

O*NET[®] Analyst Occupational Abilities Ratings: Analysis Cycle 9 Results

**Shannon Willison
Suzanne Tsacoumis**

Prepared for: National Center for O*NET Development
700 Wade Avenue, P.O. Box 27625
Raleigh, NC 27605

Prepared under: Contract Number: 2108-009
February 2009

**O*NET ANALYST OCCUPATIONAL ABILITIES RATINGS:
ANALYSIS CYCLE 9 RESULTS**

Table of Contents

| | |
|---|----|
| Introduction..... | 1 |
| Evaluation of Cycle 9 Analyst Ratings..... | 2 |
| Cycle 9 Recommended Data Flags..... | 2 |
| Cycle 9 Interrater Agreement..... | 9 |
| Cycle 9 Interrater Reliability: Across Constructs Within Occupations..... | 9 |
| Cycle 9 Interrater Reliability: Across Occupations Within Constructs..... | 10 |
| Summary..... | 12 |
| References..... | 14 |

List of Tables

| | |
|--|----|
| Table 1. Number of Times Ability Level Flagged as Not Relevant..... | 4 |
| Table 2. Importance Flags Due to Large SE_M | 6 |
| Table 3. Level Flags Due to Large SE_M | 7 |
| Table 4. Interrater Reliabilities and Standard Errors of Measurement for Abilities Across Occupations in Cycles 1 through 9..... | 10 |

O*NET ANALYST OCCUPATIONAL ABILITIES RATINGS: ANALYSIS CYCLE 9 RESULTS

Introduction

The Occupational Information Network (O*NET[®]) is a comprehensive system developed by the U.S. Department of Labor that provides information for 812 occupations within the U.S. economy. This information is maintained in a comprehensive database which was developed to replace the Dictionary of Occupational Titles (DOT) (U.S. Department of Labor, 1991). In order to keep the database current, the National Center for O*NET Development is involved in a continual data collection process aimed at identifying and maintaining current information on the characteristics of workers and jobs. The information that populates the O*NET database is collected primarily from three sources: incumbents, occupational experts, and analysts. Targeted job incumbents provide ratings on occupational tasks, generalized work activities (GWA), knowledge, education and training, work styles, and work context areas. Importance and level information regarding the abilities and skills associated with these occupations is being collected from analysts. It should be noted that there are theoretical or philosophical reasons for preferring one rater group to the other for collecting different types of data. For example, incumbents are generally more familiar with the day-to-day duties of their job; therefore, they are the best source of information regarding tasks and GWAs. In contrast, it is likely that trained analysts understand the ability and skill constructs better than incumbents and therefore should provide the ability and skills data (Tsacoumis, 2007). This report focuses on the ability results only; for Cycle 9 skill results, see Willison and Tsacoumis (2008). Abilities are "... relatively enduring attributes of an individual's capability for performing a particular range of different tasks" (Fleishman, Costanza, & Marshall-Mies, 1999, p. 175). Abilities are sometimes referred to as traits as they tend to remain stable over long periods. The 52 O*NET abilities cover performance applicable to a broad range of jobs in the world's economy. These abilities are grouped into four categories: cognitive, psychomotor, physical, and sensory-perceptual constructs.

To facilitate the ability rating process, analysts are provided relevant occupational information. Trained analysts are responsible for rating the importance and level of the 52 abilities for each of the O*NET occupations. More specifically, eight trained analysts provided ratings for each occupation. For a description of the entire analyst data collection process, including the preparation and distribution of the occupational data, the steps associated with the ratings process, and the collection and management of the ability ratings, see *O*NET Analyst Occupational Abilities Ratings: Procedures* (Donsbach, Tsacoumis, Sager, & Updegraff, 2003).

To ensure a controlled data collection and management process, occupational data is being collected in groups or "analysis cycles." This report describes the results from the data collection process for the ninth analysis cycle of 106 occupations. Reports describing each of the previous cycles are available at <http://www.onetcenter.org/resData.html#waves>. Results for subsequent cycles will be reported in separate reports. For a description of the O*NET Data Collection Publication Schedule see www.onetcenter.org. The O*NET-SOC Codes and Titles included in O*NET Analysis Cycle 9 are presented in Appendix A.

Evaluation of Cycle 9 Analyst Ratings

As previously noted, analysts provided ratings on importance and level of the 52 abilities for each of the 106 occupations in Cycle 9. The mean, standard deviation, and SE_M of the importance and level ratings were computed. These results are presented in Appendix B.

Four sets of analyses were performed to evaluate the ratings that analysts provided. First, we focused on identifying the data that may be difficult to interpret based on limited agreement among raters or because there is an indication that the ability level rating is not relevant for a specific occupation. Thus, a set of recommended suppression criteria was established which flagged: (a) an ability level rating as not relevant to an occupation because of low importance ratings, (b) an ability with too little agreement in importance ratings across raters for a particular occupation, and (c) an ability with too little agreement in level ratings across raters for a particular occupation.

