FR-10-28

*O*NET® Analyst Occupational Abilities Ratings: Analysis Cycle 11 Results*

Suzanne Tsacoumis
Shannon Willison

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

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

November 2010

HumRRO
Human Resources Research Organization
66 Canal Center Plaza, Suite 700 • Alexandria, Virginia 22314
www.humrro.org • Phone: (703) 549-3611 - Fax: (703) 519-9661

O*NET ANALYST OCCUPATIONAL ABILITIES RATINGS: ANALYSIS CYCLE 11 RESULTS

Table of Contents

List of Tables

O*NET ANALYST OCCUPATIONAL ABILITIES RATINGS: ANALYSIS CYCLE 11 RESULTS

Introduction

The Occupational Information Network (O*NET) is a comprehensive system developed by the U.S. Department of Labor that provides information for over 900 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 from three primary sources: incumbents, occupational experts, and occupational 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 occupational 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). Granted, it is imperative that the occupational analysts have detailed occupation information in order to rate the ability and skill constructs. It has also been suggested that some incumbents deliberately inflate their ratings to influence policy decisions regarding, for example, compensation and training (Harvey, 1991; Morgeson, Delaney-Klinger, Mayfield, Ferrara, & Campion, 2004). Given these considerations, occupational analysts as opposed to incumbents provide the ability and skill information in the O*NET database.

This report focuses on the ability results only. 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.

To facilitate the ability rating process, occupational analysts are provided relevant occupational information. Trained occupational analysts are responsible for rating the importance and level of the 52 abilities for each of the O*NET occupations. More specifically, eight trained occupational 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 eleventh analysis cycle of 120 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/datapublication.html. The O*NET-SOC Codes and Titles included in O*NET Analysis 11 are presented in Appendix A.

Evaluation of Cycle 11 Analyst Ratings

As mentioned above, occupational analysts provided ratings on importance and level of the 52 abilities for each of the 120 occupations in Cycle 11. The mean, standard deviation, and SE_M of the importance and level ratings were computed. These results are shown in Appendix B.

Four sets of analyses were performed to evaluate the ratings that occupational 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 11 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 occupational 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 11 data, the level ratings for the Spatial Orientation were considered not relevant for Fraud, Examiners, Investigators, and Analysts (13-2099.04) as well as Pharmacy Aides (31-9095.00) because Spatial Orientation was not considered important for the performance of these two occupations. In this cycle, there were 1,293 not relevant flags (see Table 1). To facilitate interpretation of these results, it should be

noted that there are 6,240 sets of ratings (120 occupations x 52 abilities) in the current cycle. Given this, 20.72% (1,293/6,240) of the ability ratings were flagged as not relevant.

Table 1 shows the number of not-relevant flags for ability level. The abilities with the most flags in Cycle 11 include Dynamic Flexibility, Explosive Strength, Night Vision, Peripheral Vision, Glare Sensitivity, and Sound Localization; 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. There were no instances in Cycle 11 where the mean importance rating was flagged for insufficient agreement. In fact, no importance ratings received flags for an SE_M greater than .51 since Cycle 3. The results of the suppression criteria for level for Cycles 1-11 are presented in Table 2. There were 60 insufficient agreement flags for level ratings. It should be noted that although the number of flags indicating insufficient agreement with level ratings in Cycle 11 is slightly higher than Cycle 10, the overall percentage is consistent with previous cycles.

In Cycle 11, the abilities that were flagged the most for the level criteria were: Number Facility (n = 13), Control Precision (n = 6), and Manual Dexterity (n = 5). Although still fairly low, the flag count for each of these constructs increased in Cycle 11. However, it should be noted that Control Precision was one of the most flagged abilities from Cycle 10 as well. A close look at the results revealed that the 13 occupations were in the science field (e.g., Biomedical Engineer, Robotics Engineer, Preventive Medicine Physicians) and in all cases Number Facility was rated as quite important. Since the definition of Numerical Facility includes a speeded component (i.e., "the ability to add, subtract, multiply, or divide *quickly* and correctly"), it is possible that the flags for insufficient agreement can be attributed to the differing perceptions of the extent to which people in these occupations need to perform basic math quickly.

