Statistically significant differences were noted between responding and non-responding establishments when examined by industry, geographic region or number of employees. Although the available data were sparse, comparison of responding and non-responding employees tended to indicate the possibility of a higher propensity to respond by white females, compared with other demographic sub-groups.

III.B Outliers

III.B.1 Background

This section documents the results of efforts to identify inconsistent data in order to determine whether or not some pretest questionnaire information should be excluded from the statistical analysis of the O*NET pretest experiment. Analyses to examine the data further, in order to determine what questionnaire data should be included in the updated O*NET database, are available in the ONET pretest report *Procedures for Cleaning Data*.

III.B.2 Objective

The purpose of this task was to identify any questionnaires that were completed by someone other than an employee in the occupation requested or contained an excessive amount of missing data.

III.B.3 Methods

Prior to beginning the statistical analysis tasks described in other sections of this document, the survey instruments returned in the mail and those completed by Internet were examined to determine whether or not there were any questionnaires that should not be included in the analysis of the pretest experiment. The following criteria were examined to determine if data should be disqualified from the analyses:

- misclassification of a subject to an inappropriate occupation
- failure to answer a majority of the questions.

The POC may have identified an employee as working in the requested occupation, but the classification was not applicable based on the employee data. Self-reported job titles on the background questionnaire and responses to importance questions on the task questionnaire were used to determine if a subject was classified to the appropriate occupation. As part of the informed consent procedure, participants were instructed that they could leave blank any particular questions they did not want to answer. The percentage of item non-response (questions the respondent left blank but should have answered) on the questionnaires was reviewed to determine whether missing data should lead to disqualification.

III.B.4 Results

The self-reported job titles from the background questionnaire were compared with the occupation in which the subject was sampled. For example, below is a list of every different self-reported job title from the subjects that were sampled for the occupation of underwriters:

- account underwriter
- associate director
- auto insurance underwriter

- automobile underwriter
- commercial insurance underwriter
- insurance underwriter
- property underwriter
- senior account underwriter
- senior claims rep.
- underwriter
- underwriting manager
- underwriting supervisor.

Though some of the self-reported job titles are different from the occupation requested, considering the variation and ambiguity that may be associated with many jobs and job titles, none of the subjects was disqualified from statistical analysis based on this review.

The other method that was used to determine whether a subject might have been misclassified to an inappropriate occupation was to examine their responses to the importance questions on the task questionnaire. It was hypothesized that, if an employee did not list any of the tasks as important to their job, then that person might have been misclassified. Of the task questionnaires with at least one non-missing answer for the importance questions, only one subject had a maximum value less than three (important) on the task importance ratings. Additionally, only 13 employees failed to answer any of the importance questions. None of the subjects was disqualified from further pretest analyses for this reason.

Another area studied was the percent of item non-response on the questionnaires. The criterion examined was the percentage of the questions the respondent should have answered on the questionnaire, but did not answer. To examine this question, we computed the percent missing item responses for each of the 1,662 questionnaires in the pretest database. The calculation does

not count as missing those questions that the subject should have skipped, as a result of previous responses. Results are presented in Exhibit III-7. The level of missing data seems quite reasonable, as only nine of 1,662 questionnaires had more than 50% of the questions missing. A few totally blank questionnaires were received, and they were not included in the data set for statistical analysis.

Exhibit III-7 Percent Missing Data for Individual Questionnaires

Percent Missing	Frequency	Percent
0-10	855	51.4
10-20	463	27.9
20-30	270	16.2
30-40	59	3.5
40-50	6	0.4
50-60	4	0.2
60-70	0	0.0
70-80	2	0.1
80-90	1	0.1
90-100	2	0.1

III.B.5 Conclusions

The tests employed for detecting outliers did not lead to removing any questionnaires from the statistical analyses described in this document. Examination of the reported importance ratings and job titles did not clearly identify any outliers. The item non-response testing, except for a few totally blank documents received by mail, did not identify any questionnaires with excessive incidence of missing data.

IV. Sampling Issues

IV.A Sample Size

IV.A.1 Background

The question of sample size required to describe an occupation was considered in the pretest planning. It was estimated that the maximum standard error (SE) for the mean of a domain with a sample size of 15 would be between 0.31 and 0.85 for the five-level scale questions and between 0.58 and 0.76 for the seven-level scale questions. Given this level of precision, the target sample size of 15 responses per descriptor would allow virtually all O*NET five-point and seven-point descriptors to be estimated within 1 to 1.5 scale points.

