HumRRO
HUMAN RESOURCES RESEARCH ORGANIZATION

2022 No. 084

O*NET® Analyst Ratings of Occupational Skills: Analysis Cycle 23 Results

Final Report

Prepared for:	National Center for O*NET Development 313 Chapanoke Road, Suite 130 Raleigh, NC 27603	**Prepared under:**	Subcontract Number (through RTI International): 1-312-0207142-41224L
Authors:	Timothy C. Burgoyne Matthew C. Reeder	**Date:**	December 9, 2022

Headquarters: 66 Canal Center Plaza, Suite 700, Alexandria, VA 22314 | Phone: 703.549.3611 | humrro.org

O*NET Analyst Ratings of Occupational Skills: Analysis Cycle 23 Results

Table of Contents

List of Tables

O*NET Analyst Ratings of Occupational Skills: Analysis Cycle 23 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. 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 (GWAs), knowledge, education and training, work styles, and work context areas. Importance and level information regarding the abilities and skills associated with these occupations is 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 skill 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 (Morgeson et al., 2004). Skill ratings may be particularly vulnerable to such effects given that they are more abstract and thus more difficult to verify than more observable descriptors such as job tasks (Morgeson & Campion, 1997; Morgeson et al., 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 results pertaining to the skill ratings only. Skills reflect proficiencies or competencies that are developed through training or experience (Peterson et al., 2001). The 35 O*NET skills cover performance applicable to a broad range of jobs in the world's economy and are grouped into seven categories within the O*NET content model: content, process, social, complex problem solving, technical, systems, and resource management.

To facilitate the skill rating process, occupational analysts are provided relevant occupational information. Trained occupational analysts are responsible for rating the importance and level of the 35 skills 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 skill ratings, see *O*NET Analyst Ratings of Occupational Skills: Procedures Update* (Fleisher & Tsacoumis, 2018).

To ensure a controlled data collection and management process, occupational data are being collected in groups or "analysis cycles." This report describes the results from the data collection process for the 23rd analysis cycle of 80 occupations. Reports describing each of the previous cycles are available at https://www.onetcenter.org/research.html?c=KSA. Results for subsequent cycles will be reported in separate reports. For a description of the O*NET Data Collection Publication Schedule see http://www.onetcenter.org/dataPublication.html. Appendix A includes a listing of the IDI codes and Occupational Titles addressed in Cycle 23.

Evaluation of Cycle 23 Analyst Ratings

As mentioned above, occupational analysts provided ratings on the importance and level of the 35 skills for each of the 80 occupations in Cycle 23. The mean, standard deviation, and standard error of the mean (SE_M) of the importance and level ratings were computed. These results are shown in Appendix B.

We performed four sets of analyses 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 skill level rating is not relevant for a specific occupation. Thus, a set of recommended suppression criteria was established that flagged: (a) a skill level rating as not relevant to an occupation because of low importance ratings, (b) a skill with too little agreement in importance ratings across raters for a particular occupation, and (c) a skill 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 as indicated by low reliability estimates may suggest that there is confusion about the constructs, potentially due to either the nature of the construct definition or rater training. Therefore, the second analysis involved estimating interrater agreement among the eight raters in each rating group. In the third analysis, we computed the interrater reliability of the raters 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. This analysis provides information regarding the consistency across raters in terms of how they rate the required level or relative importance of the 35 skill constructs to performance in a particular occupation. Finally, in the fourth analysis, we computed another interrater reliability estimate to examine the consistency of ratings across occupations within constructs. 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. The following sections describe each of the four sets of analyses in greater detail.