Although some abilities showed an increase in the number of level flags, most abilities received the same number of flags in Cycle 11 or decreased slightly. One of the two abilities identified as receiving the most flags in Cycle 10 (Sound Localization) decreased from four flags to one flag in Cycle 11. Overall, eight abilities showed a decrease in the number of flags and 28 abilities remained the same from Cycle 10 to 11.

With two small exceptions, a decreasing trend exists across cycles in the percentage of flagged ability level ratings. The increases in Cycle 7 and Cycle 10 were relatively small, and in the latter case, the percentage of flags was still below 1%. These findings suggest there remains a high level of agreement among the occupational analysts. 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 (*N* = 54)	Cycle 2 (*N* = 126)	Cycle 3 (*N* = 109)	Cycle 4 (*N* = 100)	Cycle 5 (*N* = 91)	Cycle 6 (*N* = 100)	Cycle 7 (*N* = 101)	Cycle 8 (*N* = 100)	Cycle 9 (*N* = 31)	Cycle 10 (*N* = 192)	Cycle 11 (*N* = 120)
Oral Comprehension	0	0	0	0	0	0	0	0	0	0	0
Written Comprehension	0	0	0	0	0	0	0	0	0	0	0
Oral Expression	0	0	0	0	0	0	0	0	0	0	0
Written Expression	0	0	0	0	0	0	0	0	0	0	0
Fluency of Ideas	0	2	0	0	0	0	0	1	0	0	0
Originality	0	7	2	0	0	1	0	1	0	0	0
Problem Sensitivity	0	0	0	0	0	0	0	0	0	0	0
Deductive Reasoning	0	0	0	0	0	0	0	0	0	0	0
Inductive Reasoning	0	0	0	0	0	0	0	0	0	0	0
Information Ordering	0	0	0	0	0	0	0	0	0	0	0
Category Flexibility	0	0	0	0	0	0	0	0	0	0	0
Mathematical Reasoning	0	6	4	1	3	4	4	6	0	1	0
Number Facility	3	5	0	1	1	3	2	4	0	1	0
Memorization	0	1	0	0	0	5	3	1	1	0	0
Speed of Closure	0	2	3	0	0	0	1	3	0	0	0
Flexibility of Closure	0	2	0	0	0	0	0	0	0	0	0
Perceptual Speed	0	1	1	0	0	0	0	3	0	0	0
Spatial Orientation	36	48	66	81	54	48	35	33	14	130	75
Visualization	0	6	3	0	2	0	0	1	0	0	0
Selective Attention	0	0	0	0	0	0	0	0	0	0	0
Time Sharing	0	0	0	0	0	0	0	2	0	0	0
Arm-Hand Steadiness	9	14	11	49	15	14	7	6	1	42	18
Manual Dexterity	9	19	9	54	16	16	7	6	1	48	27
Finger Dexterity	0	6	3	0	1	0	0	1	0	0	0
Control Precision	6	19	13	48	16	12	7	6	1	33	24
Multilimb Coordination	13	31	23	50	25	15	10	7	7	64	23
Response Orientation	30	72	50	66	39	28	16	14	9	91	44
Rate Control	35	88	57	73	43	29	18	16	9	98	53
Reaction Time	27	65	40	66	39	23	13	14	9	91	40
Wrist-Finger Speed	26	50	54	76	50	32	21	17	7	106	49
Speed of Limb Movement	28	57	49	65	47	34	20	22	12	107	67
Static Strength	21	38	33	56	36	23	15	11	8	84	33
Explosive Strength	44	104	90	93	85	93	93	80	27	160	101
Dynamic Strength	28	61	46	65	42	28	15	17	13	97	52
Trunk Strength	8	16	23	29	30	21	6	4	7	77	20
Stamina	21	42	38	58	38	25	14	12	13	90	43
Extent Flexibility	22	47	36	64	40	24	15	13	13	95	50
Dynamic Flexibility	52	104	102	98	90	99	99	92	28	182	111
Gross Body Coordination	21	46	36	58	38	25	14	13	13	90	43
Gross Body Equilibrium	27	67	53	61	44	26	14	13	13	91	47
Near Vision	0	0	0	0	0	0	0	0	0	0	0
Far Vision	0	4	3	0	0	0	0	0	0	0	0
Visual Color Discrimination	2	18	7	2	1	0	0	2	0	0	0
Night Vision	44	99	83	83	58	53	40	35	14	146	92
Peripheral Vision	44	85	79	82	55	54	41	32	14	145	91
Depth Perception	11	21	24	35	12	13	6	4	0	15	13
Glare Sensitivity	41	93	68	84	48	45	30	28	11	140	89
Hearing Sensitivity	2	39	32	16	3	0	0	3	0	0	0
Auditory Attention	2	10	4	1	2	0	0	1	0	0	1
Sound Localization	44	95	83	84	54	52	39	32	13	138	87
Speech Recognition	0	0	0	0	0	0	0	0	0	0	0
Speech Clarity	0	0	0	0	0	0	0	0	0	0	0
Total Flags out of all possible ability ratings	23.36% (656/2808)	22.74% (1490/6552)	21.67% (1228/5668)	30.75% (1599/5200)	21.70% (1027/4732)	16.25% (845/5200)	11.52% (605/5252)	10.69% (556/5200)	16.00% (258/1612)	23.66% (2362/9984)	20.72% (1293/6240)

