2022 No. 080

Development of a Revised Deviance Analysis Approach

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: Ki Ho Kim, HumRRO
Matthew C. Reeder, HumRRO
Jessica L. Harris, HumRRO
Jeffrey A. Dahlke, HumRRO
Phil Lewis, National Center for O*NET
Development

Date: August 23, 2022

O*NET Deviance Analysis

Table of Contents

List of Tables

List of Figures

Development of a Revised Deviance Analysis Approach

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 (hereafter referred to as "the Center") is involved in a continual data collection process aimed at identifying and maintaining information on the characteristics of workers and occupations. For years, the Human Resources Research Organization (HumRRO) has supported the Center's efforts to maintain the database.

The Center conducts data cleaning and preparation activities as part of its annual Job Incumbent and Occupational Expert (OE) data collection update efforts. An important part of this process is identifying respondents that are not representative of their sampled and targeted occupation (i.e., deviant respondents). A two-stage *deviance analysis* identifies and removes respondents whose task ratings and other information suggest they are not representative of their assigned occupation.

- **Stage 1**: flags respondents with profiles of task importance ratings that are statistically outlying relative to other cases in the assigned occupation. Cases that are flagged in Stage 1 are referred to as *candidate* deviant cases.

- **Stage 2**: cases flagged in Stage 1 are further evaluated based on a set of exclusion criteria to make a final deviance status designation.

The main objective of the work described in this report was to update and streamline the deviance analysis process. Specifically, the Center approached HumRRO to modernize the statistical outlier analysis and replace a rational review with a process that achieves the same aims while reducing the amount of manual review to the extent possible via automation. These updates have several broad objectives:

- *Reduced turnaround time.* Time to execute the entire deviance analysis process should be reduced from 7-8 weeks to 2-3 weeks.

- *Updated implementation.* The revised procedure was to be transitioned from SAS to an updated implementation in R. Process improvements included program streamlining, the ability to process all occupations in one run without code revisions, and the introduction of automation to Stage 2. The transition to R will also facilitate potential future enhancements to processing and procedures.

- *Comparable level of selectivity.* The revised process should not significantly increase or decrease the proportion of cases flagged as deviant each cycle.

Introducing automation to Stage 2 should reduce the time required to execute this part of the process and increase the validity, reliability, and reproducibility of the evaluations. We also adapted natural language processing (NLP) techniques applied by HumRRO for other O*NET projects (e.g., Dahlke & Putka, 2021) to reduce processing time for the write-in job title review in Stage 2. Although many of the judgments carried out at Stage 2 lend themselves to automation, some manual review will still be required (e.g., for quality control purposes).

This brief technical report summarizes work undertaken to revise the deviance analysis process. We first provide a brief description of the previous deviance analysis procedures, followed by a summary of our approach for updating the procedures. Next, we describe the revised deviance analysis process and evaluate it using input data from Analysis Cycle 22. Finally, we conclude the report by summarizing the work that was executed and reviewing pertinent findings.

Description of Deviance Analysis Procedures

The procedures employed in the previous deviance analysis have been documented in other reports (e.g., National Center for O*NET Development, 2009; U.S. Department of Labor, Employment and Training Administration, 2018, 2021). We provide a high-level description of the process below as context for the present project.

The deviance analysis has been carried out primarily using task ratings gathered in Task Questionnaires that are disseminated each cycle for a subset of occupations in the O*NET-SOC taxonomy. Task importance and relevance ratings are provided by two types of respondents, job incumbents and occupational experts (OEs), the choice of which is dictated by the sampling method applied for each occupation (U.S. Department of Labor, Employment and Training Administration, 2018). The importance ratings are gathered on a 5-point rating scale (1: Not Important, 2: Somewhat Important, 3: Important, 4: Very Important, 5: Extremely Important) whereas the relevance ratings reflect a binary judgment of Not Relevant (0) and Relevant (1).

The prior process began with empirical identification of candidate deviant cases through the execution of multivariate statistical outlier analyses. This focused on identifying respondents with unusually *low* task importance ratings, given that respondents who are unlikely to work in an occupation should be more inclined to indicate that few or no tasks are important to the work they perform. For each occupation, this included the following steps:

1. Estimating unweighted and weighted (robust) Mahalanobis Distance (D^2) values for each respondent, the latter of which is generated from an iterative minimum covariance determinant (MCD) procedure,

2. Flagging respondents as outliers based on a criterion applied to the chi-square values associated with both Mahalanobis D^2 estimates,

3. Identifying the respondent ingroup by removing all cases identified as outliers in Step 2,

4. Comparing the profile of task rating responses for each outlier case against the ingroup response profile using the regression-based "corr plus" procedure, and

5. Flagging outlier respondents with unusually low response profiles as candidate deviant cases.

Output from these analyses included files containing information about the candidate deviant cases. These files served as input to the rational review to determine each case's final disposition (i.e., deviant vs. not deviant). Cases passed to the rational review were evaluated by up to three analysts from the Center. This review began with an examination of each case's task importance and relevance ratings. Specifically, cases where 33% or more of tasks were rated 3 (Important) or higher and 50% or more of tasks were rated as Relevant were deemed not deviant and retained for inclusion in the sample that contributes data for the assigned occupation. Cases that failed to meet these criteria were further evaluated in terms of whether

one or more of the occupation's "key" tasks were rated 3 (Important) or higher. Cases that met this criterion were generally deemed not deviant and retained for inclusion in the assigned occupation. Cases who failed to meet this threshold were then evaluated based on their write-in job title, industry designation, establishment of employment, and other available information (RTI International, 2020). All cases confirmed as deviant during the rational review were subsequently excluded from providing ratings data for their assigned occupations.

Approach

We carried out a four-step approach for updating the existing deviance analysis process. The primary tasks included:

- reviewing materials that describe the previous procedures and gathering requirements from the Center,

- drafting and evaluating the Stage 1 Candidate Deviance (Statistical Outlier Analysis) program,

- drafting and evaluating the Stage 2 Final Deviance (Case-level Exclusion Criteria) program and creating procedures for the write-in job title review, and

- finalizing processes and deliverables based on feedback from the Center and complete documentation.

Primary project outcomes included R programs to execute the Stages 1 and 2 analyses, input and output template ("shell") files in Excel format, documents outlining standard operating procedures (SOPs) for executing the deviance analysis each cycle, and draft output tables in Excel containing results from a trial run of the process using input data from Cycle 22.

Task 1. Review Materials

After the project began, the Center provided HumRRO with various materials, including process documentation, existing SAS programs used for data cleaning and analysis, output tables from Cycle 22 (e.g., summary of aggregated results overall and by occupation), and input data files from Cycle 22 in SAS and Excel format. Our first task was a detailed examination of these materials to ensure understanding of the procedures, input data files, and outputs. A meeting involving staff from HumRRO and the Center was held on March 29, 2022 to discuss clarification questions and to present a path for the project moving forward.

We also evaluated the statistical methods employed (e.g., use of Mahalanobis D^2) against best practices for outlier detection in organizational settings (e.g., Aguinis et al., 2013). That evaluation led to the conclusion that we would not likely need to make significant changes to the actual statistical methods employed. Rather, the primary focus would be on transitioning the process from SAS to R, streamlining the workflow, and incorporating minor modifications to the statistical procedures as necessary.

Task 2. Draft and Evaluate Stage 1 Statistical Outlier Detection Program

Prior to drafting the Stage 1 program, we reviewed the SAS programs to understand the details at a finer-grained level than what might be reflected in standard process documentation. We then drafted an R program that carried out the same general steps involved in flagging

candidate deviant cases based on the MCD and regression-based "corr plus" procedures implemented previously.

Due to small differences in implementation (e.g., use of different missing data procedures), we sought for the Stage 1 R program to exhibit *functional* replication relative to the original SAS program as opposed to exact replication of results. The statistical methods employed in Stage 1 should be relatively robust to these differences, so although we anticipated some variability in results between the SAS and R programs, we expected the results to be highly aligned. To that end, we compared Stage 1 case-level results from the R program against those obtained using the existing SAS infrastructure from Cycle 22 to determine how consistent the two processes were.

Task 3. Draft and Evaluate Stage 2 Final Deviant Program and Create Write-in Title Review Procedures

After drafting a program to execute Stage 1, we then drafted an analogous program for Stage 2. In the original process, a manual review was carried out by analysts, so we relied on process documentation and discussions with the Center to inform our development approach. As described previously, the following pieces of information were evaluated as part of the rational reviews:

- proportion of tasks rated 3 or higher on importance and proportion of tasks rated as relevant,

- number of "key" tasks rated 3 or higher on importance, and

- other case-level information (i.e., write-in job title, industry designation, establishment of employment, and other available information).

Discussion with the Center resulted in two changes to the information considered in the revised Stage 2 process. First, whereas the write-in job titles were judged to be useful for identifying deviant cases, some of the other information evaluated late in the rational review (e.g., industry designation, establishment of employment) was not deemed to be diagnostic of each case's final deviance status. Consequently, we streamlined the process by retaining the write-in job title review in Stage 2 and removing industry designation, establishment, and other information.

Second, one piece of information that had not been formally included in Stage 2 previously was each case's global match status. The Task Questionnaires include a "global match" item that asks each respondent to read the O*NET occupation description for the assigned occupation and rate how well the description matches the respondent's job (incumbents) or how much expertise the respondent has (OEs) using a 4-point rating scale. Higher values on the scale denote poorer match, where a value of 4 indicates that the occupation description does not at all describe what the respondent does or has expertise in. The decision to include global match was based on the premise that respondents who are flagged at Stage 1 and who endorse a 4 on the global match item are unlikely to work in or have expertise pertaining to the assigned occupation and, thus, should not be included in the analysis sample.

In addition to developing a program to execute components of Stage 2 that are subject to automation, we also developed procedures for reviewing respondents' write-in job titles. Under the revised process, the write-in title review involves a comparison between the job title provided by the respondent against alternate titles associated with each occupation in the

O*NET-SOC taxonomy. The primary goal of the review is to determine whether the respondent's job title is a match (exact or very close) to any alternate titles associated with the assigned occupation.

The write-in title review is performed by two HumRRO researchers who are trained on a set of standard operating procedures. Each researcher examines all write-in job titles that are subject to review and provides their judgments (i.e., match or not) independently. Instances of reviewer disagreement are adjudicated by a third quality control (QC) reviewer. Because there are thousands of alternate titles (over 53,000 as of Database 26.3), we developed NLP-based procedures to curate the list of comparisons to only those that most closely match each write-in job title under review. The same cosine-based method employed in prior research by HumRRO for O*NET was adapted here (see Dahlke & Putka, 2021 for details).

The write-in title review only applies to job incumbents because incumbents are sampled from establishments based on employment in jobs whose work and titles should be reflective of the assigned occupation. This contrasts with OEs, who may work in qualitatively different occupations (e.g., manager, supervisor, trainer) than the assigned occupation and, thus, are inherently less likely to have titles indicative of the assigned occupation. As discussed in the next section, the write-in title review is the final step in Stage 2. Incumbents who progress to the write-in title review and whose titles match an alternate title for the assigned occupation are considered not deviant; all other write-in titles are treated as deviant. Because OEs are not subject to the write-in title review, any OE cases that make it to this component of Stage 2 are automatically considered deviant.

Task 4. Finalize Deliverables and Document Process

Draft versions of all project deliverables were provided to the Center for review. Feedback provided by the Center was used to revise the deliverables. Upon completing all revisions, final versions were submitted to the Center.

Description and Evaluation of Revised Deviance Analysis Process

The subsections below provide detail on the revised deviance analysis process, describe a small-scale pilot study conducted to evaluate the write-in title review procedures, and highlight results of analyses carried out to evaluate the revised process.

Description of Revised Process

Figure 1 illustrates the steps in the revised deviance process and their sequencing.
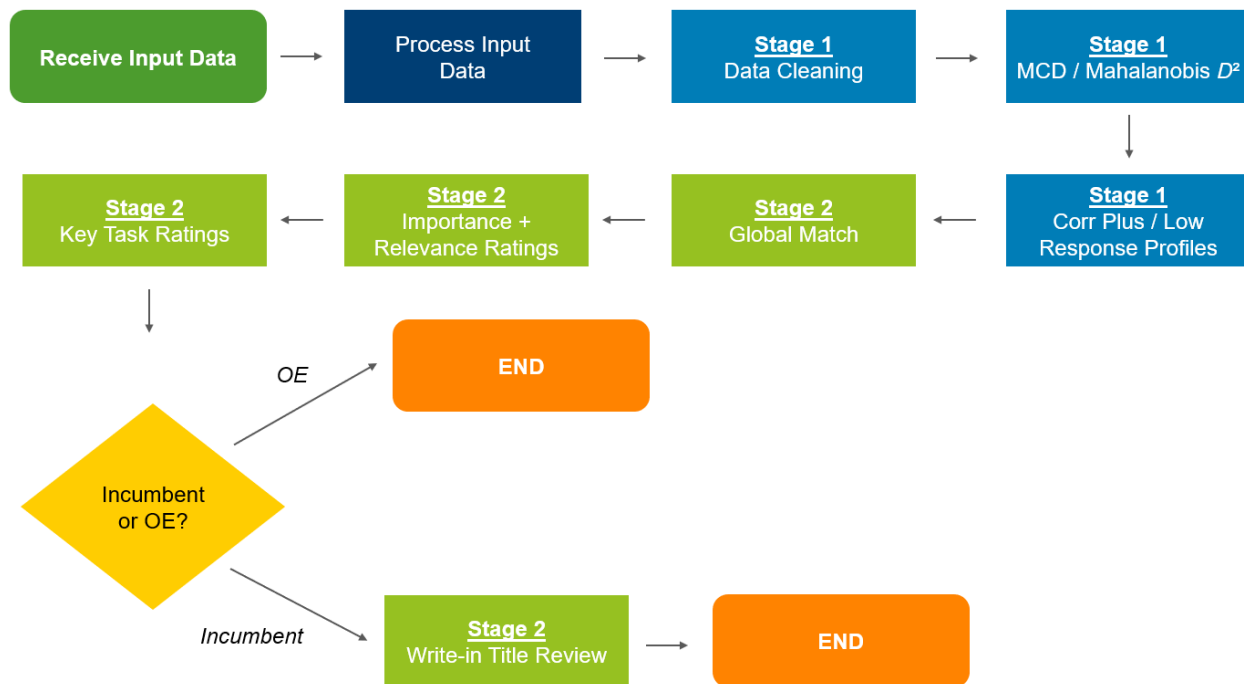

Figure 1: Diagram of revised deviance analysis procedures.

The process begins with receipt of input data needed to carry out the deviance analysis, including a respondent-level data file that contains task ratings and associated information for each respondent (e.g., write-in job title, global match item) and an occupation-level data file that includes task lists and key task designation(s) for each occupation being examined. After data receipt, initial processing and data cleaning activities occur, such as ensuring that all occupations have at least one key task and recoding indicators of missingness or non-valid response to a missing value.

Next, Stage 1 procedures are analogous to those implemented previously to identify statistically outlying respondents with low task importance ratings. Records that are flagged at Stage 1 as candidate deviant cases then proceed to Stage 2, which includes a series of sequential steps ordered as multiple hurdles for both incumbents and OEs. At each step, a case either (a) meets the criterion, is deemed not deviant, and is removed from Stage 2 processing as a valid case to be included with the assigned occupation or (b) fails to meet the criterion, is considered potentially deviant, and passes to the next Stage 2 step for further evaluation.

Table 1 shows the criteria applied at Stage 2 for both incumbents and OEs.

Table 1. Stage 2 exclusionary criteria applied to incumbents and occupational experts (OEs)

Step	Not Deviant	Deviant
Global Match	≠ 4	4
Importance & Relevance	33% or more rated 3 or higher on Importance **AND** 50% or more rated Relevant	Less than 33% rated 3 or higher on Importance **OR** less than 50% rated Relevant
Key Tasks	One or more key tasks endorsed	Zero key tasks endorsed

Decisions regarding sequencing of the criteria were based on rational and pragmatic considerations. Conceptually, the steps are ordered from broadest to narrowest, where breadth reflects how targeted (i.e., general to specific) the information under consideration is. Global match is the broadest Stage 2 decision criterion in that it is based on a general evaluation of how well the work associated with the assigned occupation aligns with the respondent's work or expertise in terms of duties and responsibilities overall. Placing the global match item at the first step also makes sense from a data quality and validity perspective because it excludes respondents who neither work in nor have expertise with the occupation by their own admission.

Following the global match step, examination of importance and relevance ratings focuses on data for individual tasks chosen specifically for each occupation and, at Step 3, importance ratings of a smaller subset of tasks deemed by experts as being "key" to each target occupation. Because both steps were part of the prior rational review and one of our stated objectives was to maintain a comparable level of selectivity as before, we adopted the standards previously applied to both steps for the revised implementation.

Another consideration for sequencing the Stage 2 steps pertains to practical matters associated with time and cost of execution. One way to reduce time and cost in a process such as this is to sequence any manual or labor-intensive steps later in the process. This allows earlier steps, which are automated, to lessen the amount of work required in the manual steps by reducing the number of cases requiring manual review.