O*NET planning work has considered a sample size of 15 responses per descriptor to be sufficient to describe an occupation. Although the O*NET pretest was limited by the short period of data collection, it was possible to use this data to corroborate the assumption regarding sample size for future data collection efforts to populate the O*NET database.

IV.A.2 Objective

The purpose of this task was to use the pretest data to further examine the question of the sample size needed to adequately describe an O*NET occupation.

IV.A.3 Methods

The O*NET *Data Collection Program* intends to obtain a minimum of 15 responses per domain for every occupation. To estimate the standard error given a sample size of 15, we use the following relationships:

$$SD = \sqrt{SE_n^2 * n / \text{deff}}$$

$$SE_{15} = \sqrt{\frac{SD^2}{15 / \text{deff}}}$$

where SD=standard deviation

SE=standard error

deff=design effect

Here SE_n is the estimated standard error of the estimate given n, the realized pretest sample size. Then SE_{15} is an estimate of the standard error of the estimate for a sample size of 15. The SE_n values were estimated in the modeling procedure using weighted pretest data. The estimated SE_{15} values were calculated by substituting the first equation above into the second equation. The design effect cancels out, and SE_{15} is estimated by multiplying the SE_n values times the square root of the quantity n divided by 15. It is assumed that the sample of 15 has the same clustering at the establishment level as the observed sample.

IV.A.4 Results

For each questionnaire, the SE_{15} estimates were calculated for each descriptor using weighted data and were then summarized by type of question (five or seven categories), yielding the summary values in Exhibit IV-1 and Exhibit IV-2.

Exhibit IV-1 SE_{15} Estimates for the Five-Point Scale by Domain

Domain	Median	95th Percentile	Maximum
A: Skills	0.20	0.45	1.12
B: Generalized Work Activities	0.23	0.62	1.90
C: Abilities	0.19	0.51	1.89
D: Work Context	0.21	0.62	2.44
E: Knowledge	0.21	0.60	1.70

Exhibit IV-2 SE_{15} Estimates for the Seven-Point Scale by Domain

Domain	Median	95th Percentile	Maximum
A: Skills	0.20	0.50	1.16
B: Generalized Work Activities	0.25	0.74	1.96
C: Abilities	0.19	0.56	1.70
E: Knowledge	0.23	0.58	1.26

To check the sensitivity of the result, the table values were also computed for sample sizes of five and 10 responses per descriptor. Exhibit IV-3 shows how much the estimated precision would be decreased for samples of size 10, and for samples of size five.

Exhibit IV-3 SE_{10} Estimates and SE_5 Estimates for the Five-Point and Seven-Point Scales, by Domain

Scale	Domain	n = 10			n = 5		
		Median	95th Percentile	Maximum	Median	95th Percentile	Maximum
Five-Point	A	0.25	0.55	1.37	0.35	0.78	1.94
	B	0.28	0.76	2.32	0.40	1.07	3.28
	C	0.23	0.62	2.31	0.32	0.88	3.27
	D	0.26	0.75	2.99	0.37	1.07	4.23
	E	0.26	0.73	2.09	0.37	1.03	2.95
Seven-Point	A	0.24	0.62	1.42	0.34	0.87	2.01
	B	0.31	0.91	2.40	0.34	1.28	3.39
	C	0.23	0.69	2.08	0.33	0.98	2.94
	E	0.28	0.71	1.55	0.39	1.01	2.19

IV.A.5 Conclusions

As would be expected given the unequal weighting inherent in the pretest design, the maximal values are larger than the SE_{15} values computed previously based on unweighted data. Although the maximal values are all above the limits expected for these scales, the 95th percentile values are all below one scale point. The half-width of the 95% confidence interval, which is equal to two times the 95th percentile values in the table, leads to the conclusion that a sample size of 15 will be adequate to describe each occupation in a domain, for at least 95% of the descriptors.

For samples of size 10, the 95th percentile values for eight of the nine instruments are adequate to describe 95% of the O*NET descriptors within 1 to 1.5 scale points. Thus, this analysis tends to confirm the adequacy of samples of

15 responses per descriptor, and shows that samples of 10 responses may even be adequate for most descriptors.