Table 2. Level Flags Due to Large SE_M

Element Name	Cycle1&2 (*N* = 180)	Cycle 3 (*N* = 109)	Cycle 4 (*N* = 100)	Cycle 5 (*N* = 91)	Cycle 6 (*N* = 100)	Cycle 7 (*N* = 101)	Cycle 8 (*N* = 100)	Cycle 9 (*N* = 31)	Cycle 10 (*N* = 192)	Cycle 11 (*N* = 120)
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	1	1	0	0	0
Fluency of Ideas	7.5	10	1	0	0	1	0	0	0	0
Originality	2	8	0	0	0	1	0	0	0	0
Problem Sensitivity	0	1	0	1	0	0	0	0	0	0
Deductive Reasoning	0	2	0	0	0	0	0	0	0	0
Inductive Reasoning	0.5	1	0	0	0	0	0	0	0	0
Information Ordering	0.5	1	0	1	0	0	0	0	0	0
Category Flexibility	1	10	0	1	0	0	0	0	0	0
Mathematical Reasoning	4	3	1	1	0	0	0	0	0	0
Number Facility	8	10	9	3	1	2	0	0	0	13
Memorization	10.5	18	1	5	3	5	2	0	1	0
Speed of Closure	18	29	5	10	4	4	2	1	1	3
Flexibility of Closure	21.5	35	22	5	1	1	1	0	1	2
Perceptual Speed	13.5	15	9	3	0	1	0	0	1	0
Spatial Orientation	5	6	1	1	1	4	3	2	0	2
Visualization	16	26	16	6	4	1	0	1	0	0
Selective Attention	1	6	0	2	0	1	0	0	0	0
Time Sharing	3	7	0	1	0	0	0	0	0	0
Arm-Hand Steadiness	2.5	3	0	0	0	0	0	0	1	0
Manual Dexterity	7	9	2	4	0	0	0	2	2	5
Finger Dexterity	10	9	0	3	0	0	0	1	1	2
Control Precision	4.5	8	4	5	1	1	1	4	4	6
Multilimb Coordination	4	5	1	5	1	0	0	0	1	1
Response Orientation	7	11	4	3	1	5	3	0	1	3
Rate Control	2.5	6	0	3	1	1	2	0	1	0
Reaction Time	12.5	19	4	4	3	13	2	0	1	3
Wrist-Finger Speed	27	7	1	2	2	15	0	1	2	4
Speed of Limb Movement	2.5	13	2	1	1	7	4	0	0	0
Static Strength	5	12	4	0	0	3	4	0	0	0
Explosive Strength	3	6	0	1	3	2	2	0	2	2
Dynamic Strength	5.5	9	2	2	2	2	3	0	0	0
Trunk Strength	1.5	0	0	0	0	0	0	0	0	0
Stamina	2.5	3	1	1	0	1	0	0	0	1
Extent Flexibility	7	14	0	5	4	10	4	0	0	1
Dynamic Flexibility	4	5	0	0	0	1	2	0	1	0
Gross Body Coordination	0	2	1	1	0	1	1	0	0	0
Gross Body Equilibrium	2	5	1	1	1	1	1	0	1	0
Near Vision	0	0	2	0	0	0	0	0	0	0
Far Vision	15	20	3	9	0	1	0	1	2	4
Visual Color Discrimination	10.5	18	7	4	1	2	8	1	1	2
Night Vision	3.5	1	0	0	3	3	4	0	1	1
Peripheral Vision	1.5	3	0	2	1	6	1	0	0	0
Depth Perception	0.5	8	2	1	0	0	0	0	0	0
Glare Sensitivity	2	9	1	0	0	2	6	1	2	1
Hearing Sensitivity	4.5	10	5	4	1	2	2	0	0	1
Auditory Attention	5	23	0	2	6	10	11	2	1	2
Sound Localization	5	8	4	3	2	5	3	1	4	1
Speech Recognition	4	3	4	2	1	3	0	0	0	0
Speech Clarity	1	6	0	1	0	0	0	0	0	0
Total Flags out of all possible ability ratings	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)	1.16% (18/1612)	0.33% (33/9984)	0.96% (60/6240)

