

# Reflections on the Growth and Impact of Large-Scale Assessments of Adult Skills

Irwin Kirsch, Mary Louise Lennon, and Laura Halderman, Educational Testing Service

From modest beginnings in the late 1950s, large-scale assessments have experienced consistent growth. Increasing interest in these surveys reflects not only the recognition of the importance of cognitive skills and skill development for both economic growth and societal well-being, but also the value in benchmarking performance against peers and economic competitors. Over this period of some 60 years, these surveys have also contributed to significant advances in new methodologies and various aspects of measurement science, as well as pioneering the full incorporation of digital technologies into survey design, development and implementation. Although much of the early work around large-scale assessments focused on student populations, since the 1980s there has been growing interest in adult surveys both at the national and international level. Over this time period, policy makers, researchers and other key stakeholders have raised new questions that focused on adults such as: What is the relationship between education and cognitive skills? How are skills related to labor force success? What factors contribute to skill acquisition and decline across age cohorts? The assessments developed to address such policy questions have contributed to a deeper understanding of the distribution of skills among adult populations and the connection those skills have to social, educational and labor market outcomes (Kirsch et al., 2017).

The latest international survey of adult skills is the second cycle of the Programme for the International Assessment of Adult Competencies (PIAAC). Adults ages 16-65 in over 30 countries/economies have participated in the survey and results are due to be released by the Organization for Economic Co-operation and Development (OECD) in December of 2024. The focus

of this paper is to highlight the innovations in the first and second cycles of PIAAC that extend the utility of these assessments beyond what was achieved in earlier surveys of adult skills and to discuss their impacts. To do so, it is helpful to have a general understanding of how large-scale assessments have evolved from their earliest iterations to the present.

## **National Large-Scale Assessments of Adult Skills: 1984 - 1993**

The first three assessments that established the foundation for future adult surveys, including PIAAC, were all paper-based assessments administered in the United States. These included: the Young Adult Literacy Survey (YALS, 1984), the assessment of adult participants in the United States Department of Labor training and unemployment programs (DOL, 1990) and the National Adult Literacy Survey (NALS, 1993).

Each of these surveys reflected an important shift from a focus on school-based reading to a broader conceptualization of literacy that reflected the diversity of tasks that adults encounter at work, home and school, and in their communities. Importantly, these surveys collected data that supported the notion of literacy as a set of complex information-processing skills that extend well beyond decoding and basic comprehension. Research based on this work helped define those information-processing skills in ways that could contribute to a deeper understanding of what is being assessed and set the stage for a move from national to international assessments of adult skills (Kirsch, 2001, 2003).

## International Large-Scale Assessments of Adult Skills: 1994-2008

As policy interest in the skills of adults grew, two paper-based international assessments focusing on adults ages 16-65 were developed and implemented: the International Adult Literacy Survey (IALS) and the survey of Adult Literacy and Lifeskills (ALL). IALS, which assessed prose, document and quantitative literacy, was conducted in multiple rounds from 1994-1999 resulting in a total of 19 participating countries. The ALL survey assessed literacy, numeracy and analytic problem solving and was conducted between 2003 and 2008 in 11 countries. Both IALS and ALL were designed to profile and explore the distribution of literacy skills among populations within and across participating countries and each expanded what was measured, both in the background questionnaires and the cognitive domains.

The work associated with developing and implementing IALS and ALL formed a knowledge base that contributed to the development and implementation of PIAAC in several important ways. Importantly, these early assessments established that it was possible to collect internationally comparable data in an assessment that is administered in multiple languages across a diverse set of countries. Additionally, these assessments demonstrated that incorporating sophisticated background questionnaires into the surveys made it possible to investigate the connections between the skills being measured and important personal and social outcomes including labor force participation, literacy activities and social participation.

## The Introduction of Technology-Based Assessments: PIAAC Cycle 1 (2012 – 2017)

PIAAC Cycle 1 was the first household survey to design and deliver an international large-scale assessment of adult skills on a technology platform. Administered in three rounds from 2012 through 2017, PIAAC was unprecedented in scope, assessing close to 200,000 adults across 39 countries. While linked by design to IALS and ALL, including sets of questions from these

previous surveys, PIAAC refined and expanded the existing assessment domains and introduced two new domains as well. The main instruments in PIAAC Cycle 1 included a comprehensive background questionnaire and cognitive assessments focused on literacy, numeracy, reading components and problem solving in technology-rich environments (PS-TRE).

The successful conduct of PIAAC was important in that it paved the way for other surveys of student and adult skills to transition from paper-based to technology-based assessments. It also improved what could be assessed both in terms of extending legacy constructs such as literacy and numeracy to include digital content and also allowing for the introduction of new constructs such as PS-TRE.

