



NATIONAL
ACADEMY
of
EDUCATION

What Use Is Educational Assessment?

The ANNALS of the American Academy of Political and Social Science Volume 683, May 2019

Special Editors: Amy I. Berman, Michael J. Feuer, and James W. Pellegrino

Educational assessments play a large and often questionable role in American teaching, learning, and schooling. We use them to assess learning and achievement, to influence curriculum and instruction, to hold students and their teachers accountable for results, to guide decisions about placement at various levels of education, and to inform cross-national comparisons of educational systems. But are assessments effective in all of these uses? Is it time to take stock of our current assessment regime, and consider the most useful roles of testing and assessment moving forward?

Our political history, growing regard for evidence-based policymaking, and advances in the sciences of learning and measurement put testing and assessment at the center of the current discussion about education policy. **The articles collected in this volume of *The ANNALS* focus on the benefits of appropriately used tests and assessment tools, as well as the risks of inappropriate applications. They reflect on how the sciences of teaching, learning, and cognition, joined with contemporary theory and measurement technologies, might lead to more appropriate and beneficial uses of assessment.**

A guide to the contents:

The **first part of the volume orients readers to the history of assessment and testing in the United States.** One chapter deals largely with K–12 schooling and the other with higher education. These histories show that governments and schools have tried many forms of examination to measure learning, evaluate teachers, and allow/limit admissions and access to education programs. The authors highlight the need to consider the application of educational assessments in the context of political and social developments and in the light of technological advancements.

The **next part of the volume identifies and examines three uses for educational assessments:**

- 1) ***Broad monitoring of the performance of school systems.*** The first chapter in this section describes applications and features of tests designed primarily as tools for monitoring the quality and progress of education systems. One chapter focuses on national large-scale system-monitoring tests, including the National Assessment of Educational Progress (NAEP; the nation's only large-scale test). Such data provide broad-gauge estimates of educational trends, with attention to regional and demographic differences across the nation and over time. The second chapter discusses the history and effectiveness of international large-scale assessments, or ILSAs, which are useful for monitoring national trends in achievement *relative to other countries* and can provide rich information about educational conditions within and across countries; the results are often used to influence policy decisions, despite their limited suitability for such tasks.

- 2) *Judging school quality and creating incentives for change.* This section includes two chapters that explain how assessments are used to measure whether schools are meeting established standards, how assessment might guide school improvements, and the economic implications of academic performance. The authors explain that the effectiveness of testing as a tool for administrative accountability depends not only on the validity, reliability, and comprehensibility of the inferences derived from the results, but also on whether assessments capture schools' contributions to learning; the evidence suggests that properly constructed standardized tests can lead to improved student outcomes and lessen race-based achievement gaps, but that they also pose risks such as narrowing of curricula and too much "teaching to the test." Academic achievement is also a determinant of economic outcomes, and the evidence here shows that test-based school accountability can lead to higher student achievement and to improvements in the teaching force, albeit with risks of unintended negative consequences.
- 3) *Deciding on placement of students or admission to different types of educational institutions.* A chapter on higher education admissions highlights the tension between standardized testing as a source of objective information to predict student performance and the possibility that overreliance on such instruments can perpetuate inequities in access and opportunity. The second chapter focuses on the role that assessments play in special education, showing how they may help to categorize learners but are not useful in identifying effective student support strategies.

The **third section of the volume looks to the future of educational assessment**, taking stock of how the sciences of learning and measurement are evolving. This section makes the case that a productive future for educational assessment relies on aligning advances in measurement science to evolving learning theory. The first chapter summarizes the emerging *situative sociocognitive (SC)* psychological perspective on education, showing how it is forcing a retooling of educational assessments. A second chapter takes that discussion into the classroom, and proposes a teaching/learning model that encompasses assessment while minimizing the potential negative effects of high-stakes testing on learning. A third chapter focuses on building coherent systems of assessments that support equity in science learning. Such systems can be enacted in the context of urban district reforms and are enabled by research-practice partnerships. The fourth chapter makes the case that digital environments could transform the assessment landscape altogether: digital assessments may be able to recreate the attributes of learning environments in ways that increase test fidelity, and enhance data collection in ways that facilitate further assessment development. The final chapter in this section illustrates such an environment by presenting a large-scale, digitally based assessment using simulations that can reliably assess students' competence to plan, execute, and analyze physical science experiments.

The **final section of the volume presents reflections by users of research evidence**. In this "policy forum," a prominent former school superintendent and former federal research official offer cautionary advice—it is clear that test-based information can and should be used, but just as clear that there are important risks of overuse or misuse that need to be taken into account.

Taken as a whole this collection of work argues that no system of measurement is error-free, and that intended benefits should be weighed against the negative consequences of using tests for various purposes. A primary consideration is that how individuals and organizations behave in response to the application of any assessment system could affect the precision, meaning, and usefulness of the data that the assessments evince.