Cycle 11 Interrater Agreement

Interrater agreement was computed to examine the level of absolute agreement among the occupational 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 agreement, we calculated the standard deviation (*SD*) of ratings across occupational 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 *Ms*" show the mean of the occupational analyst mean importance and level ratings across the 52 abilities for each occupation.[1] The columns labeled "Median of *SDs*" show the median of the *SDs* associated with 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 .64 and a median SE_M of .23. These results for importance and level reveal that raters agreed similarly in this cycle compared to previous cycles. Overall, while the values are generally greater (indicating less agreement) for level than they are for importance, the results indicate that the ratings made by the occupational analysts were consistent for both scales.

Cycle 11 Interrater Reliability: Across Constructs Within Occupations

To examine the interrater reliability of the Cycle 11 ratings we calculated the interclass correlations (ICC [3, *k*]; Shrout & Fleiss, 1979) among the occupational 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 in multiple contexts, including O*NET (e.g., Clement, Chauvot, Philipp, & Ambrose, 2003; McCloy, et al., 1999; Rase & Tognetti-Stuff, 1983).

The results of these analyses are presented in Appendix D. The data revealed high levels of interrater reliability across the 120 Cycle 11 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 .96 and .97, 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 very good level of agreement in the occupational 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 11 Interrater Reliability: Across Occupations Within Constructs

Another effective way to evaluate the reliability of the occupational 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 multiple contexts, including O*NET (e.g., Clement, Chauvot, Philipp, & Ambrose, 2003; McCloy, et al., 1999; Rase & Tognetti-Stuff, 1983).

This reliability analysis was conducted for abilities on all occupations in Cycles 1 through 11 and results are presented in Table 3. Note that a number of occupations were rated in two different cycles (e.g., Cycles 5 and 7); therefore, the reliability analyses are based on 1,124 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 achieved the target level. The majority of the abilities had high ICC(C,8) reliabilities for both importance and level. In fact, there were 20 ability levels with reliabilities greater than .90 for both importance and level (e.g., Spatial Orientation). However, there are some low reliabilities to note.

