

## **O\*NET Analyst Occupational Abilities Ratings: Analysis Cycle 6 Results**

**Suzanne Tsacoumis  
Carrie N. Byrum**

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

*Prepared under:* Contract Number: 2108-009  
December 2006

**O\*NET ANALYST OCCUPATIONAL ABILITIES RATINGS:  
ANALYSIS CYCLE 6 RESULTS**

**Table of Contents**

Introduction.....	1
Evaluation of Cycle 6 Analyst Ratings.....	2
Cycle 6 Recommended Data Flags.....	2
Cycle 6 Interrater Agreement.....	6
Cycle 6 Interrater Reliability: Across Constructs Within Occupations.....	7
Cycle 6 Interrater Reliability: Across Occupations Within Constructs.....	7
Summary.....	10
References.....	11

**List of Tables**

Table 1. Number of Times Ability Level Flagged as Not Relevant.....	3
Table 2. Ability Flags Due to Large $SE_M$ .....	5
Table 3. Interrater Reliabilities and Standard Errors of Measurement Across Cycle 1, 2, 3, 4, 5, and 6 Occupations.....	8

# O\*NET ANALYST OCCUPATIONAL ABILITIES RATINGS: ANALYSIS CYCLE 6 RESULTS

## Introduction

The Occupational Information Network (O\*NET) is a comprehensive system developed by the U.S. Department of Labor that provides information about nearly 1,000 occupations within the U.S. economy. The National Center for O\*NET Development is in the process of collecting occupational data for over 900 occupations. The data collection effort includes job incumbent ratings on occupational tasks, skills, generalized work activities (GWA), knowledge, education and training, work styles, and work context areas. Importance and level information regarding the abilities 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's likely that trained analysts understand the ability constructs better than incumbents and therefore should provide the ability data. 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 sixth analysis cycle of 100 occupations. Results for Cycle 1 are presented in Noble, Sager, Tsacoumis, Updegraff, & Donsbach (2003). Cycle 2 and Cycle 3 results are presented in Noble & Tsacoumis (2004) and Noble & Tsacoumis (2005), respectively. Finally, Cycle 4 and Cycle 5 results can be found in Byrum & Tsacoumis (2005) and Byrum & Tsacoumis (2006), respectively. 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](http://www.onetcenter.org). The O\*NET-SOC Codes and Titles included in O\*NET Analysis Cycle 6 are presented in Appendix A.

## Evaluation of Cycle 6 Analyst Ratings

As mentioned above, analysts provided ratings on importance and level of the 52 abilities for each of the 100 occupations in Cycle 6. 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 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 ability 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 within a particular occupation. That is, this analysis provides information regarding the consistency across raters in terms of how they rate the 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 6 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](http://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 ability is not important for the performance of the particular occupation. For example, in the Cycle 6 data, the level ratings for the Trunk Strength ability were considered not relevant for Statisticians (15-2041.00) as well as Medical Transcriptionists (31-9094.00) because Trunk Strength was not considered important for the performance of these two occupations. In this cycle, there were 845 not relevant flags (see Table 1). To facilitate interpretation of these results, it should be noted that there are 5,200 sets of ratings (100 occupations x 52 abilities) in the current cycle. Given this, 16.25% (845/5,200) of the ability ratings were flagged as not relevant.

As can be seen in Table 1, the most common abilities identified as not relevant remain consistent with the Cycles 1-5 results. The abilities with the most flags in Cycle 6 include Dynamic Flexibility, Explosive Strength, Peripheral Vision, Night Vision, and Sound Localization; each of these abilities has received large numbers of flags in earlier cycles. Given that these constructs capture fairly specific physical capabilities intuitively not required for many occupations, these results are not surprising.