The remaining three sets of analyses focused on computing measures of interrater agreement and interrater reliability. Poor agreement or reliability estimates may be an indication that there is confusion about the constructs, potentially due to either the nature of the definition or rater training. Specifically, the second analysis involved computing the interrater agreement among the eight raters in each rating group. Next, the interrater reliability of the raters was computed to determine the extent to which raters agreed about the order of and relative distance between constructs on a particular scale (i.e., importance or level) within a particular occupation. That is, this analysis provides information regarding the consistency across raters in terms of how they rate the required level or relative importance of the 52 ability constructs to performance in a particular occupation. Finally, another interrater reliability estimate was computed to examine the consistency of ratings across occupations within constructs. In other words, this type of interrater reliability focused on the extent to which raters agree about the order of and relative distance between occupations on a particular scale for a particular construct.

Cycle 9 Recommended Data Flags

Three distinct criteria were established to flag the ability data. All three flags affect the presentation of data within the publicly available O*NET Online (online.onetcenter.org). First, the level rating of an ability was flagged as not relevant for a particular occupation if two or fewer of the eight analysts rated its importance as two or greater. Thus, the level rating of an ability is considered not relevant when that construct is not important for the performance of the particular occupation. For example, in the Cycle 9 data, the level ratings for the Night Vision ability were considered not relevant for Shampooers (39-5093.00) as well as Watch Repairers (49-9064.00) because Night Vision was not considered important for the performance of these two occupations. In this cycle, there were 1,057 not relevant flags (see Table 1). To facilitate interpretation of these results, it should be noted that there are 5,512 sets of ratings (106 occupations x 52 abilities) in the current cycle. Given this, 19.18% (1,057/5,512) of the ability ratings were flagged as not relevant.

Table 1 shows the number of not-relevant flags for ability level. Although many of the abilities in Cycle 9 showed an increase from the most recent cycles in the number of flags, the most common abilities identified as not relevant are consistent with the Cycles 1-8 results. The abilities with the most flags in Cycle 9 include Dynamic Flexibility, Explosive Strength, Night Vision, Spatial Orientation, Peripheral Vision, Sound Localization, and Glare Sensitivity; each of these abilities has received large numbers of flags in previous cycles. Given that these constructs capture fairly specific physical capabilities intuitively not required for many occupations, these results are not surprising.

The remaining two criteria involve the recommended suppression of identifying any ability importance or level mean rating that had a standard error of the mean (SE_M) greater than .51. These criteria were established to capture those ratings deemed to have insufficient agreement across raters. The value of .51 was selected because $1.0/1.96 = .51$. An SE_M greater than .51 means that the upper and lower bounds of the confidence interval are more than one scale point away from the observed mean. The results of these two suppression criteria for importance and level for Cycles 1-9 are presented in Table 2 and 3, respectively. For ease of presentation, the data from Cycles 1 and 2 are averaged and presented in one column. As can be noted, there were no instances in Cycle 9 where the mean importance rating was flagged for insufficient agreement. There were 49 insufficient agreement flags for level ratings. It should be noted that the number of flags indicating insufficient agreement with level ratings in Cycle 9 represents the lowest percentage of insufficient agreement flags compared to all previous cycles.

In Cycle 9, the abilities that were flagged the most for the level criteria were: Spatial Orientation ($n = 5$), Control Precision ($n = 5$), Speed of Closure ($n = 4$) and Wrist-Finger Speed ($n = 4$). All of these abilities reflect small increases (i.e., 2 to 4) in the number of flags compared to Cycle 8; however, they all have received a similarly large number of flags in at least two previous cycles. In addition to these four abilities, only six other abilities increased in the number of level flags, all of which were increases of four points or less.

Although there were some small increases in the number of level flags for a few abilities, most abilities received fewer flags in Cycle 9 or remained the same. All three of the abilities identified as receiving the most flags from Cycle 8 (i.e., Auditory Attention, Visual Color Discrimination, and Glare Sensitivity) decreased dramatically in Cycle 9. For example, Auditory Attention showed a nine point decrease in the number of level flags (from 11 to 2). Overall, 19 abilities showed decreases (ranging from 1 to 9 points) and 23 abilities remained the same from Cycle 8 to Cycle 9.

Cycles 1 through 6 showed an encouraging trend reflecting a decrease in the percentage of ability level ratings receiving flags. Although Cycle 7 did not follow this trend, Cycles 8 and 9 returned to a lower percentage of flags (1.40% and .89%, respectively). These findings suggest there remains a high level of agreement among the analysts. However, it continues to be advisable to monitor the elements that have been noted as potentially problematic and to continue to address these issues via follow-up training with analysts in future cycles.

The detailed results of the recommended data flags and suppression criteria are depicted by the shaded cells in the results presented in Appendix B.

Table 1. Number of Times Ability Level Flagged as Not Relevant

| Element Name | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
|------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Oral Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Oral Expression | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written Expression | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fluency of Ideas | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Originality | 0 | 7 | 2 | 0 | 0 | 1 | 0 | 1 | 0 |
| Problem Sensitivity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Deductive Reasoning | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Inductive Reasoning | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Information Ordering | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Category Flexibility | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mathematical Reasoning | 0 | 6 | 4 | 1 | 3 | 4 | 4 | 6 | 1 |
| Number Facility | 3 | 5 | 0 | 1 | 1 | 3 | 2 | 4 | 1 |
| Memorization | 0 | 1 | 0 | 0 | 0 | 5 | 3 | 1 | 1 |
| Speed of Closure | 0 | 2 | 3 | 0 | 0 | 0 | 1 | 3 | 0 |
| Flexibility of Closure | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Perceptual Speed | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 3 | 0 |
| Spatial Orientation | 36 | 48 | 66 | 81 | 54 | 48 | 35 | 33 | 62 |
| Visualization | 0 | 6 | 3 | 0 | 2 | 0 | 0 | 1 | 0 |
| Selective Attention | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Time Sharing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| Arm-Hand Steadiness | 9 | 14 | 11 | 49 | 15 | 14 | 7 | 6 | 18 |
| Manual Dexterity | 9 | 19 | 9 | 54 | 16 | 16 | 7 | 6 | 19 |
| Finger Dexterity | 0 | 6 | 3 | 0 | 1 | 0 | 0 | 1 | 0 |
| Control Precision | 6 | 19 | 13 | 48 | 16 | 12 | 7 | 6 | 14 |
| Multilimb Coordination | 13 | 31 | 23 | 50 | 25 | 15 | 10 | 7 | 25 |
| Response Orientation | 30 | 72 | 50 | 66 | 39 | 28 | 16 | 14 | 35 |
| Rate Control | 35 | 88 | 57 | 73 | 43 | 29 | 18 | 16 | 39 |
| Reaction Time | 27 | 65 | 40 | 66 | 39 | 23 | 13 | 14 | 35 |
| Wrist-Finger Speed | 26 | 50 | 54 | 76 | 50 | 32 | 21 | 17 | 41 |
| Speed of Limb Movement | 28 | 57 | 49 | 65 | 47 | 34 | 20 | 22 | 43 |

Continued on next page

Table 1. (Continued)

| Element Name | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
|--|----------------------|-------------------------|-------------------------|-------------------------|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Static Strength | 21 | 38 | 33 | 56 | 36 | 23 | 15 | 11 | 35 |
| Explosive Strength | 44 | 104 | 90 | 93 | 85 | 93 | 93 | 80 | 87 |
| Dynamic Strength | 28 | 61 | 46 | 65 | 42 | 28 | 15 | 17 | 46 |
| Trunk Strength | 8 | 16 | 23 | 29 | 30 | 21 | 6 | 4 | 26 |
| Stamina | 21 | 42 | 38 | 58 | 38 | 25 | 14 | 12 | 42 |
| Extent Flexibility | 22 | 47 | 36 | 64 | 40 | 24 | 15 | 13 | 44 |
| Dynamic Flexibility | 52 | 104 | 102 | 98 | 90 | 99 | 99 | 92 | 102 |
| Gross Body Coordination | 21 | 46 | 36 | 58 | 38 | 25 | 14 | 13 | 42 |
| Gross Body Equilibrium | 27 | 67 | 53 | 61 | 44 | 26 | 14 | 13 | 42 |
| Near Vision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Far Vision | 0 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| Visual Color Discrimination | 2 | 18 | 7 | 2 | 1 | 0 | 0 | 2 | 0 |
| Night Vision | 44 | 99 | 83 | 83 | 58 | 53 | 40 | 35 | 64 |
| Peripheral Vision | 44 | 85 | 79 | 82 | 55 | 54 | 41 | 32 | 65 |
| Depth Perception | 11 | 21 | 24 | 35 | 12 | 13 | 6 | 4 | 4 |
| Glare Sensitivity | 41 | 93 | 68 | 84 | 48 | 45 | 30 | 28 | 59 |
| Hearing Sensitivity | 2 | 39 | 32 | 16 | 3 | 0 | 0 | 3 | 0 |
| Auditory Attention | 2 | 10 | 4 | 1 | 2 | 0 | 0 | 1 | 0 |
| Sound Localization | 44 | 95 | 83 | 84 | 54 | 52 | 39 | 32 | 65 |
| Speech Recognition | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speech Clarity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total Flags out of all possible ratings | 23.36% (656/2808) | 22.74% (1,490/6,552) | 21.67% (1,228/5,668) | 30.75% (1,599/5,200) | 21.70% (1,027/4,732) | 16.25% (845/5,200) | 11.52% (605/5,252) | 10.69% (556/5,200) | 19.18% (1057/5512) |