Incumbents who fail the key task step are then submitted to the write-in job title review described previously. If a case's write-in title is deemed a match against one or more alternate titles associated with the assigned occupation, then the case is considered not deviant. Otherwise, the case is considered deviant and removed from the pool of cases that contributes data for the assigned occupation. Because the write-in title review only applies to incumbents, OE respondents that fail the key task criterion (i.e., zero key tasks endorsed) are considered deviant and are excluded from the assigned occupation.

Upon completion of the deviance analysis, output tables are prepared, including a case-level summary containing each respondent's final disposition (i.e., deviant or not deviant) and an aggregate summary that contains frequencies of cases at each stage in the process, both overall and by occupation.

Write-in Job Title Review Pilot Study

A pilot study of the write-in job title review step was undertaken after all other procedures (i.e., Stages 1 and 2 R programs, write-in title review SOPs) were developed. The goals of the pilot study were to: (a) ensure that the reviewer instructions are clear and comprehensive, (b) examine agreement between reviewers in evaluating write-in titles, and (c) estimate how long the process will take when implemented in an actual analysis cycle. Three reviewers took part in the pilot study. All reviewers were research staff at HumRRO with backgrounds in Industrial-Organizational Psychology who had previously worked on other O*NET activities.

Write-in titles evaluated for the pilot study came from the Cycle 22 input data files. After drafting the Stage 1 and 2 R programs, we processed the Cycle 22 data using the programs. We then identified incumbent cases that were still considered deviant after the key task step in Stage 2, and these cases constituted those examined in the pilot study. A total of 74 write-in titles were reviewed, 11 of which were reviewed by all three reviewers to evaluate agreement. The remaining 63 titles were divided into three nonoverlapping groups of 21 titles, and each group was assigned to a reviewer. Thus, each reviewer evaluated a total of 32 write-in job titles: 21 unique titles and 11 common titles.

Of the 74 titles, there were 16 instances (22% of all titles reviewed) where one reviewer (unique) or all three reviewers (common) decided that the title was not deviant. Of the remaining 58 titles, 54 (73%) were evaluated as deviant. The remaining four titles were cases reviewed by all three reviewers and where there was disagreement. In two instances, two of the three reviewers judged the title as deviant while the remaining reviewer judged it as not deviant; for the other two instances, two of the three reviewers judged the title as not deviant and the remaining reviewer disagreed. Feedback solicited from the reviewers indicated that disagreement stemmed from user error (i.e., the reviewer realized their initial judgment was not what they intended) or ambiguity regarding whether a write-in title matched one or more alternate titles associated with the assigned occupation. When the process is implemented operationally in coming cycles, user error issues will be addressed by collecting decisions from two reviewers, followed by a QC reviewer who will reconcile any initial disagreements. Furthermore, we addressed the issue of ambiguous title matching by revising the standard operating procedures to instruct reviewers that ambiguous matches should be evaluated as a non-match.

Evaluation of Revised Process

The first analysis we ran to evaluate the revised process involved using the Cycle 22 data to compare respondents' Stage 1 candidate deviance status from the revised process in R against the results obtained from the prior process in SAS. As mentioned previously, although our aim was not exact replication of results, we did expect the two processes to generate results that are largely congruent since both programs contain the same broad steps.

Of the 3,747 cases available from Cycle 22, 259 (6.9%) were flagged as candidate deviant based on the SAS program and 257 (6.9%) were flagged based on the R program. Thus, the two processes are highly similar in terms of the frequency of cases flagged. Table 2 contains crosstabs of respondents' status based on the original SAS process against the revised process

in R. As shown, the two processes are highly convergent in terms of agreement in case status. Cohen's kappa, κ, estimated agreement to be 0.99. Of the 3,747 cases, there were only 38 instances of disagreement.

Table 2. Correspondence in Candidate Deviance Status for Original - SAS and Stage 1 - R programs

Original - SAS	Stage 1 - R	
	Candidate Deviant	**Not Deviant**
Candidate Deviant	240	19
Not Deviant	17	3,471

Next, we generated frequencies for the entire revised process to illustrate the flow of cases across steps. Tables 3 and 4 show the number of cases from Cycle 22 processed at each stage of the process for incumbents and OEs, respectively. At the conclusion of the process, 81 out of 3,747 incumbents (2.2%) and 7 out of 518 OEs (1.4%) were flagged as deviant. For incumbents, the total number of deviant cases comes from 16 cases that were flagged at the global match step and additional 65 cases flagged from the write-in title review. For OEs, the total number of deviant cases comes from zero cases that were flagged at the global match step and seven cases that were still considered deviant following the key task step.

