

U.S. DEPARTMENT OF LABOR
BUREAU OF LABOR STATISTICS

**Occupational Requirements
Survey
Consolidated Feasibility Tests
Summary Report**
Fiscal Year 2014



Occupational Requirements Survey (ORS)
Consolidated FY 2014 Feasibility Tests Summary Report

Table of Contents

Executive Summary.....	ii
I. Background and Tests Overview	1
II. Common Findings Across All Tests	1
III. ORS Only Efficiency Innovations Test	2
IV. Joint Collection Test	4
V. New Data Elements Test	5
VI. Central Office Collection Test	6
VII. Alternative Modes Test	8
VIII. Job Observations	9
IX. Data Analysis	10
X. Item-level Response	11

Executive Summary

In fiscal year 2014, the Bureau of Labor Statistics (BLS) National Compensation Survey (NCS) completed six feasibility tests to refine the Occupational Requirements Survey (ORS) methodology.

The feasibility tests were designed to:

- refine the methods to develop more efficient approaches for data collection as identified during fiscal year 2013 testing (ORS Only Efficiency Innovations Test);
- determine how best to collect occupational requirements data elements and NCS data elements from the same establishment (NCS/ORS Joint Collection Test);
- determine the new mental and cognitive demands for work data elements, and evaluate the use of occupational task lists as developed by ETA's O*NET program during data collection (New Data Elements Test).
- determine how best to collect occupational requirements data elements from America's largest firms and State governments (Central Office Collection Test);
- determine how best to collect occupational requirements data elements when a personal visit is not optimal due to respondent resistance, collection costs, or other factors (Alternative Modes Test); and
- capture and evaluate changes in data coding as a result of observing the work environment, sampled occupation, or both (Observations).

Five tests were conducted nationally, across all six BLS regions. The New Data Elements Test was conducted only in two Metropolitan Areas: Washington, D.C. and San Diego, CA.

For all tests, at the end of most interviews, respondents were asked questions to gauge their reactions to the survey, and BLS field economists completed a post-interview questionnaire to assess their perceptions of respondent burden, response quality, and lessons learned. Throughout the year, regular debriefing sessions were held with BLS field economists and other staff to discuss interviewing experiences, identify potential issues with the materials and procedures, and share lessons learned. The information gleaned throughout these questionnaires and debriefing sessions were used to refine procedures, training materials, and collection protocols on a flow basis.

In general, the results from these tests confirms BLS' viability at collecting data relevant to ORS and demonstrate the effectiveness of the revised materials and procedures tested. All test objectives were successfully met and these activities established a strong foundation for the Pre-Production Test, scheduled to begin in fall 2014. The body of this report provides details about key findings and issues that arose during testing. The remainder of this executive summary highlights issues and considerations gleaned from the feasibility tests.

One of the principal conclusions of these tests is that when BLS economists have a choice to select which collection materials and strategies to use, they become more efficient at collecting the data. This flexibility increases the quality of the data and reduces respondents' perceptions of burden. With increased flexibility, BLS economists structure questions and capture the information in a way that works best for respondents, both in terms of question order and timing.

Overall, when available, job observations had a positive impact on the quality of collected data. BLS economists were able to verify certain data elements with respondents after observations. Occasionally establishments had materials and resources that were used for indirect observations, such as occupational safety and recruiting videos. Job observations were not always available due to time constraints, perceptions of burden, or safety or security concerns.

Well-defined and stable procedures are essential to ensuring consistency in the data collected nationally improving inter-coder reliability and higher ORS data quality. BLS is working on more consistent methods of issuing procedures to ensure that clarifications after initial training are applied universally. In addition to procedures, reactions to the various feasibility tests material (e.g., introductory spiel, collection tools, respondent visual aids, etc.) were generally positive. BLS continues to improve respondent materials for phone and email collection since they can greatly impact data quality and respondent perceptions about the survey.

Respondents' reactions to ORS were mostly positive. Most respondents expressed interest in the topic and understood the need to collect high-quality information about the ORS elements. Several volunteered that they liked the idea of BLS partnering with SSA or that the process of reporting ORS information had a positive impact on their understanding of their own employees' tasks and work environments. Most respondents indicated that there was no undue burden placed on them by the ORS questions.