Table 2. Importance Flags Due to Large SE_M

| Element Name | Frequency SE_M Importance > .51 | | | | | | | |
|-----------------------------|-----------------------------------|---------|---------|---------|---------|---------|---------|---------|
| | Cycle1&2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
| Oral Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Oral Expression | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written Expression | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fluency of Ideas | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Originality | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Problem Sensitivity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Deductive Reasoning | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Inductive Reasoning | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Information Ordering | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Category Flexibility | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mathematical Reasoning | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Number Facility | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Memorization | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speed of Closure | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Flexibility of Closure | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Perceptual Speed | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Spatial Orientation | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Visualization | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Selective Attention | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Time Sharing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Arm-Hand Steadiness | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Manual Dexterity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Finger Dexterity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Control Precision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Multilimb Coordination | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Response Orientation | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Rate Control | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Reaction Time | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Wrist-Finger Speed | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speed of Limb Movement | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Static Strength | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Explosive Strength | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dynamic Strength | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Trunk Strength | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stamina | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Extent Flexibility | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dynamic Flexibility | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gross Body Coordination | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gross Body Equilibrium | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Near Vision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Far Vision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Visual Color Discrimination | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Night Vision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Peripheral Vision | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Continued on next page

Table 2. (Continued)

| Element Name | Frequency SE_M Importance >. 51 | | | | | | | |
|---------------------|-----------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | Cycle1&2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
| Depth Perception | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Glare Sensitivity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Hearing Sensitivity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Auditory Attention | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sound Localization | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speech Recognition | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speech Clarity | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 0% (7/9360) | 0% (0/5668) | 0% (0/5200) | 0% (0/4732) | 0% (0/5200) | 0% (0/5252) | 0% (0/5200) | 0% (0/5512) |

Table 3. Level Flags Due to Large SE_M

| Element Name | Frequency SE_M Level >. 51 | | | | | | | |
|------------------|------------------------------|---------|---------|---------|---------|---------|---------|---------|
| | Cycle1&2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
| Oral | | | | | | | | |
| Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written | | | | | | | | |
| Comprehension | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Oral Expression | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Written | | | | | | | | |
| Expression | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| Fluency of Ideas | 7.5 | 10 | 1 | 0 | 0 | 1 | 0 | 0 |
| Originality | 2 | 8 | 0 | 0 | 0 | 1 | 0 | 0 |
| Problem | | | | | | | | |
| Sensitivity | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| Deductive | | | | | | | | |
| Reasoning | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Inductive | | | | | | | | |
| Reasoning | 0.5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Information | | | | | | | | |
| Ordering | 0.5 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Category | | | | | | | | |
| Flexibility | 1 | 10 | 0 | 1 | 0 | 0 | 0 | 0 |
| Mathematical | | | | | | | | |
| Reasoning | 4 | 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Number Facility | 8 | 10 | 9 | 3 | 1 | 2 | 0 | 0 |
| Memorization | 10.5 | 18 | 1 | 5 | 3 | 5 | 2 | 0 |
| Speed of Closure | 18 | 29 | 5 | 10 | 4 | 4 | 2 | 4 |
| Flexibility of | | | | | | | | |
| Closure | 21.5 | 35 | 22 | 5 | 1 | 1 | 1 | 0 |
| Perceptual Speed | 13.5 | 15 | 9 | 3 | 0 | 1 | 0 | 0 |
| Spatial | | | | | | | | |
| Orientation | 5 | 6 | 1 | 1 | 1 | 4 | 3 | 5 |
| Visualization | 16 | 26 | 16 | 6 | 4 | 1 | 0 | 1 |
| Selective | | | | | | | | |
| Attention | 1 | 6 | 0 | 2 | 0 | 1 | 0 | 0 |
| Time Sharing | 3 | 7 | 0 | 1 | 0 | 0 | 0 | 0 |
| Arm-Hand | | | | | | | | |
| Steadiness | 2.5 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| Manual | | | | | | | | |
| Dexterity | 7 | 9 | 2 | 4 | 0 | 0 | 0 | 2 |
| Finger Dexterity | 10 | 9 | 0 | 3 | 0 | 0 | 0 | 3 |

Continued on next page

Table 3. (Continued)