For samples of size five, the 95th percentile values would be adequate to describe 95% of the O*NET descriptors only within 2 to 2.5 scale points. This level of precision does not meet the O*NET sample size criterion previously established and indicates that samples of size five would not be adequate to meet this standard.

Additional, more detailed standard error analysis was done on these data using only the pretest occupations that had at least six incumbent raters, to project reliability using standard errors of descriptor ratings for the various sample sizes described above (described in *Standard Error Analysis: Reliability and Implications for Sample Size*; RTI Report to the National O*NET Consortium, in preparation). The results indicated that only a very small proportion of descriptors in each domain had predicted standard errors that exceeded a criterion of .5 if the sample size collected were 15. The .5 criterion assures that the "true score" of any rating is within one scale point of the estimate 95% of the time when random samples of raters are drawn from the population. For a sample size of 15, only 1% of the Skills descriptors, 3% of the GWA descriptors, 2% of the Work Context descriptors, and 1% of the Knowledge descriptors had standard errors greater than .5.

Thus, this analysis tends to confirm the adequacy of samples of 15 responses per descriptor, shows that samples of 10 responses may also be adequate, but that samples of five responses would be inadequate. These results are consistent with the O*NET pretest report, *Standard Error Analysis*.

IV.B Design Effects

IV.B.1 Background

The design effect is a measure of the relative efficiency of the sampling design. When the design effect is less than one, more precise estimates are produced with a smaller sample size than would have been available from a simple random sample. Design effects are typically increased due to clustering and unequal weighting and may be decreased by stratification. Clustering is often the principal factor that increases the design effect. Design effects from 1 to 2 are common in many survey designs that involve small clusters of second-stage sampling units selected with overall probabilities that do not vary to extremes.

The design effect is defined as the ratio of the variance obtained using the actual sample design to the variance that would have been obtained with a simple random sample of the same size. If it were feasible to use a simple random sample of employees to represent an occupation, the design effect would be equal to one. In the *O*NET Data Collection Program*, the two-stage sampling design produces clustering of employees within establishments with the possibility of higher design effects.

IV.B.2 Objective

The purpose of this task was to examine the median and extreme values of the O*NET sample design effects, in order to describe their distribution.

IV.B.3 Methods

For each item on each questionnaire, design effect estimates were computed using weighted data, by dividing the estimated sampling variance for the item by the corresponding estimate of variance assuming a simple random sample. The calculated design effect estimates were then summarized across occupations and question types (five or seven categories).

IV.B.4 Results

The median design effects for the five-point scale items are shown in Exhibit IV-4, along with the 95th percentile and maximum design effect estimates for each subject matter domain. Exhibit IV-5 contains the summary values of the design effect estimates for the seven-point scale items.

Exhibit IV-4 Design Effect Estimates for the Five-Point Scale by Domain

Domain	Median	95th Percentile	Maximum
A: Skills	0.87	2.24	6.34
B: Generalized Work Activities	0.90	2.46	25.14
C: Abilities	0.82	2.24	20.24
D: Work Context	0.93	2.49	37.19
E: Knowledge	0.87	2.61	12.39

Exhibit IV-5 Design Effect Estimates for the Seven-Point Scale by Domain

Domain	Median	95th Percentile	Maximum
A: Skills	0.76	1.83	6.93
B: Work Activities	0.82	2.49	16.82
C: Abilities	0.72	2.03	16.31
E: Knowledge	0.75	2.22	8.22

IV.B.5 Conclusions

The realized employee sample sizes per sample establishment were small for most occupations, so one would not expect large design effects due to sample clustering. All of the median design effects are less than 1, indicating positive effects of the complex sampling design and the small cluster sizes. (Refer to Section IV.B.1 for typical design effect sizes.) The 95th percentile values also seem reasonable, given the unequal weighting inherent in the pretest sample design. The very high maximal values that occurred for a small number of descriptors are probably due to extremely unequal weights for a few of the occupations. This may suggest that those occupations might be more easily represented using an alternate sampling approach involving less variability in the weights.

V. Descriptive Statistics

V.A Demographics

V.A.1 Background

Demographic and disability data were collected from all 1,662 pretest respondents, in order that O*NET data users might perform analyses for sub-groups of occupations.