Analysis 1: Cycle 23 Recommended Data Flags

Three distinct criteria were established to flag the skill data. All three flags affect the presentation of publicly available data (e.g., O*NET OnLine, My Next Move, O*NET Web Services). First, the level rating of a skill was flagged as not relevant for a particular occupation if at least six of the eight occupational analysts rated its importance as one (1), the lowest possible rating. Thus, the level rating of a skill is considered "not relevant" when that construct is not important for performance in a particular occupation. For example, in the Cycle 23 data, the level ratings for Installation were considered not relevant for Political Scientists (IDI: 00217.03.1) and Pharmacists (IDI: 00318.04.1) because Installation was not considered important for performance in these occupations. In this cycle, there were 300 not relevant flags (see Table 1 for the number of not relevant flags across the past 10 cycles). To facilitate interpretation of these results, it should be noted that there are 2,800 sets of ratings (80 occupations x 35 skills) in the current cycle. Given this, 10.71% (300/2,800) of the skill ratings were flagged as not relevant. The average percentage of skill ratings flagged as not relevant across the previous 22 cycles is 13.18% (*SD* = 3.61%); thus, the percentage of ratings flagged due to this criterion in the current cycle was slightly below the average cycle. The skills with the most flags in Cycle 23 include Installation (*n* = 61), Repairing (*n* = 47), and Equipment Maintenance (*n* = 43). Each of these skills has received large numbers of flags in previous cycles. Given that these constructs capture fairly specific technical proficiencies intuitively not required for many occupations, these results are not surprising.

The remaining two criteria involve the recommended suppression of identifying any skill importance or level mean rating that had an SE_M greater than 0.51. These criteria were established to capture those ratings deemed to have insufficient agreement across raters. The value of 0.51 was selected because 1.00/1.96 = 0.51. An SE_M greater than 0.51 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 23 where the mean importance rating was flagged for insufficient agreement. In fact, no importance ratings received flags for an SE_M greater than 0.51 since Cycle 3. The results of the suppression criteria for level for the past 10 cycles (Cycles 14-23) are presented in Table 2. There were 24 insufficient agreement flags for level ratings in Cycle 23 with the highest number of flags occurring for Operations Analysis and Quality Control Analysis. The percentage of flags indicating insufficient agreement for level ratings in Cycle 23 was 0.86%, which is higher than was observed for previous cycles dating back to Cycle 15, which had 0.98%.

Dating back to Cycle 1, a decreasing trend exists across cycles with respect to the percentage of skill level ratings flagged for having a large SE_M. Exceptions in which there have been increases in flagged ratings across the cycles, such as the increase observed for Cycle 23, have been relatively rare. The increase in agreement observed in cycles over time could be attributable to the fact that most of the occupations rated have also been rated in a previous cycle, and slightly revised rating procedures were introduced to accommodate this large percentage of repeat occupations (Fleisher & Tsacoumis, 2018). In contrast, the decrease in agreement observed in Cycle 23 could be attributed to the fact that 32 of 80 occupations examined were "new" occupations arising from the recent taxonomy update (Green & Allen, 2020; Gregory et al., 2019). It seems reasonable that agreement might be slightly lower because analysts did not have prior mean ratings for these occupations as a source of information to inform their current ratings. That said, 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.

Analysis 2: Cycle 23 Interrater Agreement

Interrater agreement was assessed to determine the level of absolute agreement among the occupational analysts in ratings within a construct for a particular occupation. Measures of interrater agreement index the extent to which the eight raters provided the same rating regarding the level of a skill (e.g., Reading Comprehension) required to perform within a particular occupation. To examine 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 greater 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 occupational analyst mean importance and level ratings across the 35 skills for each occupation. The columns labeled "Median of *SD*s" show the median of the *SD*s associated with each mean importance and level rating across the 35 skills 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 35 skills for each occupation.

The importance ratings across all occupations had a median *SD* of 0.46 and a median SE_M of 0.16. The level ratings across occupations had a median *SD* of 0.52 and a median SE_M of 0.18. These values are slightly higher than Cycle 22 (median *SD* = 0.35 for importance and level, median SE_M = 0.13 for importance and level), but still reflect strong agreement.