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

Element Name	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Oral Comprehension	0	0	0	0	0
Written Comprehension	0	0	0	0	0
Oral Expression	0	0	0	0	0
Written Expression	0	0	0	0	0
Fluency of Ideas	2	0	0	0	0
Originality	7	2	0	0	1
Problem Sensitivity	0	0	0	0	0
Deductive Reasoning	0	0	0	0	0
Inductive Reasoning	0	0	0	0	0
Information Ordering	0	0	0	0	0
Category Flexibility	0	0	0	0	0
Mathematical Reasoning	6	4	1	3	4
Number Facility	5	0	1	1	3
Memorization	1	0	0	0	5
Speed of Closure	2	3	0	0	0
Flexibility of Closure	2	0	0	0	0
Perceptual Speed	1	1	0	0	0
Spatial Orientation	48	66	81	54	48
Visualization	6	3	0	2	0
Selective Attention	0	0	0	0	0
Time Sharing	0	0	0	0	0
Arm-Hand Steadiness	14	11	49	15	14
Manual Dexterity	19	9	54	16	16
Finger Dexterity	6	3	0	1	0
Control Precision	19	13	48	16	12
Multilimb Coordination	31	23	50	25	15
Response Orientation	72	50	66	39	28
Rate Control	88	57	73	43	29
Reaction Time	65	40	66	39	23
Wrist-Finger Speed	50	54	76	50	32
Speed of Limb Movement	57	49	65	47	34
Static Strength	38	33	56	36	23
Explosive Strength	104	90	93	85	93
Dynamic Strength	61	46	65	42	28
Trunk Strength	16	23	29	30	21
Stamina	42	38	58	38	25
Extent Flexibility	47	36	64	40	24

Continued on next page.

Element Name	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Dynamic Flexibility	104	102	98	90	99
Gross Body Coordination	46	36	58	38	25
Gross Body Equilibrium	67	53	61	44	26
Near Vision	0	0	0	0	0
Far Vision	4	3	0	0	0
Visual Color Discrimination	18	7	2	1	0
Night Vision	99	83	83	58	53
Peripheral Vision	85	79	82	55	54
Depth Perception	21	24	35	12	13
Glare Sensitivity	93	68	84	48	45
Hearing Sensitivity	39	32	16	3	0
Auditory Attention	10	4	1	2	0
Sound Localization	95	83	84	54	52
Speech Recognition	0	0	0	0	0
Speech Clarity	0	0	0	0	0
Total Flags out of all possible ratings	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)

The remaining two criteria involve the recommended suppression of identifying any ability mean or level importance 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 are presented in Table 2. As can be noted, there were no instances where the mean importance rating was flagged for insufficient agreement. There were 49 insufficient agreement flags for level ratings, 17 of these flagged constructs also had ability level ratings flagged as not relevant (34.69% of 49). It should be noted that the number of flags indicating insufficient agreement substantially declined between Cycle 5 and Cycle 6. As such, Cycle 5 had 91 occupations and 109 level flags for insufficient agreement (2.30%) while in Cycle 6 there were 100 occupations and only 49 flags for level (.94%).

In Cycle 6, the abilities that were flagged the most for the level criteria included: Auditory Attention (n=6), Extent Flexibility (n=4), Speed of Closure (n=4), and Visualization (n=4). In some cases, the abilities with the most flags in Cycle 6 also received many flags in the previous five cycles. However, there are a few points of note. First, there was a substantial drop in the number of flags for Speed of Closure (n=4) and Far Vision (n=0). In Cycle 5, Speed of Closure (n=10), Far Vision (n=9), and Extent Flexibility (n=5) all experienced an increase of at least five ability flags due to large  $SE_M$ . All three of these constructs have been shown to be problematic for raters over the previous five cycles. As a result, the small increase in flags during Cycle 5 may have been attributable, in part, to the introduction of new analysts. The decrease in Cycle 6 of flags for Speed of Closure (n=4) and Far Vision (n=0) may be due to the experience, targeted-training, and feedback received by raters. However, while the number of flags is small, Extent Flexibility remains one of the abilities with the most flags in Cycle 6. This ability should be monitored in Cycle 7 to determine if further training is required.