| Element Name | Frequency SE_M Level > 51 | | | | | | | |
|-------------------|-----------------------------|---------------------|---------------------|---------------------|--------------------|---------------------|--------------------|--------------------|
| | Cycle1&2 | Cycle 3 | Cycle 4 | Cycle 5 | Cycle 6 | Cycle 7 | Cycle 8 | Cycle 9 |
| Control | | | | | | | | |
| Precision | 4.5 | 8 | 4 | 5 | 1 | 1 | 1 | 5 |
| Multilimb | | | | | | | | |
| Coordination | 4 | 5 | 1 | 5 | 1 | 0 | 0 | 2 |
| Response | | | | | | | | |
| Orientation | 7 | 11 | 4 | 3 | 1 | 5 | 3 | 1 |
| Rate Control | 2.5 | 6 | 0 | 3 | 1 | 1 | 2 | 2 |
| Reaction Time | 12.5 | 19 | 4 | 4 | 3 | 13 | 2 | 2 |
| Wrist-Finger | | | | | | | | |
| Speed | 27 | 7 | 1 | 2 | 2 | 15 | 0 | 4 |
| Speed of Limb | | | | | | | | |
| Movement | 2.5 | 13 | 2 | 1 | 1 | 7 | 4 | 3 |
| Static Strength | 5 | 12 | 4 | 0 | 0 | 3 | 4 | 1 |
| Explosive | | | | | | | | |
| Strength | 3 | 6 | 0 | 1 | 3 | 2 | 2 | 0 |
| Dynamic | | | | | | | | |
| Strength | 5.5 | 9 | 2 | 2 | 2 | 2 | 3 | 1 |
| Trunk Strength | 1.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stamina | 2.5 | 3 | 1 | 1 | 0 | 1 | 0 | 0 |
| Extent | | | | | | | | |
| Flexibility | 7 | 14 | 0 | 5 | 4 | 10 | 4 | 1 |
| Dynamic | | | | | | | | |
| Flexibility | 4 | 0 | 0 | 0 | 0 | 1 | 2 | 0 |
| Gross Body | | | | | | | | |
| Coordination | 0 | 2 | 1 | 1 | 0 | 1 | 1 | 0 |
| Gross Body | | | | | | | | |
| Equilibrium | 2 | 5 | 1 | 1 | 1 | 1 | 1 | 0 |
| Near Vision | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| Far Vision | 15 | 20 | 3 | 9 | 0 | 1 | 0 | 2 |
| Visual Color | | | | | | | | |
| Discrimination | 10.5 | 18 | 7 | 4 | 1 | 2 | 8 | 1 |
| Night Vision | 3.5 | 1 | 0 | 0 | 3 | 3 | 4 | 1 |
| Peripheral | | | | | | | | |
| Vision | 1.5 | 3 | 0 | 2 | 1 | 6 | 1 | 0 |
| Depth Perception | 0.5 | 8 | 2 | 1 | 0 | 0 | 0 | 0 |
| Glare Sensitivity | 2 | 9 | 1 | 0 | 0 | 2 | 6 | 2 |
| Hearing | | | | | | | | |
| Sensitivity | 4.5 | 10 | 5 | 4 | 1 | 2 | 2 | 1 |
| Auditory | | | | | | | | |
| Attention | 5 | 23 | 0 | 2 | 6 | 10 | 11 | 2 |
| Sound | | | | | | | | |
| Localization | 5 | 8 | 4 | 3 | 2 | 5 | 3 | 2 |
| Speech | | | | | | | | |
| Recognition | 4 | 3 | 4 | 2 | 1 | 3 | 0 | 0 |
| Speech Clarity | 1 | 6 | 0 | 1 | 0 | 0 | 0 | 0 |
| TOTAL | 5.81% (544/9360) | 7.82% (443/5668) | 2.31% (120/5200) | 2.30% (109/4732) | 0.94% (49/5200) | 2.27% (119/5252) | 1.40% (73/5200) | 0.89% (49/5512) |

Cycle 9 Interrater Agreement

Interrater agreement was computed to examine the level of absolute agreement among the analysts in ratings within a construct for a particular occupation. For example, these indices identified the extent to which eight raters provided the same rating regarding the level of the ability *Written Comprehension* required to perform a particular occupation. To look at the agreement, we calculated the standard deviation (*SD*) of ratings across analysts for a given construct and scale for each occupation and the *SE_M* of these ratings. For both indices, lower values indicate higher agreement, and vice versa.

A summary of these results is shown in Appendix C. The columns labeled “Mean of *M_s*” show the mean of the analyst mean importance and level ratings across the 52 abilities for each occupation.¹ The columns labeled “Median of *SD_s*” show the median of the *SD_s* associated with each mean importance and level rating across the 52 abilities for each occupation. Finally, the columns labeled “Median of *SE_{M_s}*” show the median of the *SE_{M_s}* associated with each mean importance and level rating across the 52 abilities for each occupation.

The importance ratings across all occupations had a median *SD* of .52 and a median *SE_M* of .18. The level ratings across occupations had a median *SD* of .67 and a median *SE_M* of .24. These results for importance and level reveal that raters agreed similarly in this cycle compared to agreement in Cycle 8; moreover, these values are consistent with agreement in Cycles 1-7. Overall, while the values are generally greater (indicating less agreement) for the level than they are for the importance, the results indicate that the ratings made by the analysts were consistent for both scales.

Cycle 9 Interrater Reliability: Across Constructs Within Occupations

To examine the interrater reliability of the Cycle 9 ratings we calculated the interclass correlations (*ICC* [3, *k*]; Shrout & Fleiss, 1979) among the analyst’s ratings to look at consistency across constructs within occupations. As mentioned previously, this calculation examines the similarity in the rank ordering and relative distance between the abilities on a particular scale within an occupation. Our target level of interrater reliability is a median *ICC* (3, *k*) of .80 or greater. The value of .80 is judged to be a good rule-of-thumb that has been used previously in the O*NET context (e.g., McCloy, Waugh, & Medsker, April 1998).