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

	Element Name	Cycle 14 (*N* = 106)	Cycle 15 (*N* = 126)	Cycle 16 (*N* = 102)	Cycle 17 (*N* = 116)	Cycle 18 (*N* = 110)	Cycle 19 (*N* = 90)	Cycle 20 (*N* = 100)	Cycle 21 (*N* = 100)	Cycle 22 (*N* = 100)	Cycle 23 (*N* = 80)
1	Reading Comprehension	0	0	0	0	0	0	0	0	0	0
2	Active Listening	0	0	0	0	0	0	0	0	0	0
3	Writing	0	0	0	0	0	0	0	0	0	0
4	Speaking	0	0	0	0	0	0	0	0	0	0
5	Mathematics	1	1	0	0	1	0	1	0	1	1
6	Science	26	26	33	33	38	29	29	27	16	26
7	Critical Thinking	0	0	0	0	0	0	0	0	0	0
8	Active Learning	0	0	0	0	0	0	0	0	0	0
9	Learning Strategies	0	0	0	0	0	0	0	0	1	0
10	Monitoring	0	0	0	0	0	0	0	0	0	0
11	Social Perceptiveness	0	0	0	0	0	0	0	0	0	0
12	Coordination	0	0	0	0	0	0	0	0	0	0
13	Persuasion	0	0	0	0	0	0	0	0	0	0
14	Negotiation	0	0	0	0	0	0	0	0	0	0
15	Instructing	0	0	0	2	1	0	0	0	1	0
16	Service Orientation	0	0	0	0	0	0	0	0	0	0
17	Complex Problem Solving	0	0	0	0	0	0	0	0	0	0
18	Operations Analysis	5	19	8	10	17	21	13	12	4	11
19	Technology Design	8	14	10	14	15	12	8	12	7	5
20	Equipment Selection	45	58	32	50	47	45	59	47	48	31
21	Installation	80	105	70	90	88	72	88	88	86	61
22	Programming	27	30	34	30	28	27	23	30	19	12
23	Quality Control Analysis	6	8	3	5	8	8	9	11	4	5
24	Operations Monitoring	2	0	1	2	1	0	5	5	2	3
25	Operation and Control	20	28	15	19	24	29	44	22	25	20
26	Equipment Maintenance	56	71	36	56	52	55	69	55	59	43
27	Troubleshooting	17	30	16	23	21	23	37	24	21	21
28	Repairing	59	73	40	59	55	55	70	57	61	47
29	Systems Analysis	0	0	0	0	0	0	0	0	0	0
30	Systems Evaluation	1	0	0	0	0	0	0	0	1	0
31	Judg. and Dec. Making	0	0	0	0	0	0	0	0	0	0
32	Time Management	0	0	0	0	0	0	0	0	0	0
33	M. of Financial Resources	13	9	9	7	8	7	9	11	8	8
34	M. of Material Resources	9	7	2	5	5	5	4	8	5	6
35	M. of Personnel Resources	1	0	0	1	0	0	0	0	0	0
	Total Flags out of all possible skill ratings	10.10% (376/3710)	10.86% (479/4410)	8.66% (309/3570)	10.00% (406/4060)	10.62% (409/3850)	12.32% (388/3150)	13.37% (468/3500)	11.69% (409/3500)	10.54% (369/3500)	10.71% (300/2800)