**Table 2. Ability Flags Due to Large  $SE_M$**

Element Name	Frequency $SE_M$ Importance > 51					Frequency $SE_M$ Level > 51				
	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Oral Comprehension	0	0	0	0	0	0	0	0	0	0
Written Comprehension	0	0	0	0	0	0	0	0	0	0
Oral Expression	0	0	0	0	0	0	0	0	0	0
Written Expression	0	0	0	0	0	0	0	0	0	0
Fluency of Ideas	0	0	0	0	0	11	10	1	0	0
Originality	0	0	0	0	0	3	8	0	0	0
Problem Sensitivity	0	0	0	0	0	0	1	0	1	0
Deductive Reasoning	0	0	0	0	0	0	2	0	0	0
Inductive Reasoning	0	0	0	0	0	1	1	0	0	0
Information Ordering	0	0	0	0	0	1	1	0	1	0
Category Flexibility	0	0	0	0	0	2	10	0	1	0
Mathematical Reasoning	0	0	0	0	0	7	3	1	1	0
Number Facility	0	0	0	0	0	15	10	9	3	1
Memorization	0	0	0	0	0	18	18	1	5	3
Speed of Closure	0	0	0	0	0	32	29	5	10	4
Flexibility of Closure	2	0	0	0	0	29	35	22	5	1
Perceptual Speed	0	0	0	0	0	15	15	9	3	0
Spatial Orientation	1	0	0	0	0	9	6	1	1	1
Visualization	0	0	0	0	0	19	26	16	6	4
Selective Attention	0	0	0	0	0	2	6	0	2	0
Time Sharing	0	0	0	0	0	6	7	0	1	0
Arm-Hand Steadiness	0	0	0	0	0	2	3	0	0	0
Manual Dexterity	0	0	0	0	0	8	9	2	4	0
Finger Dexterity	0	0	0	0	0	20	9	0	3	0
Control Precision	0	0	0	0	0	5	8	4	5	1
Multilimb Coordination	0	0	0	0	0	8	5	1	5	1
Response Orientation	0	0	0	0	0	8	11	4	3	1
Rate Control	0	0	0	0	0	2	6	0	3	1
Reaction Time	0	0	0	0	0	19	19	4	4	3
Wrist-Finger Speed	0	0	0	0	0	33	7	1	2	2
Speed of Limb Movement	1	0	0	0	0	4	13	2	1	1
Static Strength	0	0	0	0	0	6	12	4	0	0
Explosive Strength	1	0	0	0	0	3	6	0	1	3
Dynamic Strength	0	0	0	0	0	7	9	2	2	2
Trunk Strength	0	0	0	0	0	1	0	0	0	0
Stamina	0	0	0	0	0	3	3	1	1	0
Extent Flexibility	0	0	0	0	0	13	14	0	5	4
Dynamic Flexibility	0	0	0	0	0	5	0	0	0	0
Gross Body Coordination	0	0	0	0	0	0	2	1	1	0
Gross Body Equilibrium	0	0	0	0	0	0	5	1	1	1

Element Name	Frequency $SE_M$ Importance > 51					Frequency $SE_M$ Level > 51				
	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Near Vision	0	0	0	0	0	0	0	2	0	0
Far Vision	0	0	0	0	0	14	20	3	9	0
Visual Color Discrimination	0	0	0	0	0	16	18	7	4	1
Night Vision	0	0	0	0	0	4	1	0	0	3
Peripheral Vision	0	0	0	0	0	2	3	0	2	1
Depth Perception	0	0	0	0	0	0	8	2	1	0
Glare Sensitivity	0	0	0	0	0	2	9	1	0	0
Hearing Sensitivity	0	0	0	0	0	6	10	5	4	1
Auditory Attention	0	0	0	0	0	9	23	0	2	6
Sound Localization	1	0	0	0	0	9	8	4	3	2
Speech Recognition	0	0	0	0	0	8	3	4	2	1
Speech Clarity	0	0	0	0	0	2	6	0	1	0
TOTAL	0% (6/6552)	0% (0/5668)	0% (0/5200)	0% (0/4732)	0% (0/5200)	5.91% (387/6552)	7.82% (443/5668)	2.31% (120/5200)	2.30% (109/4732)	.94% (49/5200)

While the frequency of flagging an ability level rating was higher than the importance rating, it should be noted that the total number of level flags reflected only .94% of the 5,200 total ratings. In addition, this value continues to reflect a decrease in the percentage of ability level ratings receiving flags across the first five cycles. These findings suggest there remains a high level of agreement among the analysts and that constructs that earlier appeared problematic may not require additional training. Moreover, the heterogeneity of occupations that was found in Cycle 5 continued in Cycle 6; Cycle 6 was comprised of occupations that represented multiple O\*NET job families. This indicates that the reduced number of flags for level should not be attributed to the similarity among the occupations. However, it continues to be advisable to monitor the elements that were previously problematic to determine if additional training is warranted.

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

### ***Cycle 6 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.<sup>1</sup> The columns labeled “Median of  $SD_s$ ” show the median of the  $SD_s$  associated with

<sup>1</sup> While the mean is not a measure of agreement, it can affect the potential range of the  $SD$  and  $SE_M$ .



each mean importance and level rating across the 52 abilities for each occupation. Finally, the columns labeled “Median of  $SE_{MS}$ ” show the median of the  $SE_{MS}$  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 .71 and a median  $SE_M$  of .25. These results for importance and level remained the same as those found in Cycle 5. As such, they also represent a decrease in interrater agreement from that found in Cycles 1-4. The continuation of this small decrease in interrater agreement may have been caused by the introduction of new analysts in Cycle 5. However, while the values are generally greater 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 6 Interrater Reliability: Across Constructs Within Occupations***

To examine the interrater reliability of the Cycle 5 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 100 Cycle 6 occupations. Specifically, the mean  $ICC$  for importance ratings for the abilities across the occupations was .94 ( $SD = .03$ ). The mean  $ICC$  for the level ratings was .94 ( $SD = .04$ ). The reliability for both the importance and level ratings exceeded the target coefficient value of .80. Interrater reliability did not vary greatly across occupations and the mean coefficient for importance ratings was identical to the mean coefficient for level ratings. Results also indicate that 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.