The results of these analyses are presented in Appendix D. The data revealed high levels of interrater reliability across the 106 Cycle 9 occupations. Specifically, the mean and median *ICC* for importance ratings for the abilities across the occupations was .95 and .96 (*SD* = .03), respectively. The mean and median *ICC* for the level ratings were also .95 and .96, respectively (*SD* = .03). The reliability for both the importance and level ratings exceeded the median target coefficient value of .80. Results also indicate that for the most part, occupations with the lowest reliability coefficients for importance had the lowest values for level ratings. This may be due to the skip pattern which forces a “0” for level if the ability is rated not important. Overall, the results support a good level of agreement in the analysts’ ratings.

¹ While the mean is not a measure of agreement, it can affect the potential range of the *SD* and *SE_M*.

Cycle 9 Interrater Reliability: Across Occupations Within Constructs

Another effective way to evaluate the reliability of the analyst's ratings is to look at the consistency across occupations within constructs. This type of reliability is the extent to which raters agree about the order of and relative distance among occupations on a particular scale for a particular construct. For example, is there consistency across raters in how they differentiate among occupations on the required level of the ability *Oral Comprehension*? To make this evaluation, Shrout and Fleiss' (1979) $ICC(3, k)$ is calculated for each construct on each scale (instead of for each occupation on each scale as described above). For example, each of the 52 ability importance scale ratings will have a reliability value. The target level of interrater reliability for this coefficient is that the median $ICC(3, k)$ across the construct ratings for a particular domain on a particular scale be .80 or greater (e.g., the median reliability across 52 ability level ratings should be at least .80). The value of .80 is judged to be a good rule-of-thumb that has been used in the O*NET context before (e.g., McCloy, Waugh, & Medsker, April 1998).

This reliability analysis was conducted for abilities on all 886 occupations from Cycles 1 through 9 and results are presented in Table 4². Note that one occupation was rated in two different cycles; therefore, the reliability analyses are based on 887 rating targets. The values in the columns titled $ICC(C,1)$ reflect the single rater reliabilities, whereas the values in the columns titled $ICC(C,8)$ reflect the reliability for eight raters. Overall for the abilities, the median $ICC(C,8)$ across the construct ratings for importance was .87 ($M = .84, SD = .11$) and for level was .90 ($M = .87, SD = .09$). This indicates that on the whole, the reliabilities for achieved the target level. However, there are some low reliabilities to note.

Table 4. Interrater Reliabilities and Standard Errors of Measurement for Abilities Across Occupations in Cycles 1 through 9

| Ability | | Cycles 1 through 9 ($N = 887$) | | | | | |
|---------|------------------------|----------------------------------|------------|-------|------------|------------|-------|
| | | Importance | | | Level | | |
| | | $ICC(C,1)$ | $ICC(C,8)$ | s_E | $ICC(C,1)$ | $ICC(C,8)$ | s_E |
| 1 | Oral Comprehension | 0.41 | 0.85 | 0.18 | 0.50 | 0.89 | 0.21 |
| 2 | Written Comprehension | 0.50 | 0.89 | 0.19 | 0.60 | 0.92 | 0.21 |
| 3 | Oral Expression | 0.49 | 0.88 | 0.18 | 0.55 | 0.91 | 0.20 |
| 4 | Written Expression | 0.49 | 0.89 | 0.20 | 0.63 | 0.93 | 0.24 |
| 5 | Fluency of Ideas | 0.40 | 0.84 | 0.21 | 0.47 | 0.88 | 0.28 |
| 6 | Originality | 0.46 | 0.87 | 0.21 | 0.53 | 0.90 | 0.27 |
| 7 | Problem Sensitivity | 0.36 | 0.82 | 0.19 | 0.49 | 0.88 | 0.23 |
| 8 | Deductive Reasoning | 0.35 | 0.81 | 0.18 | 0.51 | 0.89 | 0.23 |
| 9 | Inductive Reasoning | 0.41 | 0.85 | 0.19 | 0.53 | 0.90 | 0.23 |
| 10 | Information Ordering | 0.21 | 0.68 | 0.20 | 0.33 | 0.80 | 0.23 |
| 11 | Category Flexibility | 0.20 | 0.66 | 0.20 | 0.29 | 0.77 | 0.26 |
| 12 | Mathematical Reasoning | 0.46 | 0.87 | 0.23 | 0.55 | 0.91 | 0.31 |
| 13 | Number Facility | 0.38 | 0.83 | 0.23 | 0.49 | 0.88 | 0.32 |
| 14 | Memorization | 0.19 | 0.65 | 0.23 | 0.25 | 0.72 | 0.35 |
| 15 | Speed of Closure | 0.22 | 0.70 | 0.25 | 0.28 | 0.76 | 0.36 |

Continued on next page

² The current total of occupations contained within the O*NET system is 812. A number of occupations for which ratings were developed in initial cycles were rolled-up or aggregated during the transition to the 2006 O*NET taxonomy.