Occupational Requirements Survey (ORS)

Consolidated FY 2014 Feasibility Tests Summary Report

I. Background and Tests Overview

In fiscal year 2012, the Bureau of Labor Statistics (BLS) signed an interagency agreement with the Social Security Administration (SSA) to design, develop, and conduct a series of tests using the National Compensation Survey (NCS) platform. The purpose was to assess the feasibility of using the NCS to accurately and reliably capture data relevant to SSA's disability program. The resulting initiative—the Occupational Requirements Survey (ORS)—was launched to capture data elements new to NCS using the NCS survey platform. In fiscal year 2013, BLS completed three initial phases of ORS testing: a proof-of-concept test, a collection protocol development test, and a broad scale testing of the various protocols. In fiscal year 2014, BLS conducted six feasibility tests with the following objectives:

- ORS Only Efficiency Innovations Test – refine the methods to develop more efficient approaches for data collection as identified during fiscal year 2013 testing;
- NCS/ORS Joint Collection Test – determine how best to collect occupational requirements data elements and NCS data elements from the same establishment;
- New Data Element Tests – determine the new mental and cognitive demands of work data elements and evaluate the use of occupational task lists as developed by ETA's O*NET program during data collection;
- Central Office Collection Test – determine how best to collect occupational requirements data elements from America's largest firms and State governments;
- Alternative Modes Test – determine how best to collect occupational requirements data elements when a personal visit is not optimal due to respondent resistance, collection costs, or other factors via phone, email, or fax; and
- Observations – capture and evaluate changes in data coding as a result of observing the work environment, sampled occupation, or both.

All fiscal year 2014 tests included goals to refine prior collection methods and tools, collect results in a uniform and efficient manner, and ensure the data collected would meet SSA's requirements. BLS also tracked when field economists were able to observe jobs on collection appointments and the impact of those observations on data collection.

No establishments were included in more than one of the feasibility tests listed above.

II. Common Findings Across All Tests

Although each test had a unique purpose, BLS found some similarities across the tests in the areas listed below.

Mode of Collection: Field economists (i.e., BLS economists who collect the data) maintain the personal visit is the most effective method for collecting ORS data. While some tests were specifically designed to look into alternative methods of collection, both field economists and respondents perceived personal visits resulted in better data collection.

Collection tools: The collection tools were found to be effective in all tests. For the fiscal year 2014 tests, the tools developed for collection were widely viewed as efficient and usable. Materials sent to respondents in advance of the appointment to allow them to prepare for the interview also were perceived as efficient and usable.

Respondents: The feasibility tests have reinforced the need to identify the appropriate respondent for the ORS elements upfront. This has been an ongoing challenge for our field economists.

Training: In nearly all cases, the participating field economists indicated training was a big part of test success. In addition, the debriefing sessions, where economists discuss collection experiences after the fact, proved valuable across all tests to identify issues and best practices and to gauge how the tests were progressing.

Elements collected were uniform across the tests with some individual test variations. The elements common to all tests included: description of establishment operations for the purpose of assigning the proper industry, name and title of all respondents, total employment, occupational employment, work setting, worker characteristics, occupational work schedules, non-supervisory/supervisory/lead worker designations, job duties to classify the occupation, specific vocational preparation, physical demands, work environment, job observation, and interview start/end times.

III. ORS Only Efficiency Innovations Test

The primary goal of the ORS Only Efficiency Innovations Test was to develop alternative methods for collecting ORS data that improved or maintained the level of data quality while shortening the interview and decreasing respondent burden. BLS used feedback collected from field economists and respondents during previous phases of testing to develop data collection tools that incorporated various best practices and strategies.

There were 90 schedules collected in the Efficiency Innovations test, evenly distributed across the six BLS regions. This resulted in 434 quotes¹ from 16 State and local governments and 74 private industry schedules. The service-providing industry was the largest in terms of schedules and quotes collected for both the public and private sectors. There were 182 unique occupations (at the 8-digit SOC level) collected during the course of the test.

The expectation was to collect these data via personal visits. If a respondent refused or was unable to schedule an in-person appointment, phone and email collection fallbacks were permissible. Phone and email collection protocols were very similar to those for in-person collection. In all cases, field economists were required to use one of the three collection strategies — conversational, sit/stand, or white collar/blue collar — described below.