### ***Cycle 5 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 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)  $ICC(3, k)$  must be 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 type of reliability was first used to evaluate the raters after combining results of Cycle 1 and Cycle 2 data collection because it requires a reasonable number of occupations. With the completion of Cycle 5, there were 480 occupations included in the reliability analysis. The current reliability analysis was conducted on all 580 occupations from Cycles 1 through 6 and results are presented in Table 3. 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. The lowest ICC(C,8) reliabilities were found for Selective Attention and Time Sharing; none of the reliabilities for these two constructs had reliabilities over .66 on either level or importance. These reliabilities may be due to low variation in the importance or the required level of these abilities across jobs or disagreement among raters.

In addition, Near Vision, Speech Recognition, Memorization, Information Ordering, Category Flexibility, and Flexibility of Closure all had ICC(C,8) reliabilities for Importance that were less than .70 (ICC(C,8) reliabilities for level were all .70 or greater). These differences in reliabilities for importance and level likely reflect high agreement but lack of variability in the ratings of these constructs across occupations on importance and high agreement and high variation in the ratings of these constructs across jobs on level.

However, comparisons with interrater reliabilities obtained from Cycle 1, 2, 3, 4, and 5 data indicate that for some elements ICC(C,8) reliability improved with the addition of Cycle 6 data. For example, the construct Speech Recognition demonstrated increases in ICC(C,8) reliability of .04 and .05 for importance and level, respectively. Increases in the size of reliability coefficients are limited because of the relatively large coefficients already obtained on the majority of constructs.

**Table 3. Interrater Reliabilities and Standard Errors of Measurement Across Cycle 1, 2, 3, 4, 5, and 6 Occupations**

Ability	Cycles 1 through 6 (N = 580)					
	Importance			Level		
	ICC(C,1)	ICC(C,8)	s <sub>E</sub>	ICC(C,1)	ICC(C,8)	s <sub>E</sub>
1 Oral Comprehension	0.36	0.82	0.18	0.47	0.88	0.21
2 Written Comprehension	0.49	0.88	0.19	0.61	0.93	0.22
3 Oral Expression	0.45	0.87	0.18	0.51	0.89	0.20
4 Written Expression	0.48	0.88	0.20	0.63	0.93	0.24
5 Fluency of Ideas	0.40	0.84	0.22	0.45	0.87	0.30
6 Originality	0.49	0.89	0.20	0.54	0.90	0.27
7 Problem Sensitivity	0.34	0.81	0.19	0.49	0.89	0.24
8 Deductive Reasoning	0.32	0.79	0.19	0.50	0.89	0.23
9 Inductive Reasoning	0.40	0.84	0.19	0.54	0.90	0.24
10 Information Ordering	0.20	0.67	0.21	0.32	0.79	0.24
11 Category Flexibility	0.20	0.67	0.21	0.28	0.76	0.27
12 Mathematical Reasoning	0.48	0.88	0.23	0.57	0.91	0.31
13 Number Facility	0.39	0.84	0.24	0.48	0.88	0.35

