2024 No. 160

O*NET® Analyst Ratings of Occupational Skills: Analysis Cycle 25 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 10, 2024

O*NET Analyst Ratings of Occupational Skills: Analysis Cycle 25 Results

Table of Contents

List of Tables

O*NET Analyst Ratings of Occupational Skills: Analysis Cycle 25 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 a variety of sources, including 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 are 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 jobs; 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 occupational analysts have detailed occupation information 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 skill ratings. 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.

Occupational analysts are provided with relevant occupational information to facilitate the skill rating process. 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 25th analysis cycle of 101 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 https://www.onetcenter.org/ombclearance.html#schedule. Appendix A includes a listing of the IDI codes and Occupational Titles addressed in Cycle 25.

Evaluation of Cycle 25 Analyst Ratings

As mentioned above, occupational analysts provided ratings on the importance and level of the 35 skills for each of the 101 occupations in Cycle 25. 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 confusion about the constructs, potentially due to 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 focuses 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 25 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 six or more 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 25 data, the level ratings for Installation were considered not relevant for Compliance Officers (IDI: 01189.00.1) and Statisticians (IDI: 00200.04.1) because Installation was not considered important for performance in these occupations. In this cycle, there were 395 not relevant flags (see Table 1 for the number of not relevant flags across the past 10 cycles). To facilitate the interpretation of these results, it should be noted that there are 3,535 sets of ratings (101 occupations x 35 skills) in the current cycle. Given this, 11.17% (395/3,535) of the skill ratings were flagged as not relevant. The average percentage of skill ratings flagged as not relevant across the previous 24 cycles is 12.96% (*SD* = 3.52%); thus, the percentage of ratings flagged in the current cycle is below the average across previous cycles. The skills with the most flags in Cycle 25 include Installation (*n* = 86), Repairing (*n* = 58), and Equipment Maintenance (*n* = 56). 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 with 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 25 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 16-25) are presented in Table 2. There were 34 insufficient agreement flags for level ratings in Cycle 25, with the highest number of flags occurring for Operations Analysis and Science. The percentage of flags indicating insufficient agreement for level ratings in Cycle 25 was 0.96%. This value is higher than Cycle 24 (0.00%) and similar to Cycle 23, which had 0.86%.

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 25 and 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 and Cycle 25 could be attributed to the fact that more “new” occupations were rated. In particular, in Cycle 25, 21 of the 101 occupations examined were treated as new occupations due to the O*NET-SOC 2019 taxonomy update (Green & Allen, 2020; Gregory et al., 2019). In Cycle 23, 32 of 80 occupations examined were new occupations arising from the same taxonomy update. 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 for Cycle 25 and prior 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.

Analysis 2: Cycle 25 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.

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

	Element Name	**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)	**Cycle 24** (*N* = 90)	**Cycle 25** (*N* = 101)
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	0	1	0	1	0	1	1	1	0
6	Science	33	33	38	29	29	27	16	26	22	25
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	1	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	2	1	0	0	0	1	0	0	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	8	10	17	21	13	12	4	11	16	16
19	Technology Design	10	14	15	12	8	12	7	5	13	9
20	Equipment Selection	32	50	47	45	59	47	48	31	40	47
21	Installation	70	90	88	72	88	88	86	61	73	86
22	Programming	34	30	28	27	23	30	19	12	24	19
23	Quality Control Analysis	3	5	8	8	9	11	4	5	6	13
24	Operations Monitoring	1	2	1	0	5	5	2	3	1	2
25	Operation and Control	15	19	24	29	44	22	25	20	19	29
26	Equipment Maintenance	36	56	52	55	69	55	59	43	44	56
27	Troubleshooting	16	23	21	23	37	24	21	21	18	28
28	Repairing	40	59	55	55	70	57	61	47	46	58
29	Systems Analysis	0	0	0	0	0	0	0	0	0	0
30	Systems Evaluation	0	0	0	0	0	0	1	0	0	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	9	7	8	7	9	11	8	8	3	4
34	M. of Material Resources	2	5	5	5	4	8	5	6	3	3
35	M. of Personnel Resources	0	1	0	0	0	0	0	0	0	0
	Total Flags out of all possible skill ratings	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)	10.44% (329/3150)	11.17% (395/3535)