Table 4. (Continued)

| | | Cycles 1 through 9 (N = 887) | | | | | |
|---------|-----------------------------|------------------------------|----------|-------|----------|----------|-------|
| | | Importance | | | Level | | |
| Ability | | ICC(C,1) | ICC(C,8) | s_E | ICC(C,1) | ICC(C,8) | s_E |
| 16 | Flexibility of Closure | 0.21 | 0.68 | 0.26 | 0.27 | 0.74 | 0.34 |
| 17 | Perceptual Speed | 0.28 | 0.75 | 0.25 | 0.27 | 0.75 | 0.32 |
| 18 | Spatial Orientation | 0.56 | 0.91 | 0.20 | 0.57 | 0.91 | 0.28 |
| 19 | Visualization | 0.39 | 0.83 | 0.24 | 0.42 | 0.85 | 0.33 |
| 20 | Selective Attention | 0.13 | 0.55 | 0.21 | 0.15 | 0.59 | 0.26 |
| 21 | Time Sharing | 0.18 | 0.63 | 0.23 | 0.21 | 0.68 | 0.28 |
| 22 | Arm-Hand Steadiness | 0.73 | 0.95 | 0.20 | 0.71 | 0.95 | 0.26 |
| 23 | Manual Dexterity | 0.71 | 0.95 | 0.20 | 0.66 | 0.94 | 0.30 |
| 24 | Finger Dexterity | 0.44 | 0.86 | 0.23 | 0.44 | 0.86 | 0.30 |
| 25 | Control Precision | 0.72 | 0.95 | 0.20 | 0.69 | 0.95 | 0.29 |
| 26 | Multilimb Coordination | 0.70 | 0.95 | 0.21 | 0.70 | 0.95 | 0.28 |
| 27 | Response Orientation | 0.62 | 0.93 | 0.20 | 0.66 | 0.94 | 0.30 |
| 28 | Rate Control | 0.69 | 0.95 | 0.18 | 0.69 | 0.95 | 0.26 |
| 29 | Reaction Time | 0.71 | 0.95 | 0.19 | 0.71 | 0.95 | 0.30 |
| 30 | Wrist-Finger Speed | 0.43 | 0.86 | 0.22 | 0.43 | 0.86 | 0.34 |
| 31 | Speed of Limb Movement | 0.58 | 0.92 | 0.19 | 0.60 | 0.92 | 0.27 |
| 32 | Static Strength | 0.69 | 0.95 | 0.20 | 0.74 | 0.96 | 0.28 |
| 33 | Explosive Strength | 0.38 | 0.83 | 0.13 | 0.40 | 0.84 | 0.22 |
| 34 | Dynamic Strength | 0.59 | 0.92 | 0.19 | 0.63 | 0.93 | 0.27 |
| 35 | Trunk Strength | 0.63 | 0.93 | 0.20 | 0.66 | 0.94 | 0.26 |
| 36 | Stamina | 0.66 | 0.94 | 0.18 | 0.65 | 0.94 | 0.26 |
| 37 | Extent Flexibility | 0.72 | 0.95 | 0.18 | 0.73 | 0.96 | 0.29 |
| 38 | Dynamic Flexibility | 0.19 | 0.66 | 0.11 | 0.21 | 0.68 | 0.18 |
| 39 | Gross Body Coordination | 0.63 | 0.93 | 0.19 | 0.67 | 0.94 | 0.25 |
| 40 | Gross Body Equilibrium | 0.62 | 0.93 | 0.18 | 0.62 | 0.93 | 0.26 |
| 41 | Near Vision | 0.15 | 0.58 | 0.19 | 0.33 | 0.80 | 0.24 |
| 42 | Far Vision | 0.35 | 0.81 | 0.23 | 0.32 | 0.79 | 0.33 |
| 43 | Visual Color Discrimination | 0.38 | 0.83 | 0.24 | 0.41 | 0.85 | 0.34 |
| 44 | Night Vision | 0.51 | 0.89 | 0.16 | 0.53 | 0.90 | 0.24 |
| 45 | Peripheral Vision | 0.58 | 0.92 | 0.16 | 0.59 | 0.92 | 0.23 |
| 46 | Depth Perception | 0.60 | 0.92 | 0.20 | 0.60 | 0.92 | 0.28 |
| 47 | Glare Sensitivity | 0.63 | 0.93 | 0.16 | 0.67 | 0.94 | 0.24 |
| 48 | Hearing Sensitivity | 0.43 | 0.86 | 0.24 | 0.44 | 0.86 | 0.33 |
| 49 | Auditory Attention | 0.35 | 0.81 | 0.23 | 0.36 | 0.82 | 0.34 |
| 50 | Sound Localization | 0.49 | 0.88 | 0.17 | 0.52 | 0.90 | 0.26 |
| 51 | Speech Recognition | 0.20 | 0.67 | 0.21 | 0.24 | 0.72 | 0.30 |
| 52 | Speech Clarity | 0.37 | 0.83 | 0.19 | 0.38 | 0.83 | 0.28 |

Note. These ICCs indicate how consistently raters rated (rank ordered) occupations on a given ability.

s_E = Standard error of measurement = Observed score variance times the square root of one minus ICC(C,8).