V.A.2 Objective

The purpose of this task was to infer, if possible from the pretest results, whether or not the pretest sample respondents were demographically similar to the corresponding inferential populations.

V.A.3 Methods

Background questionnaire data on age, gender, race, ethnicity, and disability conditions were tabulated for the pretest respondents. Ideally, the demographics and disability information for the sample respondents for each pretest occupation would be compared with independent population data describing all employees for that occupation.

V.A.4 Results

In the aggregate, the pretest respondent cohort was 58% male with a median age of 42 years (range: 17, 82). Hispanics represented 5% of the sample respondents. Races were distributed as follows: White, 90%; Pacific Islander, 0.2%; Black ,7%; Asian, 3%; American Indian, 0.7%. These numbers appear to be very roughly within range of the corresponding percentages for the total United States population. However, results for the aggregate of employees in the 50 occupations is not meaningful for comparisons with the total population of the United States.

Information on disability status was collected in the O*NET pretest for blindness or deafness (1.2%) and for other conditions that severely limit physical activity (2.9%). A small percentage of responders also reported difficulty in learning (1.7%), dressing (0.1%), going outside the home (0.1%) or working (0.5%).

The most relevant demographic employment data identified, containing age, gender, race and ethnicity information for specific occupations, was extracted from the 1990 census files on gender and race for occupations. Year 2000 Census data are not available at this time. The 1990 Census data were available for 500 occupational codes that are typically aggregates of O*NET occupations. An effort was made to collapse the O*NET occupations into Census occupation

codes, but the matching was imperfect. The effort was abandoned due to the age of the Census data and the difficulties matching the occupation codes. Independent data on disability status of job employees by occupation were not located.

V.A.5 Conclusions

Comparable demographic and disability data that are independent of the pretest survey and presented by occupation were not located. One must also consider the difficulty encountered in matching the O*NET occupations with the Census occupations, the changes in the employment patterns occurring since 1990, and the small sample sizes for each O*NET occupation. From this analysis, it is not possible to conclude whether or not the pretest sample respondents were demographically different from the corresponding inferential populations.

V.B Item Estimates

V.B.1 Background

Researchers and other data users performing analyses using the O*NET pretest data will often need access to unbiased estimates of the item means and standard errors.

V.B.2 Objective

The purpose of this task was to compute estimated means and standard errors for the variables in the five O*NET domain questionnaires:

A: Skills Questionnaire

B: Generalized Work Activities Questionnaire

C: Abilities Questionnaire

D: Work Context Questionnaire

E: Knowledge Questionnaire.

V.B.3 Methods

Separate estimates for each questionnaire item for each subject matter domain were computed for each of the 50 occupations that were included in the pretest. SUDAAN software, which properly computes sampling errors for statistics based on complex sample designs, was used to compute the estimated means and standard error for each questionnaire item. Prior to computation of these estimates, the pretest data had been examined for exceptional cases and outliers, as described in Chapter III, Section B of this report. No cases were excluded from the calculation of item summary statistics. Missing data were not a concern, since every item was completed by 92% or more of the respondents on all five questionnaires.

V.B.4 Results

The weighted estimates and sample sizes are shown as appendices to this report. Each appendix contains the mean and standard error (SE) estimates for each questionnaire item in the domain. There is a text descriptor for each item, a code for each occupation, and the sample size (n) indicating the number of employees from that occupation who responded to the questionnaire. The five appendices are labeled as follows:

Appendix V-1 Skills Questionnaire estimates
Appendix V-2 Generalized Work Activities Questionnaire estimates
Appendix V-3 Abilities Questionnaire estimates
Appendix V-4 Work Context Questionnaire estimates
Appendix V-5 Knowledge Questionnaire estimates.

V.B.5 Conclusions

Unbiased estimates of the item means and standard errors are now available for researchers and other data users who need this information.

VI. Concluding Statement to Pretest Analysis

The procedure used within the O*NET pretest can serve as the primary model for future data collections. In addition, the experimentation included within the pretest, along with knowledge gained during the collection experience itself, identified a number of enhancements potentially improving employer eligibility, employer and employee response rates, as well as the overall efficiency of the project. In total, the results of the pretest provided critical information that informed the best design for the proposed, full-scale implementation of O*NET data collection.