Table 2. Level Flags Due to Large SE_M

	Element Name	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)	Cycle 24 (N = 90)	Cycle 25 (N = 101)
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	0	0	0	0	0	0	0	0	1
6	Science	2	0	0	1	0	1	0	2	0	5
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	0	0	0	0	0	0	1	0	0
18	Operations Analysis	6	6	2	1	1	0	0	7	0	11
19	Technology Design	3	2	0	1	0	0	0	0	0	3
20	Equipment Selection	2	0	0	3	0	1	0	1	0	1
21	Installation	2	3	1	1	0	0	0	2	0	1
22	Programming	2	0	1	0	0	0	0	1	0	3
23	Quality Control Analysis	3	0	2	3	0	0	0	4	0	2
24	Operations Monitoring	0	0	0	0	0	0	0	1	0	0
25	Operation and Control	1	0	0	1	0	0	0	1	0	1
26	Equipment Maintenance	0	1	1	0	0	0	0	0	0	2
27	Troubleshooting	1	1	0	0	0	0	0	1	0	1
28	Repairing	0	0	0	0	0	0	0	1	0	0
29	Systems Analysis	0	0	0	0	0	0	0	1	0	1
30	Systems Evaluation	0	0	0	0	0	0	0	1	0	0
31	Judg. and Dec. Making	0	0	0	0	0	0	0	0	0	1
32	Time Management	0	0	0	0	0	0	0	0	0	0
33	M. of Financial Resources	1	0	1	1	0	0	0	0	0	1
34	M. of Material Resources	0	0	0	0	0	0	0	0	0	0
35	M. of Personnel Resources	1	0	0	0	0	0	0	0	0	0
	Total Flags out of all possible skill ratings	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)	0.00% (0/3150)	0.96% (34/3535)

A summary of these results is shown in Appendix C. The columns labeled "Mean of *M*s" show the mean of the occupational analysts' mean importance and level ratings across the 35 skills 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 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 also had a median *SD* of 0.46 and a median SE_M of 0.16. These values are identical to the level values in Cycle 24 but slightly higher for importance (median *SD* = 0.35 and median SE_M = 0.13 for importance in Cycle 24). The importance values are identical to those observed in Cycle 23, and the level values are slightly lower (median *SD* = 0.52 and median SE_M = 0.18 for level in Cycle 23). As noted previously, Cycle 23 had slightly lower agreement, likely due to the increased number of "new" occupations rated in that cycle. Cycle 25 also had an increased number of "new" occupations and the values still reflect strong agreement.

Analysis 3: Cycle 25 Interrater Reliability—Across Constructs within Occupations

To examine the interrater reliability of the Cycle 25 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 101 Cycle 25 occupations. Specifically, the median *ICC* for importance ratings for the skills across the occupations was 0.97 (*M* = 0.96, *SD* = 0.03). The median *ICC* for the level ratings was 0.98 (*M* = 0.97, *SD* = 0.03). The reliability for both the importance and level ratings exceeded the median target coefficient value of 0.80. In fact, all the reliability estimates were greater than 0.90, except for the importance reliability for three occupations (Electrical and Electronic Engineering Technologists and Technicians; Nanotechnology Engineering Technologists and Technicians; Jewelers and Precious Stone and Metal Workers) and the level reliability for four occupations (Electrical and Electronic Engineering Technologists and Technicians; Nanotechnology Engineering Technologists and Technicians; Geological Technicians, Except Hydrologic Technicians; Nuclear Technicians). Overall, the results support a very good level of reliability in the occupational analysts' ratings.

Analysis 4: Cycle 25 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).

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

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 across all occupations for the past 10 cycles,[2] and results are presented in Table 3. The reliability analyses are based on 989 rating targets.[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 median *ICC*(*C*,8) was 0.93 (*M* = 0.93, *SD* = 0.03) across the importance ratings and 0.95 (*M* = 0.95, *SD* = 0.02) across the level ratings. Overall, these values indicate that 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 30 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.