Table 2. Level Flags Due to Large SE_M

	Element Name	Cycle 14 (*N* = 106)	Cycle 15 (*N* = 126)	Cycle 16 (*N* = 102)	Cycle 17 (*N* = 116)	Cycle 18 (*N* = 110)	Cycle 19 (*N* = 90)	Cycle 20 (*N* = 100)	Cycle 21 (*N* = 100)	Cycle 22 (*N* = 100)	Cycle 23 (*N* = 80)
1	Reading Comprehension	0	0	0	0	0	0	0	0	0	0
2	Active Listening	0	0	0	0	0	0	0	0	0	0
3	Writing	0	0	0	0	0	0	0	0	0	0
4	Speaking	0	0	0	0	0	0	0	0	0	0
5	Mathematics	0	1	0	0	0	0	0	0	0	0
6	Science	3	2	2	0	0	1	0	1	0	2
7	Critical Thinking	0	0	0	0	0	0	0	0	0	0
8	Active Learning	0	0	0	0	0	0	0	0	0	0
9	Learning Strategies	0	0	0	0	0	0	0	0	0	0
10	Monitoring	0	0	0	0	0	0	0	0	0	0
11	Social Perceptiveness	0	0	0	0	0	0	0	0	0	0
12	Coordination	0	0	0	0	0	0	0	0	0	0
13	Persuasion	0	0	0	0	0	0	0	0	0	0
14	Negotiation	0	0	0	0	0	0	0	0	0	0
15	Instructing	0	0	0	0	0	0	0	0	0	0
16	Service Orientation	0	0	0	0	0	0	0	0	0	0
17	Complex Problem Solving	0	1	0	0	0	0	0	0	0	1
18	Operations Analysis	12	11	6	6	2	1	1	0	0	7
19	Technology Design	5	5	3	2	0	1	0	0	0	0
20	Equipment Selection	6	2	2	0	0	3	0	1	0	1
21	Installation	5	2	2	3	1	1	0	0	0	2
22	Programming	10	4	2	0	1	0	0	0	0	1
23	Quality Control Analysis	4	5	3	0	2	3	0	0	0	4
24	Operations Monitoring	0	0	0	0	0	0	0	0	0	1
25	Operation and Control	1	1	1	0	0	1	0	0	0	1
26	Equipment Maintenance	0	0	0	1	1	0	0	0	0	0
27	Troubleshooting	1	1	1	1	0	0	0	0	0	1
28	Repairing	1	1	0	0	0	0	0	0	0	1
29	Systems Analysis	0	0	0	0	0	0	0	0	0	1
30	Systems Evaluation	0	0	0	0	0	0	0	0	0	1
31	Judg. and Dec. Making	0	0	0	0	0	0	0	0	0	0
32	Time Management	0	0	0	0	0	0	0	0	0	0
33	M. of Financial Resources	1	4	1	0	1	1	0	0	0	0
34	M. of Material Resources	1	3	0	0	0	0	0	0	0	0
35	M. of Personnel Resources	0	0	1	0	0	0	0	0	0	0
	Total Flags out of all possible skill ratings	1.30% (50/3710)	0.98% (43/4410)	0.67% (24/3570)	0.32% (13/4060)	0.21% (8/3850)	0.38% (12/3150)	0.03% (1/3500)	0.06% (2/3500)	0.00% (0/3500)	0.86% (24/2800)

Analysis 3: Cycle 23 Interrater Reliability—Across Constructs within Occupations

To examine the interrater reliability of the Cycle 23 ratings, we calculated intraclass correlations (*ICC*[C, *k*]; McGraw & Wong, 1996) among the occupational analysts' ratings to assess consistency across constructs within occupations. This statistic indicates the degree of similarity in the rank ordering and relative distance between the skills on a particular scale within an occupation. Our target level of interrater reliability is a median *ICC*(C, *k*) of 0.80 or greater. The value of 0.80 is judged to be a good rule-of-thumb that has been used in multiple contexts, including O*NET (e.g., Clement et al., 2003; McCloy et al., 1999; Rase & Tognetti-Stuff, 1983).

The results of these analyses are presented in Appendix D. The results revealed high levels of interrater reliability across the 80 Cycle 23 occupations. Specifically, the median *ICC* for importance ratings for the skills across the occupations was 0.97 (*M* = 0.95, *SD* = 0.04). The median *ICC* for the level ratings was 0.97 (*M* = 0.96, *SD* = 0.03). The reliability for both the importance and level ratings exceeded the median target coefficient value of 0.80. All the reliability estimates were greater than 0.80, with the exception of the importance reliability for Farmworkers and Laborers, Crop, Nursery, and Greenhouse (*ICC* = 0.77). Overall, the results support a very good level of reliability in the occupational analysts' ratings.

Analysis 4: Cycle 23 Interrater Reliability—Across Occupations within Constructs

Another way to evaluate the reliability of the occupational analysts' ratings is to examine the consistency of the ratings 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 skill Reading Comprehension? To make this evaluation, McGraw and Wong's (1996) *ICC*(C, *k*) is calculated for each construct on each scale (instead of for each occupation on each scale as described above). Consequently, each of the 35 skill importance scale ratings will have a reliability value. A median *ICC*(C, *k*) across the construct ratings for a particular domain on a particular scale of 0.80 or greater is the target interrater reliability for this coefficient (e.g., the median reliability across 35 skill level ratings should be at least 0.80). Again, the value of 0.80 has been judged to be a good rule-of-thumb.