Conversational Strategy

This approach was designed to decrease appointment length and respondents' perception of burden by maximizing information capture as it arose. This strategy used a hybrid of single and multi-quote approaches. It also utilized efficiency strategies developed in last year's Phase 3 testing, such as

¹ A quote is the unit of observation in NCS. It is defined by its Standard Occupational Classification code (SOC) and characteristics, such as work schedule (full- or part-time) and pay (time or incentive). For ORS, BLS is collecting the 8-digit SOC code defined by O*NET.

branching questions, combination elements, and task based questions. After the initial quote-by-quote job description discussion, the strategy transitioned to a multi-quote style, incorporating elements that have worked with multi-quote ORS collection in past phases.

Sit/Stand Strategy

This approach was intended to test the collection of ORS data by grouping elements performed while sitting from elements performed while standing, walking, or both. Through this approach, field economists used the respondent's duration responses to group the number of elements. Field economists followed up with yes-no questions to confirm the groupings where necessary. The data-collection tool allowed certain ORS data elements to be collected using the single quote approach and the remaining elements by the multi-quote approach.

White-collar/Blue-collar Strategy

This approach was meant to decrease respondents' perception of burden and the time taken to administer the survey by creating two distinct versions of the survey instrument. Each version tailored the order and format of the questions to either blue-collar (e.g., require manual labor or take place outside of an office setting) or white-collar (e.g., do not require manual labor) occupations. This strategy engaged respondents early in the interview by giving them a chance to respond affirmatively to questions. The tool attempted to concentrate applicable elements early in the interview to maximize respondent focus and therefore response quality.

Key findings

The conversational strategy was used in 29.4 percent of interviews conducted during the ORS Only Efficiency Innovations Test, the sit/stand strategy was used in 27.5 percent of the interviews, and the white collar/blue collar strategy was used in 43.1 percent of the interviews. Test guidance indicated that each approach should be used once and then field economists were free to use any approach they wanted. When free to make the choice, the majority of field economists chose to use the White-collar/Blue-collar strategy.

On average, interviews took just about an hour to collect an average of 4 quotes: the Conversational strategy took 64.2 minutes, the Sit/Stand strategy took 55.2 minutes, and the White-collar/Blue-collar strategy took 58.2 minutes. There was considerable variation among individual schedules.

Field economists agreed that it was relatively easy to determine if an element was present. The difficulty arose in determining the exact duration associated with the element, which appeared to slow down the collection process. In addition, the perception of burden decreased when field economists asked questions in their own words rather than as written in the tools. The flexibility in asking the questions helped the interview flow more smoothly and made the respondent more at ease throughout the interview, although the interview length was about the same.

Another finding is that interviews are respondent-driven, so no one approach will work universally. Field economists' ability to probe, ask follow-up questions, and employ skip patterns depended upon respondent willingness to invest the time to listen to the questions carefully and give thoughtful answers.

Several respondents volunteered that they had not anticipated the level of detail of the questions. Among this group, reactions were mixed: some were surprised by how quickly the field economist was able to collect all of the information needed; others expressed irritation about the process or

uncertainty about some of the details. The most common negative reaction was that it was difficult to provide accurate duration estimates. Respondent identified reasons for the difficulty including: they did not know the job well enough; durations that varied across people in the job; and jobs with a wide range of duties that were done intermittently.

The majority of respondents in this test reported that the survey was “not at all burdensome;” only seven individuals indicated the survey was “a little burdensome.” Some of the respondents, who acknowledged some burden, felt it would have been helpful to have materials in advance to allow them to gather information and be more efficient and accurate during the collection appointment. One respondent noted that the most burdensome aspect of the process was getting corporate headquarters to approve the ORS personal visit.

Field economists’ overall opinion was the various strategies improved efficiency over the collection procedures used in earlier testing. They liked that the interview was not structured and permitted them to address the various elements as they naturally came up instead of trying to force the respondent to follow a prescribed question order.

IV. Joint Collection Test

The primary goal of the Joint Collection Test was to develop methods for collecting ORS data and assess if they could be used concurrently with NCS without decreasing survey data quality. Two other goals of this test were to determine: 1) if traditional NCS respondents can provide the ORS data for all occupations or if the data needs to be collected from another source, and 2) the optimal timing for collection of ORS data elements in order to minimize respondent burden and maximize useable data for both product lines.

In an attempt to achieve these goals, four test groups were created over two phases. Phase 1 of the NCS Joint Collection Test included the collection of current NCS production schedules (in estimation) and schedules that had recently rotated out. Phase 2 of the test included NCS initiation units and recently initiated NCS establishments currently in their seasoning quarters (not yet included in published ECI estimates).

