



Notes in Brief

The new data sciences of education bring substantial legal, political, and ethical questions about the management of information about learners. This piece provides a synoptic view of recent scholarly discussion in this domain and calls for a proactive approach to the ethics of learning research.

AUTHOR

Mitchell L. Stevens, Ph.D.
Stanford University

An Ethically Ambitious Higher Education Data Science

The work assembled in this issue leaves little doubt that postsecondary assessment is in a sea change. Digitally mediated instruction provides data whose fidelity to processes of learning are superior to any available to this field in the history of quantitative inquiry. The papers and reviews collected here provide a tantalizing early sense of the scientific promise of these new empirics and a glimpse of their implications for the improvement of higher education.

Yet the opening of a vast new scientific frontier is not the only sea change in postsecondary assessment, or even the most important one. During the same few years that digitally mediated instruction has become a data science, the spiraling cost of attending college in the United States has become a political crisis. During these same few years, the goal of raising stubbornly low rates of college completion has become a major priority for prominent philanthropies. And also during these years, the question of what and how much students actually learn in college has become a major research and policy concern. In sum, the emergence of education data science is simultaneous with an accountability revolution in the postsecondary sector, with many new voices in government and business joining researchers and policy analysts in calls for new means of measuring success in higher education.

Educational measurement is political. It changes the way people make sense of the world and what things count as facts and expertise. It changes relationships between those who produce education, pay for it, and regulate it. It makes educational processes comparable that might long have been regarded as distinct and incommensurate. And it produces information about individuals and groups that can be used by third parties to sell products and distribute fateful opportunities and credentials. This is why the educational data streams now available to scientific inquiry must be considered and managed with thoughtful care.

CORRESPONDENCE

Email
mitchell.stevens@stanford.edu

These are the issues that encouraged some 50 educators, scientists, and legal/ethical scholars to convene at the Asilomar Conference Grounds near Monterey, California, in June 2014. Their task was to specify the ethical challenges and obligations that accompany research on higher education in the era of big data. The convening was modeled after a 1975 event at the same site, during which 140 biologists, lawyers and physicians met to write voluntary guidelines for ensuring the safety of recombinant DNA technology. Another precedent was a 1978 meeting at the Belmont Conference Center in ElkrIDGE, Maryland, which produced a document informing ethical considerations of research with human subjects.

The outcome was a heroically brief document affirming the importance of pursuing education data science for the improvement of higher education in an open, urgent, and ethically considered way. The Asilomar participants concurred that the political implications of measurement in higher education should not inhibit its pursuit, since the prospect of improving higher education with new science was too important a goal to inhibit inquiry.

The Asilomar Convention for Learning Research in Higher Education includes two basic tenets:

Advance the science of learning for the improvement of higher education.

The science of learning can improve higher education and should proceed through open, participatory, and transparent processes of data collection and analysis that provide empirical evidence for knowledge claims.

Share. Maximizing the benefits of learning research requires the sharing of data, discovery, and technology among a community of researchers and educational organizations committed, and accountable to, principles of ethical inquiry held in common.

The Convention additionally specifies six principles to inform decisions about data use and knowledge sharing in the field: *Respect for the rights and dignity of learners; beneficence; justice; openness; the humanity of learning; and continuous consideration* of the ethical dimensions of learning research. The entire document is available at asilomar-highered.info. By way of informing the discussion represented in this issue of *Research & Practice in Assessment*, I add a brief word here about the final principle.

Anyone who pursues education data science quickly learns that there is considerable uncertainty about just how inherited norms and routines for ethical oversight should be applied to data from digitally mediated instruction. IRB protocols that require active consent (rather than a continuous flow of data collection) and prior specification of research questions (rather than iterative inquiry), university proprietary rules that presume data have single owners or trustees (rather than multiple ones), and legal rules applying specifically to students (rather than learners) are but a few features of standard regulatory architecture that fit only awkwardly, if at all, to research with data from digitally mediated instruction. What to do?

One option would be wait until our IRB officers, attorneys, government and foundation officials, and politicians figure out how to rewrite the inherited rules. In light of the inherent complexity of this problem it is unclear just how long that wait might be. A second option is to move forward with research with an explicit commitment to what the Asilomar Convention calls *continuous consideration*. “In a rapidly evolving field there can be no last word on ethical practice” it reads. “Ethically responsible learner research requires ongoing and broadly inclusive discussion of best practices and comparable standards among researchers, learners, and educational institutions.”¹

I believe that the second option is by far the ethically more ambitious one. It recognizes the complexity of the current historical moment while keeping sight of the extraordinary opportunity for new science to improve the quality of instruction and learning in college. It recognizes that ongoing peer review is an essential component of responsible scientific conduct. And it enables us to inform the ongoing development of ethical tradition with the wisdom and caution that comes only with practice.

Moving forward quickly and ambitiously with higher education data science will not be uncontroversial. As this mode of inquiry gains intellectual space and analytic sophistication, it will almost surely direct attention away from currently preponderant modes of measuring value in the sector: persistence and completion rates, accreditation review protocols, rating and ranking schemes, and the myriad social sciences of higher education that have been built with student-level survey and census data. Each of these measurement regimes has partisans and profiteers who will pay attention to any change in what counts as valid and reliable assessment. Add all of this to the more general ethical questions confronting use and integration of big data generally, and we have research frontier whose obstacles are hardly for the faint of heart.

Thankfully the work itself is thrilling and the possibilities for educational improvement profound. Hang on, keep moving, and steer.

The Asilomar Convention for Learning Research in Higher Education... specifies six principles to inform decisions about data use and knowledge sharing in the field: Respect for the rights and dignity of learners; beneficence; justice; openness; the humanity of learning; and continuous consideration of the ethical dimensions of learning research.

Educational measurement is political. It changes the way people make sense of the world and what things count as facts and expertise.

¹ <http://asilomar-highered.info/>