The lowest ability ICC(C,8) reliabilities were found for Selective Attention, Near Vision, and Time Sharing; none of the reliabilities for these constructs had importance reliabilities over .63. The level reliabilities for Selective Attention and Time Sharing were also low, failing to exceed .68. Low variance is a contributing factor to low reliabilities. Also, these reliabilities are consistent with the reliabilities for these constructs in the previous cycle.

In addition, a number of abilities had ICC(C,8) reliabilities for importance that were less than .70: Information Ordering, Category Flexibility, Memorization, Flexibility of Closure, Dynamic Flexibility, and Speech Recognition. ICC(C,8) reliabilities for these ability level ratings were .70 or greater for all of these abilities except Dynamic flexibility (.68). These differences in ability reliabilities for importance and level likely reflect high agreement but lack of variability in the ratings of these constructs across occupations on importance, yet high agreement and high variation in the ratings of these constructs across jobs on level.

Despite a few cases of low reliability, 85% of the ability constructs had high ICC(C,8) reliabilities for both importance and level (i.e., $\geq .70$). In fact, there were 19 ability levels with reliabilities greater than .90 for both importance and level (e.g., Spatial Orientation).

Comparisons with interrater reliabilities for abilities obtained from all previous cycles indicate that for almost all elements, the ICC(C,8) reliability estimates for both importance and level from Cycles 1-9 were consistent (within .01) with those from previous cycles. The two exceptions were small decreases in reliability for Selective Attention (importance decreased .03 and level decreased .02) and a small increase in level reliability for Explosive Strength (.02). Although there were no large increases to note, it is important to realize that increases in the size of reliability coefficients are limited because of the relatively large coefficients already obtained on the majority of constructs.

Keep in mind that some variation in calculated values is likely to occur by chance. As previously described, the goal was for the ICC(C,8) reliabilities to have a median value across constructs of .80 or greater, which was achieved for both importance and level (.87 and .90, respectively). These results suggest that there was a good level of agreement among the raters with respect to the order and relative distance among occupations on particular constructs for importance and level.

Summary

The main findings of the analysis of Cycle 9 analyst ratings were as follows:

- Approximately 19 % of the ability ratings were flagged because the construct was considered not important for performance. This is an increase from the most recent cycles, but less than cycles 1-5. The constructs that were flagged were very similar to those flagged in previous cycles.
- Less than 1 percent of the level ratings were flagged based on a SE_M greater than .51; this is the lowest percentage compared to all previous cycles.
- There was strong interrater agreement for this cycle as evidenced by the overall low medians of SE_{MS} .
- All within-occupation ICC reliabilities were well above the target value of .80 (McCloy, Waugh, & Medsker, April 1998). These high levels of interrater reliability indicate that the analysts rank ordered the abilities within each occupation similarly on both importance and level.

- Index interrater reliability calculated at the end of Cycle 9 was high and did not vary greatly from one occupation to the next.
- The importance and level median across-occupation ICC reliabilities were above the target value of .80. These high levels of interrater reliability indicate that analysts rank ordered occupations within each ability similarly on both importance and level.

These results indicate that the analysts were well trained and understand the abilities and associated definitions. Review training for returning analysts and, if required, new analyst training will continue to occur prior to each new cycle, focusing on particular constructs identified in the four sets of analyses. Agreement was high and there is clear evidence regarding the quality of the data.

References

- Donsbach, J., Tsacoumis, S., Sager, C., & Updegraff, J. (2003). *O*NET analyst occupational abilities ratings: Procedures* (DFR-03-22). Alexandria, VA: Human Resources Research Organization.
- Fleishman, E.A., Costanza, D. P., & Marshall-Mies, J. (1999). Abilities. In N.G. Peterson, M.D. Mumford, W. C. Borman, P. R. Jeanneret, & E. A. Fleishman (Eds.), *An occupational information system for the 21st century: The development of O*NET* (p.175-195). Washington D.C.: American Psychological Association.
- McCloy, R., Waugh, G., & Medsker, G. (1998, April). *Determining the occupational reinforcer patterns for O*NET occupational units*. Alexandria, VA: Human Resources Research Organization.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428
- Tsacoumis, S. (2007). The feasibility of using O*NET to study skill changes. Presented at the Workshop on Research Evidence Related to Future Skill Demands organized by the National Academies Center for Education, Washington, D.C.
- Willison, S., & Tsacoumis, S. (2008). *O*NET analyst occupational skills ratings: Analysis Cycle 9 results*. Alexandria, VA: Human Resources Research Organization.
- U.S. Department of Labor (1991). Dictionary of occupational titles. (Rev. 4th ed.). Washington, DC: U.S. Government Printing Office.