Continued on next page

**Cycles 1 through 5 (N = 480)**

Ability	Importance			Level		
	ICC(C,1)	ICC(C,8)	$s_E$	ICC(C,1)	ICC(C,8)	$s_E$
14 Memorization	0.18	0.64	0.24	0.23	0.70	0.36
15 Speed of Closure	0.22	0.70	0.26	0.27	0.74	0.39
16 Flexibility of Closure	0.22	0.69	0.27	0.28	0.75	0.37
17 Perceptual Speed	0.26	0.74	0.26	0.25	0.73	0.34
18 Spatial Orientation	0.59	0.92	0.18	0.58	0.92	0.26
19 Visualization	0.40	0.84	0.25	0.43	0.86	0.37
20 Selective Attention	0.16	0.60	0.21	0.17	0.63	0.26
21 Time Sharing	0.18	0.64	0.23	0.20	0.66	0.29
22 Arm-Hand Steadiness	0.71	0.95	0.20	0.70	0.95	0.27
23 Manual Dexterity	0.70	0.95	0.20	0.65	0.94	0.31
24 Finger Dexterity	0.43	0.86	0.24	0.43	0.86	0.32
25 Control Precision	0.71	0.95	0.20	0.68	0.94	0.31
26 Multilimb Coordination	0.70	0.95	0.20	0.69	0.95	0.28
27 Response Orientation	0.65	0.94	0.18	0.66	0.94	0.29
28 Rate Control	0.70	0.95	0.17	0.68	0.95	0.25
29 Reaction Time	0.71	0.95	0.18	0.70	0.95	0.30
30 Wrist-Finger Speed	0.43	0.86	0.21	0.41	0.85	0.35
31 Speed of Limb Movement	0.59	0.92	0.17	0.59	0.92	0.26
32 Static Strength	0.71	0.95	0.19	0.75	0.96	0.27
33 Explosive Strength	0.42	0.85	0.12	0.43	0.86	0.20
34 Dynamic Strength	0.62	0.93	0.18	0.65	0.94	0.26
35 Trunk Strength	0.63	0.93	0.20	0.65	0.94	0.27
36 Stamina	0.67	0.94	0.18	0.65	0.94	0.26
37 Extent Flexibility	0.74	0.96	0.17	0.74	0.96	0.29
38 Dynamic Flexibility	0.26	0.74	0.10	0.29	0.77	0.16
39 Gross Body Coordination	0.65	0.94	0.18	0.67	0.94	0.25
40 Gross Body Equilibrium	0.67	0.94	0.15	0.63	0.93	0.24
41 Near Vision	0.15	0.59	0.19	0.33	0.80	0.25
42 Far Vision	0.39	0.83	0.23	0.33	0.80	0.35
43 Visual Color Discrimination	0.41	0.85	0.24	0.42	0.85	0.35
44 Night Vision	0.59	0.92	0.14	0.56	0.91	0.22
45 Peripheral Vision	0.63	0.93	0.14	0.60	0.92	0.21
46 Depth Perception	0.62	0.93	0.20	0.61	0.93	0.28
47 Glare Sensitivity	0.68	0.94	0.14	0.69	0.95	0.22
48 Hearing Sensitivity	0.47	0.88	0.23	0.45	0.87	0.33
49 Auditory Attention	0.35	0.81	0.22	0.39	0.83	0.34
50 Sound Localization	0.55	0.91	0.15	0.55	0.91	0.24
51 Speech Recognition	0.17	0.62	0.22	0.22	0.70	0.30
52 Speech Clarity	0.34	0.81	0.20	0.38	0.83	0.29

*Note.* These ICCs indicate how consistently raters rated occupations on a given ability.

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

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. Median ICC(C,8) reliabilities for importance and level were .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 6 analyst ratings were as follows:

- The not-relevance and suppression criteria did not generate any results reflecting poorly on the overall quality of the Cycle 6 ratings.
- The decline in the percentage of abilities flagged for level ratings due to a  $SE_M$  greater than .51 indicates that elements once considered problematic in earlier cycles may not require additional training.
- While interrater agreement was higher for importance than for level ratings, overall results indicate that the ratings made by the analysts were consistent for both scales across occupations. Cycle 6 analyst agreement results were consistent with those found in Cycle 5.
- 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 6 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.

Given these results, it appears as though 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. Agreement was high and there is clear evidence regarding the quality of the data.

## References

- Byrum, C.N., & Tsacoumis, S. (2006). *O\*NET analyst occupational abilities ratings: Analysis Cycle 5 results*. Alexandria, VA: Human Resources Research Organization.
- Byrum, C.N., & Tsacoumis, S. (2005). *O\*NET analyst occupational abilities ratings: Analysis Cycle 4 results*. Alexandria, VA: Human Resources Research Organization.
- 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.
- Noble, C.L., Sager, C., Tsacoumis, S., Updegraff, J. & Donsbach, J. (2003). *O\*NET analyst occupational abilities ratings: Wave 1 results*. Alexandria, VA: Human Resources Research Organization.
- Noble, C.L., & Tsacoumis, S. (2004). *O\*NET analyst occupational abilities ratings: Analysis Cycle 2 results*. Alexandria, VA: Human Resources Research Organization.
- Noble, C.L., & Tsacoumis, S. (2005). *O\*NET analyst occupational abilities ratings: Analysis Cycle 3 results*. 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