All standard ORS data elements were collected in the Joint Collection Test. For establishments in estimation and seasoning, the field economist was to collect both NCS update and ORS data elements. For establishments that recently rotated out of NCS, the field economist was to collect only the ORS data elements. A major test constraint was to do no harm to the NCS schedules since these were live production units.

Field economists were provided a list of establishments eligible for inclusion in the NCS Joint Collection Test in each region and each test group identified above. Ninety-six establishments were targeted: 16 in each of the six regions. Each region was to collect eight schedules in each phase of joint collection with a further breakout of four schedules from each test group. Since a sample was not drawn for this test, each region was responsible for the selection of specific establishments for collection based on the criteria below.

For Phase 1 of the Joint Collection Test:

- Test units that had already rotated out of the NCS were in private industry.

- For the establishments currently in estimation, one of the four units collected in each region was targeted to be a State or local government unit; the other three units collected were to be in private industry.

For Phase 2 of the Joint Collection Test

- Test units for initiations and schedules in their seasoning quarters were in private industry.

Key findings

The Joint Collection Test results provided useful information for the development of more efficient ORS data collection strategies. Eighty-six schedules were collected with roughly 6 percent coming from State and local government establishments and 94 percent coming from private industry establishments. The service-providing industry was the largest in both the public and private sectors, while education was the smallest in terms of both schedules and quotes collected. In all, occupational requirements data were collected for 447 quotes and 187 unique occupations at the 8-digit SOC level.

Staff identified collection strategies and best practices for joint collection from this test. Since field economists had the option of choosing which respondents to pick, BLS cannot determine a refusal rate with any sort of statistical accuracy. Additionally, test response rates are inconclusive for determining if ORS data and NCS data can be collected together without impacting quality of either survey.

The majority of the ORS data collected during the NCS/ORS Joint Collection Test came from the existing NCS respondent. Most respondents said they were the best person in their organization to provide the information requested. To support their assessments, respondents cited knowing jobs very well due to the size of the company, or the difficulty of scheduling appointments with other individuals who might have more specific knowledge of some jobs. When respondents provided alternate contacts, they typically were human resource managers or department heads. The reasons cited for collecting data from a non-NCS respondent related to the existing respondents' lack of knowledge about physical demands or environmental conditions for the selected occupations.

Respondents expressed a high degree of confidence in their answers, although several admitted they did not have adequate knowledge about some of the jobs or the duration estimates. This is mostly because of variations from day-to-day work across incumbents. One NCS respondent at first had difficulty adjusting to the nature of ORS questions, but indicated that he gained confidence as he went through the interview and revised earlier answers. As in previous tests, a few respondents indicated that they were confused by the weight questions and had difficulty reporting durations for some elements.

V. New Data Elements Test

The ORS New Data Elements Test was designed to test collection of occupational mental and cognitive demands of work data elements and occupational task lists. BLS divided the test into two phases and used feedback from the first phase to refine questions and procedures for the second. While the New Data Elements Test focused on the mental and cognitive demands of work data elements and task lists, it also tested for revised SVP wording. All other ORS elements and NCS leveling were collected for a subset of units in the second phase. Collection for this subset was referred to as the blended approach.

The first phase of the New Data Elements 1 was conducted with 25 private industry and State and local government establishments from the Washington, D.C. Metropolitan Area. The second phase was

conducted with 122 private industry and State and local government establishments from the San Diego, CA Metropolitan Area. All of the first phase test establishments and half of the test visits in the second phase of collection were observed by a BLS or SSA employee, who captured information about the interview process and questions as well as the interview time.

Field economists selected occupations for each establishment as respondent time and cooperation allowed. Occupations were not randomly selected. The criteria used were based on occupational employment, SOC codes, and respondent ability to provide occupation information. There were 268 unique occupations at the 8-digit SOC level collected during the course of both phases of the test.

The analysis of the New Data Elements Test compared data collected to similar data collected in prior ORS collection phases, the Dictionary of Occupational Titles (DOT), the O*NET, and the National Compensation Survey with the goal of identifying and understanding patterns in the data to create new review edits and make recommendations for future collection.

Key findings

Traditional NCS respondents have access to task list information and an understanding of the complexity, frequency of change, supervisory controls, and type of contact of the survey occupations. ORS respondents seemed split about whether they would be able to provide this information remotely and expressed concerns about the quality of their own responses (i.e., give more attention in person). Further, some observers noted how field economists used the task list information as a basis to direct or clarify a respondent answer.

Respondents had concerns regarding the wording of some mental and cognitive demands of work questions and the definition of terms such as “routine,” “familiar contacts,” and “general methods.” Respondents indicated the questions covered most of the main mental demands of the job. A few mentioned jobs can also have emotional demands (such as high levels of stress) or require soft/people skills that were not addressed in this test.

Most field economists agreed that the revised SVP questions and mental and cognitive demands of work data questions worked either “well” or “very well” in this test. However since several of the mental and cognitive elements were too wordy and awkward to ask, field economists are unsure whether some respondents were truly grasping the meaning of the questions. Field economists noted it is important to have a professional understanding of the intent of the questions in order to appropriately guide respondents and interpret their responses. Between the overlaps with current NCS leveling concepts and the ORS mental and cognitive questions, most field economists expressed comfort with the questions and the quality of the information that they are collecting. Continuing to use a more conversational approach when collecting data and allowing flexibility to ask questions in the way most conducive to an individual respondent’s understanding was also mentioned as a best practice not only for the mental and cognitive questions but for the ORS survey as a whole.

VI. Central Office Collection Test

Large firms and State governments make special arrangements with BLS about the best way to collect data from their units. In some cases, these arrangements require BLS to obtain collection authorization from a central office (i.e., headquarters) and often require data collection to occur at the central office location. The primary objective of the Central Office Collection (COC) test was to determine how to collect ORS data from these central office collection firms while balancing data quality—as measured by

overall and item level response and cost of collection. The COC Test included other goals to determine if: 1) traditional NCS respondents can provide the data for all occupations at all locations or if the data needs to be collected from another source, 2) responses for ORS data elements are the same for a given occupation at all locations in the firm and when responses may vary by location, and 3) existing COC protocols need to be revised for collection of ORS data elements.

Key findings

The COC Test started in November 2013 and concluded in June 2014. Firms were selected across the BLS Regional Offices based on existing COC arrangements for NCS. BLS field economists completed interviews with 32 firms representing 89 establishments. Personal visit interviews were conducted for 56.2 percent of the schedules, 12.4 percent were conducted by telephone, 13.5 percent were conducted via email, and 17.9 used a combination of reporting modes. During these interviews, field economists collected ORS data for 548 sampled occupations, representing 172 eight-digit classification levels within the SOC structure.

This test used modified NCS occupation selection methods. Field economists selected between four and eight unique occupations for each establishment up to a maximum of 24 occupations per firm. If the field economist was unable to conduct a probability-selection-of-occupations statistical routine at the establishment level, it was permissible to conduct it at the narrowest level available (e.g., division, state, corporate-wide).

Unlike other tests, the COC Test did not determine if the ORS concepts and procedures worked well. This test identified the preliminary procedures for collecting ORS data from COC firms. BLS recognized protocols that need evaluation and modification to ensure high quality data is obtained from these firms without negatively impacting cooperation or respondent burden. Collection of these specific ORS data elements from these central office establishments is viable but more work remains.

Although a few firms refused to participate in this test, it is encouraging to note that there were very few permanent refusals. The sense amongst the regional test coordinators with regard to these non-participating firms is that given a little more time to work with them most could be pursued and convinced to cooperate in ORS. Many of the challenges for COC firm participation are common in the NCS collection experience as well. Some examples are: contact was not acknowledged, lack of time and resources within the firm, confidentiality concerns, anti-government sentiment, and concerns over legal implications. Many COC respondents were concerned with how often this process would have to be repeated for the various establishments within their firms.

Locating the appropriate respondent was one of the most prevalent issues in the COC test. Due to the complexity of the organizations and specialization of human resource employees, traditional respondents can have limited knowledge of the other operations within the firm. Many field economists contacted their regular NCS respondent and found the respondent was not knowledgeable about the ORS elements and was not comfortable providing the data. In some cases, the field economist needed to receive additional authorization from a higher level official or the legal department prior to collecting the ORS elements. Sometimes respondents could not identify the appropriate contact or even to point the field economist to the right place within the organization to collect ORS data. Collecting ORS data from these COC firms often required contacts with multiple individuals to cover all the jobs being collected.

Corporate vs. Local Collection

One goal of the COC Test was determining the amount of variation in occupational requirements across the firm's locations. These variations impact the need to collect data at the local individual unit level. The information from respondents is inconclusive about this. When variations were reported across locations, respondents cited the size of the location as a contributing factor. For example, a larger retail store location may have full-time warehouse positions, so sales clerks do not have to stock materials. Variation of missions across units in different locales was another factor cited. For example, some locations may focus on different functions within the organization, serve regional needs, or reflect regional regulations. Respondents generally indicated it was appropriate to collect ORS data at the local level and BLS would have permission.

VII. Alternative Modes Test

In fiscal year 2013, the primary method of collection for ORS schedules was personal visit (PV), with phone and email used as collection fallbacks. Although field economists successfully collected data without conducting a PV at the request of respondents, no wide-scale, remote-collection efforts were attempted. Personal visit collection is expensive for a national sample survey and respondents do not always agree to a PV, so obtaining quality data through other methods is important for the success of any survey program. The Alternative Modes Test (AMT) developed and evaluated protocols and procedures for collecting ORS data by modes other than PV. The primary objectives for the AMT were to: 1) determine how to collect high-quality ORS data via phone, email, or fax, and 2) to compare the quality of data obtained via the different collection modes.

This test was started in February 2014 and concluded in June 2014. Establishments were selected from a national frame clustered by ownership, collection area, employment size class, and NAICS code. The method of collection for each establishment was randomly assigned—half were designated PV units and the remaining were non-PV or remote collection units.

NCS random selection of occupations was not used in this test. Field economist selected four unique occupations reflecting a broad mix at the 2-digit SOC level at the first establishment contacted within the cluster. The same four occupations were collected from the second establishment in the cluster, creating paired observations. A replacement occupation, matching the original detailed SOC as closely as possible, was selected for occupations that could not be matched at the second establishment.

Key Findings

The main objective of the test was determining how to collect high quality ORS data via phone, email, or fax. The results of this feasibility test suggest that remote collection of ORS data is viable. Of the 147 establishments collected, personal visit interviews were conducted for 49 percent of the schedules and 51 percent were conducted by telephone, email, fax or a combination of these methods. Field economists collected ORS data for 831 sampled occupations. The occupations represented 182 eight-digit classification levels within the SOC structure and contained 125 pairs of quotes evenly split between personal visit and remote collection types.

Collection by remote methods had a higher percentage of schedules with at least one non-responding quote than did schedules collected by PV (18.6 percent compared to 4.2 percent). During the test, best practices for multiple-mode collection and modifications needed to tools and procedures were identified for future research and testing.

Respondents and field economists both have a definite preference for PV collection overall. Approximately half of the field economists reported that respondents often expressed an unwillingness or inability to spend the time needed for full ORS collection by phone, which might negatively impact survey cooperation. A number of respondents also indicated that they ordinarily would delay response to or ignore requests for remote collection.

Thirty-eight percent of remote respondents cited time constraints, which negatively impacted ORS collection. None of the PV respondents reported feeling rushed during the interviews. Field economists offered a number of reasons for the perceived reductions in efficiency and data quality in remote collection. These included: respondents' unwillingness to spend adequate time on the phone; field economists' inability to demonstrate and clarify ORS concepts; the absence of respondents' nonverbal cues to sense potential confusion and maintain engagement; and the inability to see the work environment.

Personal-visit respondents reported being less burdened than their remote-respondent counterparts. Respondents' perceptions of burden and their willingness to report on these perceptions are impacted by the mode of collection. When burden was reported, the most frequently cited reasons were time pressure, feeling rushed, or competing work demands. Although respondent engagement was generally good across modes, occasionally phone respondents' attention waned relatively quickly (e.g., after 15 minutes) and they were obviously distracted by interruptions. Respondents cited similar concerns with remote collection. Respondents who endorsed phone or mixed collection primarily perceived those approaches to take less time than PVs. It also allowed them to send job descriptions in advance to the field economists to streamline collection.

Personal-visit and non-PV units' respondents expressed a good degree of confidence in most of their answers, except for duration estimates. Good confidence occurred when respondents knew the jobs well, reviewed ORS materials sent in advance, prepared for the interview, or felt that any confusion was effectively clarified by the field economist. Respondents' low confidence in their time estimates, particularly for the gross and fine manipulation and keyboarding items, occurred because often these small actions accumulate in a way not easily observed or estimated. Distinguishing between manipulation and keyboarding required considerable thought and time that was not available in the phone interviews. Phone respondents were particularly likely to express low confidence in their estimated durations.