## PIAAC Cycle 2 (2018-2024)

The second cycle of PIAAC is the culmination of all that was learned in the process of developing, delivering and interpreting the results of previous large-scale assessments of adult skills. Like the first cycle, the main instruments in Cycle 2 include a comprehensive background questionnaire and cognitive assessments of literacy and numeracy skills. Adaptive problem solving, the new cognitive domain in Cycle 2, focuses on problem solving in the context of dynamically changing situations. Finally, to improve the measurement of adults with lower skills, measures of reading and numeracy component skills are included in the assessment, with the latter being introduced for the first time in a large-scale assessment of adults.

In addition to extending what was assessed, PIAAC Cycle 2 introduced other innovations that expanded the impact of the survey. For example, the introduction of a simple tablet-based design for the administration of all survey instruments meant that the need for paper administration was eliminated. As a result, more data could be collected about the digital literacy skills of all participating adults. Additionally, performance data, including timing and response patterns could be captured in the log files for each task. While the potential of log file data and analyses is yet to be fully realized, these data have already provided valuable insights into how adults at different skill levels approach literacy, numeracy and problem-solving tasks.

## Impact of the PIAAC Surveys

International large-scale assessments of adults have proven over the years to be a reliable source of credible information about the distributions of skills and how they may be changing over time. In addition, PIAAC data provide participating countries with an opportunity to examine how performance on the set of skills assessed relates to a host of social, educational and labor market outcomes. As such, there has been an increased interest in using PIAAC data for a variety of purposes including secondary analyses, the creation of derivative products, estimating skills for subpopulations, and connecting the assessment materials to targeted interventions.

### Secondary Analyses

The availability of large-scale assessment data in general, and PIAAC in particular, has led to a variety of secondary analyses along with the development and publication of research papers and policy reports that address a variety of issues. For example,

- The GESIS Leibniz Institute for the Social Sciences has been developing a bibliography that contains more than 900 references to papers and reports relating to PIAAC and its precursor assessments including available abstracts and links to the various publications. The current version covers the period from 2008, when the first cycle of PIAAC began, through 2023 (Maehler et al., 2023).
- Work within ETS has focused on the development of policy reports that use PIAAC data, along with other assessment data, to examine a variety of issues that center around the growing importance of skills and their distribution both within and across participating countries. Among other findings, these reports reflect the growing relationship between human and social capital and their connections to opportunity and overall well-being. These policy reports also include a series on labor market outcomes associated with skills, educational attainment and income as well as information on a representative sample of the incarcerated population and older adults between the ages of 66 and 75. Access to these reports can be found at: <https://www.ets.org/research/policy/human-capital-education/reports>.

### Derivative Products

Work associated with PIAAC Cycle 1 and earlier international assessments has been extended through derivative products that make use of the content, development processes and procedures, and data from these assessments for new purposes. For example, national large-scale assessments in the U.S. in the 1990s formed the basis for several derivative products including: the Test of Applied Literacy Skills (TALS), a paper-and-pencil test that yielded individual-level results; a multi-media group-based instructional system for adults that focused on prose, document and quantitative literacy; and the PDQ Profile series, an adaptive computer-based assessment of literacy proficiency for individuals that was linked to the NALS and IALS scales.

Following a similar model, Education & Skills Online (ESO) was developed in conjunction with the OECD and several interested countries as an online adaptive assessment designed to provide individual-level results in some 16 language versions that are linked to PIAAC Cycle 1. Measures of literacy and numeracy are included in this derivative product, as well as optional assessments of reading components and problem solving in technology-rich environments. Because of its link to PIAAC, results from ESO can be benchmarked against national and international results for participating countries. An optional assessment of non-cognitive skills is also included in the product. The primary purpose of ESO is to provide information about the skills of individuals, either to inform training efforts or for research purposes. As such, the OECD identifies potential users and ways in which the ESO can be used.

### Estimating Skills for Subpopulations

While PIAAC provides a direct measure of the skills and background characteristics of large national samples, researchers and policy makers are often interested in estimates of the proficiency distribution of selected subpopulations that go beyond the selected sample of adults who participated in the main study. One way to obtain this information is through indirect estimates based on information collected from the direct assessment along with other known characteristics of the population in a desired geographical area or areas. The methods used to obtain this information are called “small area estimates.” Small area estimation is a methodology that

can be used to leverage that relationship in order to estimate skills for subpopulations who have not taken the cognitive assessment but for whom other sources can provide some of the same background data as that collected in PIAAC. For example, small area estimation models were used in the US to develop the Skills Map, which provides reliable estimates of adult literacy and numeracy skills in all 50 states and the District of Columbia. See <https://nces.ed.gov/surveys/piaac/skillsmap/>

### Connecting the Assessment to Targeted Interventions

In addition to ESO and small area estimation, there are other potential benefits for countries participating in PIAAC. One of these involves connecting the assessment materials to targeted interventions that can be used with various groups of adult learners to help them improve their skills. Another relies on the materials and trainings that are provided by the contractors and the OECD that can extend the capabilities of current and new participating countries. Each is discussed briefly here.