Table 3. Interrater Reliabilities and Standard Errors of Measurement for Abilities Across Occupations in Cycles 1 through 11

		Cycles 1 through 11 (*N* = 1,124)					
		Importance			Level		
	Ability	ICC(C,1)	ICC(C,8)	s_E	ICC(C,1)	ICC(C,8)	s_E
1	Oral Comprehension	0.38	0.83	0.18	0.50	0.89	0.20
2	Written Comprehension	0.49	0.89	0.19	0.60	0.92	0.21
3	Oral Expression	0.46	0.87	0.18	0.53	0.90	0.20
4	Written Expression	0.49	0.88	0.20	0.63	0.93	0.23
5	Fluency of Ideas	0.40	0.84	0.21	0.47	0.88	0.28
6	Originality	0.45	0.87	0.20	0.52	0.90	0.27
7	Problem Sensitivity	0.35	0.81	0.18	0.49	0.89	0.23
8	Deductive Reasoning	0.35	0.81	0.18	0.50	0.89	0.23
9	Inductive Reasoning	0.41	0.85	0.18	0.52	0.90	0.23
10	Information Ordering	0.21	0.68	0.19	0.33	0.80	0.23
11	Category Flexibility	0.20	0.67	0.19	0.30	0.78	0.25
12	Mathematical Reasoning	0.48	0.88	0.22	0.57	0.91	0.30
13	Number Facility	0.38	0.83	0.23	0.49	0.88	0.32
					Table Continued on Next Page		
14	Memorization	0.18	0.63	0.23	0.24	0.71	0.34
15	Speed of Closure	0.21	0.68	0.25	0.28	0.75	0.35

		Cycles 1 through 11 (N = 1,124)					
		Importance			Level		
	Ability	ICC(C,1)	ICC(C,8)	s_E	ICC(C,1)	ICC(C,8)	s_E
16	Flexibility of Closure	0.22	0.69	0.25	0.29	0.76	0.32
17	Perceptual Speed	0.26	0.74	0.24	0.27	0.75	0.30
18	Spatial Orientation	0.55	0.91	0.19	0.56	0.91	0.28
19	Visualization	0.38	0.83	0.23	0.43	0.86	0.32
20	Selective Attention	0.12	0.51	0.21	0.15	0.58	0.25
21	Time Sharing	0.17	0.62	0.22	0.20	0.66	0.27
22	Arm-Hand Steadiness	0.73	0.96	0.19	0.72	0.95	0.26
23	Manual Dexterity	0.71	0.95	0.20	0.66	0.94	0.30
24	Finger Dexterity	0.42	0.86	0.23	0.44	0.86	0.29
25	Control Precision	0.72	0.95	0.19	0.69	0.95	0.29
26	Multilimb Coordination	0.71	0.95	0.20	0.71	0.95	0.27
27	Response Orientation	0.63	0.93	0.19	0.67	0.94	0.29
28	Rate Control	0.70	0.95	0.18	0.69	0.95	0.26
29	Reaction Time	0.72	0.95	0.19	0.71	0.95	0.30
30	Wrist-Finger Speed	0.45	0.87	0.21	0.45	0.87	0.33
31	Speed of Limb Movement	0.61	0.93	0.18	0.62	0.93	0.26
32	Static Strength	0.70	0.95	0.19	0.75	0.96	0.27
33	Explosive Strength	0.40	0.84	0.13	0.41	0.85	0.21
34	Dynamic Strength	0.61	0.93	0.18	0.65	0.94	0.26
35	Trunk Strength	0.64	0.94	0.20	0.68	0.94	0.26
36	Stamina	0.68	0.94	0.18	0.67	0.94	0.25
37	Extent Flexibility	0.74	0.96	0.17	0.75	0.96	0.28
38	Dynamic Flexibility	0.21	0.68	0.10	0.23	0.71	0.17
39	Gross Body Coordination	0.64	0.94	0.18	0.68	0.95	0.24
40	Gross Body Equilibrium	0.63	0.93	0.17	0.62	0.93	0.25
41	Near Vision	0.14	0.57	0.19	0.32	0.79	0.23
42	Far Vision	0.32	0.79	0.23	0.31	0.78	0.32
43	Visual Color Discrimination	0.37	0.82	0.23	0.41	0.85	0.33
44	Night Vision	0.53	0.90	0.16	0.55	0.91	0.23
45	Peripheral Vision	0.60	0.92	0.15	0.60	0.92	0.22
46	Depth Perception	0.58	0.92	0.20	0.58	0.92	0.29
47	Glare Sensitivity	0.65	0.94	0.15	0.69	0.95	0.23
48	Hearing Sensitivity	0.40	0.84	0.24	0.41	0.85	0.33
49	Auditory Attention	0.33	0.80	0.22	0.36	0.82	0.33
50	Sound Localization	0.50	0.89	0.16	0.52	0.90	0.25
51	Speech Recognition	0.19	0.66	0.21	0.25	0.73	0.28
52	Speech Clarity	0.34	0.81	0.19	0.37	0.83	0.27

Note. These ICCs indicate how consistently raters rated (rank ordered) occupations on a given ability. sE = Standard error of measurment = Observed score variance times the square root of one minus ICC(C,8).

The lowest importance ICC(C,8) reliabilities were found for Selective Attention and Near Vision, both of which were under .60. The constructs with the lowest level reliabilities was Selective Attention and Time Sharing (.58 and .66, respectively). The results for these constructs are consistent with those in the previous cycles. Close examination revealed low variation in the ratings, which is a contributing factor to low reliability and therefore is a plausible explanation for the results, even though these are not the only abilities with low variation in the importance or required level across jobs.

Eight additional abilities had ICC(C,8) reliabilities for importance that were less than .70: Time Sharing, Memorization, Speech Recognition, Category Flexibility, Dynamic Flexibility, Flexibility of closure, Information Ordering, and Speed of Closure. ICC(C,8) reliabilities for the level ratings of these abilities (except for Time Sharing) were greater than .70. The observed 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 occupations on level.

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 very 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 11 analyst ratings were as follows:

- About 20% of the ability ratings were flagged because the construct was considered not important for performance. This is a decrease from Cycle 10, but it is still higher than other recent cycles. The constructs that were flagged were very similar to those flagged in previous cycles and conceptually it is understandable that these constructs may be considered not-relevant for the given occupations.
- Less than 1 percent of the level ratings were flagged based on a SE_M greater than .51.
- There was strong interrater agreement for this cycle as evidenced by the overall low medians of $SE_M s$.
- All within-occupation ICC reliabilities were well above the target value of .80. These high levels of interrater reliability indicate that the occupational analysts rank ordered the abilities within each occupation similarly on both importance and level.
- Index interrater reliability calculated at the end of Cycle 11 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.

Given these results, it appears as though the analysts were well trained and understand the abilities and associated definitions. Agreement was high and there is clear evidence regarding the high quality of the data. Nevertheless, project staff will continue to review the constructs and process with returning analysts prior to each new cycle and as needed, throughout a cycle. Also, staff will thoroughly train new analysts, if any.

References

Clement, L., Chauvot, J., Philipp, R., & Ambrose, R. (2003). A method for developing rubrics for research purposes. In N. A. Pateman, B. J. Dougherty, & J. T. Zilliox (Eds.), *Proceedings of the 2003 joint meeting of PME and PMENA* (Vol. 2, pp. 221–227). Honolulu: CRDG, College of Education, University of Hawaii.

Donsbach, J., Tsacoumis, S., Sager, C., & Updegraff, J. (2003). *O*NET analyst occupational abilities ratings: Procedures* (FR-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.

Rase, C.W. & Tognetti-Stuff, R. K. (1983). Reliability of the Auditing Process at the University of Montana's Physical Therapy Department. *Physical Therapy, 64*(7), 1088-1090.

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.

U.S. Department of Labor (1991). Dictionary of occupational titles. (Rev. 4th ed.). Washington, DC: U.S. Government Printing Office.