Table 3. Number of Incumbents Processed at Each Stage of Revised Deviance Analysis Process

Stage	Step	Description	Status	n	Percent At Step	Percent Of Total
1	1	Statistical Outlier Analysis	Candidate Deviant	257	6.9%	6.9%
			Not Deviant	3,490	93.1%	93.1%
2	1	Global Match	It does not at all describe what I do (4)	16	6.2%	0.4%
			Else	241	93.8%	6.4%
2	2	Importance & Relevance	33% or more rated $\geq$ 3 on Importance & $\geq$ 50% or more rated Relevant	71	29.5%	1.9%
			Less than 33% rated $\geq$ 3 on Importance OR < 50% rated Relevant	170	70.5%	4.5%
2	3	Key Tasks	At least 1 key task rated $\geq$ 3 on Importance	88	51.8%	2.3%
			Zero key tasks rated $\geq$ 3 on Importance	82	48.2%	2.2%
2	4	Write-in Job Title Review	Match	17	20.7%	0.5%
			No Match	65	79.3%	1.7%
		FINAL DEVIANCE STATUS	Deviant	81	2.2%	2.2%
			Not Deviant	3,666	97.8%	97.8%

Table 4. Number of Occupational Experts (OEs) Processed at Each Stage of Revised Deviance Analysis Process

Stage	Step	Description	Status	n	Percent At Step	Percent Of Total
1	1	Statistical Outlier Analysis	Candidate Deviant	33	6.4%	6.4%
			Not Deviant	485	93.6%	93.6%
2	1	Global Match	No expertise on the work performed in the occupation (4)	0	0.0%	0.0%
			Else	33	100.0%	6.4%
2	2	Importance & Relevance	33% or more rated $\geq$ 3 on Importance & $\geq$ 50% or more rated Relevant	22	66.7%	4.2%
			Less than 33% rated $\geq$ 3 on Importance OR < 50% rated Relevant	11	33.3%	2.1%
2	3	Key Tasks	At least 1 key task rated $\geq$ 3 on Importance	4	36.4%	0.8%
			Zero key tasks rated $\geq$ 3 on Importance	7	63.6%	1.4%
		FINAL DEVIANCE STATUS	Deviant	7	1.4%	1.4%
			Not Deviant	511	98.6%	98.6%

Conclusion

The work characterized in this report was carried out to update and streamline the deviance analysis process used to identify respondents who are not representative of their assigned occupation. To that end, we developed a revised set of procedures implemented in R that includes the statistical outlier analyses executed at Stage 1 and automates the case-level exclusionary criteria previously examined during the rational review. In addition, we developed an approach for executing the write-in job title review for job incumbents as part of the revised Stage 2 steps. The revised approach will reduce the amount of time required to execute the deviance analysis each cycle through a process designed to be standardized, reproducible, and efficient.

The deviance analysis will be implemented by HumRRO staff each year as a standard cyclical activity. A brief summary report will be produced for each Analysis Cycle. We will continue to monitor the process and identify potential areas of improvement that may further reduce turnaround time and ensure the procedures identify deviant respondents in a valid and reliable manner.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*(2), 270-301. https://doi.org/10.1177/1094428112470848.

Dahlke, J. A., & Putka, D. J. (2021). *Streamlining the identification of emerging tasks in the O*NET system using natural language processing (NLP): Technical summary (Report 2021 No. 100).* Alexandria, VA: Human Resources Research Organization. https://www.onetcenter.org/dl_files/EmergingTasksNLP.pdf

National Center for O*NET Development. (2009). *O*NET Data Collection Program: Statistical procedures for deviant case detection.* Raleigh, NC: Author.

RTI International. (2020). *O*NET 2019 methodology report*. Research Triangle Park, NC: Author.

U.S. Department of Labor, Employment and Training Administration. (2018). *O*NET® Data Collection Program: Office of Management and Budget clearance package supporting statement.* Washington, DC: Author. https://www.onetcenter.org/dl_files/omb2018/Supporting_StatementA.pdf

U.S. Department of Labor, Employment and Training Administration. (2021). *O*NET® Data Collection Program: Supporting statement for O*NET Data Collection Program.* Washington, DC: Author. https://www.onetcenter.org/dl_files/omb2021/Supporting_StatementA.pdf

Van Buren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1-67. https://doi.org/10.18637/jss.v045.i03.

Yuan, Y. (2011). Multiple Imputation Using SAS Software. *Journal of Statistical Software*, *45*(6), 1-25. https://doi.org/10.18637/jss.v045.i06.