This reliability analysis was conducted for skills on all occupations for the past 10 cycles[1] and results are presented in Table 3. The reliability analyses are based on 1,030 rating targets[2]. 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 skills, the median *ICC*(*C*,8) across the construct ratings for importance was 0.92 (*M* = 0.92, *SD* = 0.03) and for level was 0.95 (*M* = 0.95, *SD* = 0.02). This indicates that on the whole, the reliabilities exceeded the target level. The majority of the skills had high *ICC*(*C*,8) reliabilities for both importance and level. In fact, there were 27 skills with reliabilities equal to or greater than 0.90 for importance (e.g., Science) and all skills had reliabilities equal to or greater than 0.90 for level.

[1] Starting in Cycle 22, interrater reliability analyses across occupations were limited to the past 10 cycles to reflect more recent trends. Previous reports (e.g., Reeder et al., 2020) included all cycles.
[2] A rating target refers to a unique instance of an occupation. An occupation can contribute more than one rating target if it has been rated more than once across data collection cycles.

Table 3. Interrater Reliabilities and Standard Errors of Measurement for Skills Across Occupations in Cycles 14 through 23

		Cycles 14 through 23 (*N* = 1,030)					
		Importance			Level		
	Skill	*ICC(C,1)*	*ICC(C,8)*	*SE*	*ICC(C,1)*	*ICC(C,8)*	*SE*
1	Reading Comprehension	0.64	0.93	0.13	0.78	0.97	0.15
2	Active Listening	0.57	0.91	0.13	0.70	0.95	0.14
3	Writing	0.66	0.94	0.15	0.79	0.97	0.15
4	Speaking	0.60	0.92	0.13	0.74	0.96	0.14
5	Mathematics	0.66	0.94	0.14	0.75	0.96	0.18
6	Science	0.80	0.97	0.16	0.83	0.98	0.22
7	Critical Thinking	0.56	0.91	0.13	0.66	0.94	0.15
8	Active Learning	0.56	0.91	0.15	0.70	0.95	0.17
9	Learning Strategies	0.61	0.93	0.15	0.71	0.95	0.16
10	Monitoring	0.43	0.86	0.14	0.59	0.92	0.16
11	Social Perceptiveness	0.55	0.91	0.14	0.66	0.94	0.16
12	Coordination	0.48	0.88	0.13	0.54	0.90	0.16
13	Persuasion	0.56	0.91	0.15	0.62	0.93	0.18
14	Negotiation	0.58	0.92	0.14	0.65	0.94	0.17
15	Instructing	0.65	0.94	0.14	0.71	0.95	0.16
16	Service Orientation	0.59	0.92	0.14	0.63	0.93	0.16
17	Complex Problem Solving	0.51	0.89	0.16	0.66	0.94	0.16
18	Operations Analysis	0.59	0.92	0.18	0.67	0.94	0.27
19	Technology Design	0.44	0.86	0.15	0.57	0.91	0.22
20	Equipment Selection	0.73	0.96	0.14	0.76	0.96	0.21
21	Installation	0.71	0.95	0.11	0.72	0.95	0.17
22	Programming	0.52	0.89	0.16	0.61	0.93	0.22
23	Quality Control Analysis	0.62	0.93	0.18	0.69	0.95	0.23
24	Operations Monitoring	0.73	0.95	0.15	0.76	0.96	0.20
25	Operation and Control	0.80	0.97	0.15	0.83	0.97	0.20
26	Equipment Maintenance	0.85	0.98	0.12	0.88	0.98	0.16
27	Troubleshooting	0.79	0.97	0.14	0.83	0.97	0.19
28	Repairing	0.86	0.98	0.12	0.88	0.98	0.16
29	Systems Analysis	0.60	0.92	0.15	0.72	0.95	0.17
30	Systems Evaluation	0.60	0.92	0.15	0.72	0.95	0.18
31	Judg. and Dec. Making	0.50	0.89	0.14	0.68	0.94	0.15
32	Time Management	0.40	0.84	0.13	0.56	0.91	0.15
33	M. of Financial Resources	0.55	0.91	0.15	0.67	0.94	0.23
34	M. of Material Resources	0.50	0.89	0.15	0.63	0.93	0.22
35	M. of Personnel Resources	0.57	0.91	0.15	0.65	0.94	0.17

Note. These *ICC*s indicate how consistently raters rated (rank ordered) occupations on a given skill. *SE* = Standard error of measurement = Observed score standard deviation times the square root of one minus *ICC*(*C*,8).

