

ENHANCING O*NET'S USEFULNESS FOR ACADEMIC RESEARCHERS

Michael J. Handel

Sociology and Anthropology Department

Northeastern University

m.handel@neu.edu

March 23, 2009

O*NET represents a monumental effort to collect and codify occupational information for use by students, job seekers, career counselors, government agencies, and human resource, education, and training professionals, among others. The O*NET web site has a "Developer's Corner" with resources for third parties to develop additional products for career exploration using the O*NET database.

Unfortunately, another potential user group, academic researchers, has not figured prominently in O*NET's dissemination efforts. O*NET has great potential interest for sociologists, economists, policy researchers, and other social scientists interested in job skill requirements, workplace autonomy and decision making, occupational safety and health, and the myriad other job characteristics represented in the O*NET database. Given researchers' widespread use of O*NET's predecessor, the Dictionary of Occupational Titles (DOT) (Miller et al. 1980), there is great opportunity to widen the circle of users for this resource.

One useful way to think about making O*NET more accessible to the research community is to think about creating a Researcher's Corner on the O*NET web site that would be a central repository for the kinds of files and documentation academic researchers find useful and important. The rest of this paper describes how O*NET could be made more useful to researchers and offers recommendations for the content of a Researcher's Corner on the O*NET web site.

1. Original O*NET microdata. Currently, the individual-level survey data collected by O*NET is used only to construct occupational means. Researchers would be very interested in using the individual-level data to model O*NET variables as functions of the respondent characteristics collected on the Background survey, such as age, gender, race/ethnicity, education, job tenure, disability, and employer type (e.g., for-profit, non-profit, government). Understanding these relationships with aggregate-level data is much more problematic. Adding Background questions on wage and salary levels modeled on the Current Population Survey would also make the O*NET microdata an invaluable resource for researchers interested in labor markets. Currently, the full research potential of the O*NET database is unrealizable because the original survey data is unavailable.

Recommendation: Add wage and salary questions to the Background survey and make available all individual-level data files in ascii format on a Researcher's Corner section of the O*NET web site.

2. Sample information. Even in the absence of O*NET microdata, there are informational gaps that are potentially troubling for researchers. It is very difficult for a researcher to determine some basic facts about the O*NET sample, such as total number of respondents and number of respondents per occupation per questionnaire. Response rates are somewhat more accessible if one searches for them. Likewise, O*NET collects a great deal of background information on its respondents, but a search of the web site does not reveal much detail on the subject. O*NET has made great efforts to achieve a more representative sample than its predecessor, the DOT. However, most researchers

would like some short, formal documentation of O*NET sample characteristics to understand how closely it reflects the distribution of the overall workforce with respect to age, gender, race, ethnicity, education, industry, region, and English proficiency, as proxied by number of respondents using Spanish versions of the surveys. To take the most obvious example, if response rates within occupations are higher for more educated workers, which might be expected from the experience of other surveys, the estimates of mean occupational skill requirements might be higher than the population values. Obviously, this is an issue for all voluntary surveys, which usually compensate with sampling weights, but the main point is that researchers using the O*NET database need to have some basic information on the characteristics of the underlying sample to understand its strengths and limitations.

Recommendation: Publish a brief note on the sample size, response rates, and distribution of O*NET respondents across age, gender, race, ethnic, language use, education, industry, and region groups on a Researcher's Corner section of the O*NET web site.

3. Researcher-friendly data files. Currently the O*NET database is available in various formats, including SAS-PC on a linked web site. Unfortunately, the data set is distributed across numerous files, perhaps reflecting O*NET's origins as a relational database. While some researchers may appreciate the current, modular organization, many will prefer a consolidated file. I suspect there will be considerable duplication of effort as numerous researchers construct their own consolidated files. The absence of such a file probably hinders dissemination of the O*NET database to the research community.

Recommendation: Construct a single O*NET file consolidating the information from the component files currently available using a recognized format (e.g., SAS, Stata) and make available for download on a Researcher's Corner section of the O*NET web site. A codebook written especially for this file would also be helpful.

4. Maintain prior waves for longitudinal research. One of the most valuable aspects of the O*NET data collection design for researchers is the resurveying of occupations every five years. A longstanding frustration with the DOT was the absence of longitudinal information on trends in skill requirements and other occupational characteristics. The possibility of such data for O*NET represents a major advance.

Recommendation: All editions or versions of the O*NET database should be available on a Researcher's Corner section of the O*NET web site for researchers interested in longitudinal studies of changing occupational characteristics.

5. Occupational crosswalk enhancements. Most social science researchers would like to merge O*NET measures onto other datasets, such as the Current Population Survey, which use occupation codes that differ from O*NET, which is itself a modified version of the standard SOC classification system. It would be very helpful if a crosswalk were available that resolved a number of small but nettlesome issues in collapsing O*NET codes to more standard occupational classification systems. Existing crosswalks still leave the researchers with a number of coding decisions for which it would be helpful to have an authoritative resolution.

Recommendation: Make available files that crosswalk O*NET occupation codes to both standard SOC codes and Census occupation codes on a Researcher's Corner section of the O*NET web site.

6. Tools and technology (T2) data file. This is a relatively recent addition to O*NET and potentially quite useful for researchers interested in job skill requirements and how they are changing over time. It is not clear to me that any of this information is available except through the O*NET OnLine interface (though I may be mistaken). Apparently over 32,000 tools and technology are already represented in the database, so making this information available to researchers might be a challenge. One simple solution is to create another variable for each occupation that is simply a count of the number of tools and technology objects used. However, since the database apparently uses UN Standard Products and Services Codes (UNSPSC), it may be possible to create a file with complete information on T2 objects for research users.

Recommendation: Study possibilities for creating a tools and technology file containing fields for a count of the total number of T2 objects used by each occupation and all individual objects by UNSPSC code.

O*NET represents an impressive effort to develop a systematic body of information on occupational characteristics, but it remains under-utilized among academic researchers. Hopefully, the preceding recommendations will help widen the circle of users for this important resource. I would be happy to consult with anyone who would like further clarification on any of these issues.

Reference

Miller, Ann R., Donald J. Treiman, Pamela S. Cain, and Patricia A. Roos. 1980. Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles. Washington, D.C.: National Academy Press.