VIII. Job Observations

The primary objective of the Job Observation test was to determine the feasibility of observing occupations and to determine if data coding changed as a result of observing the work environment, the sampled occupation, or both. This test was conducted concurrently with all other feasibility tests by asking if any employees were observed performing the job, the general type of observation, and whether the field economist changed how the occupation data were coded based on information from the observation.

Key Findings

Ninety eligible schedules (17 percent) had at least one quote observed by the field economist and 187 individual jobs were recorded as being observed. While this is not a large percentage, the scope of the

observations covered a variety of industries and occupations. Almost all observations were included as part of a PV to the work location either to collect data or establish initial contact.

When field economists were asked about job observations, most commented that when they were able to observe the jobs they were collecting, these observations assisted greatly with their understanding of the firm or jobs activities. Specifically, observations conducted prior to the interview helped field economists to plan the interview flow and clarify the respondents' answers during the interview. When conducted after the interview, field economists mentioned that job observations allowed them to verify elements and responses and provided support when documenting the interview in the data capture system. Nonetheless, field economists acknowledge that it was difficult to observe jobs in certain scenarios, especially due to issues of time constraints. Several field economists suggested best practices for job observations including possible synergies with the task lists and conversational strategy, notifying the respondent in advance of the request for a tour and explaining to the respondent how a tour could speed up the interview process by allowing the field economist to ask fewer questions.

When a job observation resulted in data being changed, field economists entered comments in the data capture system to document the change. The data elements that changed most commonly were stooping, crouching, climbing stairs and ladders, sitting/standing, exposure to heat or cold, keyboarding, hearing and visual acuity, and exposure to loud noises. The reasons for making the changes included being able to probe the respondent on the elements observed or challenge responses that seemed inaccurate. Additionally the field economists mentioned that observing the jobs allowed them to focus in on certain jobs and elements and ask more clarifying questions based on the tasks that they had seen being performed. It also gave the field economists a sense of whether certain activities (like exposure to heat or loud noises) met the threshold for that element established by the ORS procedures.

IX. Data Analysis

BLS used a variety of review methods to ensure the feasibility test data were captured accurately and any unusual data were sufficiently documented. The methods developed were designed to verify implemented procedures were being followed, review the data for unexpected values, and analyze data to support other ORS development work.

Data Review Procedures

Review processes developed in fiscal year 2013 to improve data quality and consistency were further refined for the fiscal year 2014 feasibility tests. The edits used in fiscal year 2013 Phase 3 test were improved and extended to additional elements. The simplest and most well-tested edits were moved into the data capture system while newer and more complicated edits were still run in a secondary program. When an edit flagged in the data capture system (e.g., based on a conflict or expected relationship between two or more elements), the field economist was alerted and required to either verify that the data in question was correct and add documentation, or change the data to clear the edit before marking a schedule complete.

BLS developed a set of tools to flag potentially anomalous values based on expected relationships between elements and similar external data sets such as the DOT. These development environment tools were used on all schedules to supplement the manual review process. For all types of review, reviewers sent questions to the field economist to check the values and coding for inconsistencies identified. Field economists checked interview notes and could call the respondent if needed. Then the field economist added supporting documentation, changed the data, or both. Once the field economist

and the reviewer were satisfied the data were valid, the review was marked complete and the final data was saved to the ORS database.

X. Item-level Response

Item-level response rates were computed for data elements collected in four tests: ORS Only Efficiency Innovations Test, NCS/ORS Joint Collection Test , Central Office Collection Test , and Alternative Modes Test. The item-level response rates were lower in the COC Test than the other tests. The disparity between the COC and other tests was most marked among the arms/legs elements. The table below shows elements with response rates above 97 percent and elements with response rates below 95 percent across all tests.

Item-level response rates were above 97 percent in all four tests for:	Item-level response rates were below 95 percent in all four tests for:
<ul style="list-style-type: none"> • Hearing Conversational Speech • Near Visual Acuity • Driving • Humidity • Noise Intensity • Sitting and/or Standing 	<ul style="list-style-type: none"> • Reaching overhead • Reaching below shoulder • Lift/carry constantly • Lift/carry frequently • Stooping