The lowest skill *ICC*(*C*,8) importance reliabilities were found for Time Management (0.84), Technology Design (0.86), and Monitoring (0.86). Even though these skills (Time Management, Technology Design, Monitoring) had the lowest reliabilities compared to other skills this cycle, the reliabilities were still considerably high; the lowest reliability coefficient was above the threshold of 0.80.

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 of 0.80 or greater across constructs, which was achieved for both importance and level (0.92 and 0.95, 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 specific constructs for importance and level.

Summary

The main findings of the analysis of Cycle 23 analyst ratings were as follows:

- More than 89% of the skill ratings were considered important for performance in a given occupation. Constructs that were flagged as not relevant for performance were very similar to those flagged in previous cycles and are conceptually understandable given the specificity of those skills.
- None of the importance ratings were flagged based on a SE_M greater than 0.51.
- Although still low in an absolute sense, a higher percentage of level ratings (0.86%) was flagged for having an SE_M greater than 0.51 compared to recent cycles. This is likely due to many of the occupations examined this cycle having not been rated previously due to updates stemming from the transition to the 2019 O*NET-SOC taxonomy.
- There was strong interrater agreement this cycle as evidenced by the overall low medians of SE_M values.
- All but one of the within-occupation *ICC* reliabilities were above the target value of 0.80. These high levels of interrater reliability indicate that the occupational analysts rank ordered the skills within each occupation similarly on both importance and level.
- All across-occupation *ICC* reliabilities were above the target value of 0.80. These high levels of interrater reliability indicate that analysts rank ordered occupations within each skill similarly on both importance and level.

Given these results, it appears that the analysts are calibrated with one another and understand the skills 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 data collection process with returning analysts prior to each new cycle and as needed, throughout a cycle.

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.

Fleisher, M.S., & Tsacoumis, S. (2018). *O*NET® analyst occupational skills ratings: Procedures update* (FR-11-67). Alexandria, VA: Human Resources Research Organization. https://www.onetcenter.org/reports/AOSkills_ProcUpdate.html.

Green, J.P., & Allen, M.T. (2020). *O*NET-SOC 2019 Taxonomy Development (2020 No. 045).* Human Resources Research Organization. https://www.onetcenter.org/reports/TaxonomyDev2019.html.

Gregory, C., Lewis, P., Frugoli, P. & Nallin, A. (2019). *Updating the O*NET-SOC Taxonomy: Incorporating the 2018 SOC Structure*. National Center for O*NET Development. https://www.onetcenter.org/reports/Taxonomy2019.html.

McCloy, R., Waugh, G., Medsker, G., Wall, J., Rivkin, D., & Lewis, P. (1999, July). *Determining the occupational reinforce patterns for O*NET occupational units (Volume I: Report)*. Raleigh, NC: National Center for O*NET Development. https://www.onetcenter.org/reports/ORP.html.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46. https://doi.org/10.1037/1082-989X.1.1.30

Morgeson, F. P., & Campion, M. A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology, 82,* 627-655. https://doi.org/10.1037/0021-9010.82.5.627

Morgeson, F. P., Delaney-Klinger, K., Mayfield, M. S., Ferrara, P., & Campion, M. A. (2004). Self-presentation processes in job analysis: A field experiment investigating inflation in abilities, tasks, and competencies. *Journal of Applied Psychology, 89*, 674-686. https://doi.org/10.1037/0021-9010.89.4.674

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., Campion, M. A., Mayfield, M. S., Morgeson, F. P., Pearlman, K., Gowing, M. K., Lancaster, A. R., Silver, M. B., & Dye, D. M. (2001). Understanding work through the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology, 54,* 451-492. https://doi.org/10.1111/j.1744-6570.2001.tb00100.x

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. https://doi.org/10.1093/ptj/64.7.1088

Reeder, M. C., Burgoyne, T. C., & Allen, M. T. (2020). O*NET® analyst ratings of occupational skills: Analysis cycle 21 results (*2020 No. 078*). Human Resources Research Organization.

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.