[2] 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.
[3] 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 16 through 25

		Cycles 16 through 25 (*N* = 989)					
		Importance			Level		
	Skill	*ICC(C,1)*	*ICC(C,8)*	*SE*	*ICC(C,1)*	*ICC(C,8)*	*SE*
1	Reading Comprehension	0.66	0.94	0.13	0.79	0.97	0.15
2	Active Listening	0.59	0.92	0.13	0.72	0.95	0.13
3	Writing	0.69	0.95	0.14	0.81	0.97	0.14
4	Speaking	0.62	0.93	0.13	0.75	0.96	0.14
5	Mathematics	0.68	0.95	0.14	0.76	0.96	0.18
6	Science	0.80	0.97	0.16	0.84	0.98	0.22
7	Critical Thinking	0.57	0.91	0.13	0.68	0.94	0.15
8	Active Learning	0.57	0.91	0.15	0.72	0.95	0.17
9	Learning Strategies	0.63	0.93	0.15	0.72	0.95	0.16
10	Monitoring	0.43	0.86	0.14	0.59	0.92	0.16
11	Social Perceptiveness	0.56	0.91	0.14	0.68	0.94	0.16
12	Coordination	0.50	0.89	0.13	0.55	0.91	0.15
13	Persuasion	0.58	0.92	0.14	0.64	0.93	0.18
14	Negotiation	0.60	0.92	0.14	0.66	0.94	0.16
15	Instructing	0.66	0.94	0.14	0.72	0.95	0.16
16	Service Orientation	0.60	0.92	0.14	0.64	0.94	0.16
17	Complex Problem Solving	0.54	0.90	0.15	0.68	0.95	0.16
18	Operations Analysis	0.61	0.93	0.18	0.68	0.95	0.26
19	Technology Design	0.45	0.87	0.15	0.58	0.92	0.22
20	Equipment Selection	0.74	0.96	0.14	0.77	0.96	0.20
21	Installation	0.72	0.95	0.11	0.74	0.96	0.16
22	Programming	0.55	0.91	0.16	0.64	0.93	0.22
23	Quality Control Analysis	0.63	0.93	0.17	0.72	0.95	0.22
24	Operations Monitoring	0.74	0.96	0.15	0.77	0.96	0.19
25	Operation and Control	0.80	0.97	0.15	0.84	0.98	0.19
26	Equipment Maintenance	0.84	0.98	0.13	0.87	0.98	0.17
27	Troubleshooting	0.79	0.97	0.14	0.84	0.98	0.18
28	Repairing	0.85	0.98	0.12	0.88	0.98	0.16
29	Systems Analysis	0.62	0.93	0.15	0.73	0.96	0.17
30	Systems Evaluation	0.62	0.93	0.15	0.72	0.95	0.18
31	Judg. and Dec. Making	0.52	0.90	0.14	0.69	0.95	0.15
32	Time Management	0.41	0.85	0.13	0.56	0.91	0.15
33	M. of Financial Resources	0.55	0.91	0.15	0.67	0.94	0.22
34	M. of Material Resources	0.50	0.89	0.15	0.63	0.93	0.21
35	M. of Personnel Resources	0.58	0.92	0.14	0.66	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 (0.85), Technology Design (0.87), 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 (0.93 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 25 analyst ratings were as follows:

- About 89% of the skill ratings were considered important for performance in a given occupation. Constructs that were flagged as not relevant were very similar to those flagged in previous cycles and, at least facially, appear to be specific in their applicability to certain types of work.
- No 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 (0.96%) were flagged for having an SEM greater than 0.51 compared to the prior cycle (0.00%). This is likely due to some of the Cycle 25 occupations having not been rated previously due to the O*NET-SOC 2019 taxonomy updates (Green & Allen, 2020; Gregory et al., 2019).
- There was strong interrater agreement this cycle, as evidenced by the overall low medians of SE_M values.
- All within-occupation *ICC* reliability estimates 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* reliability estimates 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.

These results suggest 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 before each new cycle and as needed throughout the 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. https://www.onetcenter.org/reports/AOSkills_21.html.

Tsacoumis, S. (2007). *The feasibility of using O*NET to study skill changes* [Presentation]. Workshop on Research Evidence Related to Future Skill Demands organized by the National Academies Center for Education, Washington, D.C.