- Current assessments such as PIAAC represent recent advances in measurement science and provide a model for the design and development of cognitive instruments that focuses on the collection of validity evidence to support both their development and appropriate use. Developing a deeper understanding of what is being assessed and how it relates to various outcomes can contribute to the development of more effective interventions. That is, linking learning materials with these assessments requires the same rigor and understanding as needed to develop the assessments. The approach used to develop this type of assessment system is referred to as construct based, or evidence-centered design, and offers a critical roadmap for the development of coherent and quality interventions (Kirsch et al., 2021).
- In addition to a deeper understanding of what is being assessed and how to connect this to the development of learner materials and instructor development, PIAAC produces other materials such as a technical report and a set of standards and guidelines that can also help provide participating countries with information that can expand their capabilities and experience with developing and administering these types of assessments, as well as how to use and

analyze the data. Data files are provided by the OECD, as well as by individual countries, as public use and restricted use files. As an example, after the first cycle of PIAAC, the U.S. government sponsored a multiyear program where interested researchers learned how to use the PIAAC data and submitted short proposals on topics they wanted to study. These were reviewed and feedback was provided as needed to support their research efforts. This led to the development of numerous reports that were reviewed and selected to be presented at a 2-day conference.

### Summary

As societies continue to evolve, there has been a growing recognition about the importance of human capital to support social and economic development. And, as the ability to develop and deliver international surveys has evolved and met the information needs of a wide range of countries, there has been a growing interest from middle- and low-income countries in participating in these surveys as well as from international donors who are willing to support them. Following each cycle and round, the OECD meets and talks with a variety of countries and funding agencies to gauge their interest in participating in PIAAC and other international surveys. As such, each round generally includes both returning and new participants.

Not surprisingly, the increased participation in PIAAC and other international surveys has also led to a number of challenges and concerns, which are both technical and political in nature. For example, increased interest leads to the need for systematic efforts that will enhance capabilities to support expanding participation by a wider range of countries. There will also be a desire to broaden the range of constructs that can be assessed as well as the need to accommodate rising costs and the desired use of multiple digital devices. On the political side, there are concerns around the misuse of league tables and desire to frame policy goals in terms of country rankings. There are also challenges associated with how best to disseminate findings from secondary analyses and, in some places, an apparent decline in the trust of evidence and expertise. Collectively, these challenges present healthy tensions that can and should be discussed and debated among policy makers, researchers and other



key stakeholders. Although we are reasonably certain that international large-scale assessments will continue to face these and other challenges, it is also likely that new technical and methodological innovations will follow from new questions and requirements raised by the various

stakeholders. Like the evolution we have seen from the first large-scale assessments of adult skills to the present, we expect that current and future innovations will lead to increased appreciation and impact for PIAAC and other international surveys.

## References

- Kirsch, I. S. (2003). Measuring literacy in IALS: A construct centered approach. *International Journal of Educational Research*, 39, 181-190.
- Kirsch, I. S. (2001). The framework used in developing and interpreting the International Adult Literacy Survey (IALS). *European Journal of Psychology of Education*, 16, 335-361.
- Kirsch, I., Lennon, M.L., Yamamoto, K., & von Davier, M. (2017). Large-scale assessments of adult literacy. In R. Bennett & M. von Davier (Eds.), *Advancing human assessment. Methodology of educational measurement and assessment* (pp. 285-310). Springer, Cham. [https://doi.org/10.1007/978-3-319-58689-2\\_9](https://doi.org/10.1007/978-3-319-58689-2_9)
- Kirsch, I, Sands, A.M., Robbins, S.R., Goodman, M.J., & Tannenbaum, R.J. (2021). *Buttressing the middle: A case for reskilling and upskilling America's middle-skill workers in the 21st century*. Educational Testing Service.
- Maehler, D. B., Konradt, I., Morozova, D., Dickson, J., Braun, R.-K., & Jakowatz, S. (2023). PIAAC Bibliography - 2008-2022. (GESIS Papers, 2023/04). GESIS - Leibniz-Institut für Sozialwissenschaften.
- Organization for Economic Co-operation and Development. (2013). *Technical report of the Survey of Adult Skills (PIAAC)*. OECD Library. [https://www.oecd.org/skills/piaac/\\_Technical%20Report\\_17OCT13.pdf](https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf)