Standardized Student Learning Outcomes Assessments for Higher Education:
Obstacles and Opportunities

Thomas Van Essen

Executive Director

Center for External Research

Educational Testing Service

4/26/2016

William James (1842-1910)

- "We are practical beings, each of us with limited functions and duties to perform. Each is bound to feel intensely the importance of his own duties and the significance of the situations that these call forth. . . .The others are too much absorbed in their own secrets to take an interest in ours. Hence the stupidity and injustice of our opinions, so far as they deal with the significance of alien lives. Hence the falsity of our judgments, so far as they presume to decide in an absolute way on the value of other persons' conditions or ideals."

- " . . . .neither the whole of truth, nor the whole of good, is revealed to any single observer, although each observer gains a partial superiority of insight from the peculiar position in which he stands."

    --"On a Certain Blindness of Human Beings" (1899)

My Perspective

- I've spent a large part of my career worrying about large scale testing programs.

- I've been concerned with how these programs can be fair, valid, and reliable.

- And sustainable over time.

- I was involved in the AHELO project.

Two Recent Articles

- Paul Ashwin (2015) "Missionary zeal: some problems with the rhetoric, vision and approach of the AHELO project," European Journal of Higher Education, published online 5 October 2015

- Dirk van Damme (2015) "Global higher education in need of more and better learning metrics. Why OECD's AHELO project might help fill the gap," *European Journal of Higher Education,* published online 22 October 2015

AHELO

- A feasibility study carried out by the OECD to determine if it is "practically and scientifically feasible to assess what students in higher education **know and can do upon** graduation."

- Planning began in 2008

- Field work and reporting in 2011 and 2012

- A top down approach in which the OECD worked with ministries of education to fund the project

- Economics

- Engineering

- Generic Skills
  - Collegiate Learning Assessment (CLA) developed by the Council for Aid to Education (CAE)

Measuring the Power of Learning.™

AHELO Status

- The OECD tried to move to a second phase, with only generic skills (the CLA) in 2014-2015.

- The feasibility study did not convince countries that this approach was going to work.
  - There was a great deal of opposition from colleges and universities, especially in the United States.

- The project is currently on hold.

van Damme

- An argument for "an international assessment of learning outcomes."
  - Counter arguments boil down to "it can't be done" or "it shouldn't be done."
  - Focus of the paper is on why it should be done.

- Higher education is a "quasi-market" that is a cross between a "public good" model and a "private consumption model" and, as such, is characterized by "information asymmetries."
  - Students value higher education, but don't know what they are purchasing in the higher education market.
  - Hence the rise of rankings, which we know are flawed.

- Skills are more and more important
  - Employers are starting to question that university graduates have the requisite skills.

van Damme (continued)

- A number of national initiatives have been launched
  - CLA+, HEIghten (ETS), KoKoHS, a UK project on learning gains
- But because of globalization an international approach is required.
- Higher Education is very diverse, but there are some common attributes associated with a university education so we should focus, as PISA does, on generic skills as a way to at least partially address the "information gap."
- "It should be done. . . .the feasibility question is ultimately a matter of will or expertise. . . .Perfect solutions are very difficult to achieve, but good solutions are feasible. And good solutions could bring us already some answers to the information gaps in the system. Doing nothing is not much of an alternative."

It's a good argument . . .

- But it is an argument from the perspective of *policymakers*.
  - *And the OECD*

There are a number of other actors in this space, and they see things from a very different perspective.

Ashwin

- Explicitly a critique of the AHELO project
- Asserts that "there is a genuine need for quality measures that can help to ensure students have equitable access to high-quality higher education wherever they study."
- Such measures should not, like most of the current ranking schemes, privilege historical reputation.

Ashwin (continued)

- Identifies three problems with AHELO:
    - The arguments for AHELO (as advanced by the OECD) are based on expedience rather than principle.
    - AHELO assumes that higher education is about the development of generic skills.
        - And generic skills are problematic in that "problem solving" in the disciplines is different than generalized "problem solving."
        - These constructs are very dependent on familiarity with the kind of test that measures them.
    - Negative wash back: faculty will teach to the test and there will be a narrowing of the curriculum.
    - Conclusion: there cannot be an assessment that unproblematically tells us what students have gained from higher education.

It's a good argument . . .

- But it is an argument from the perspective of a university researcher focused on teaching and learning in the university.

- There are a number of other actors in this space, and they see things from a very different perspective.

- My point is that the solutions proposed depend very much on the perspective of the individual making the argument.

- It is not that they are wrong; it's just that they are limited.

Problem and Solution

- If the problem we want to solve is how can we get valid, reliable, and actionable information about higher education outcomes, we need to take into consideration the often competing perspectives of multiple stakeholders.

Measuring the Power of Learning.™

Stakeholders

- Perspectives I want to consider
  - Policymakers
  - Administrators within institutions of higher education
  - University faculty
  - Students
  - Parents
  - Employers
  - Test makers
    - These are all artificial groups

- What kind of data and information do these various stakeholders want? What kind of actions would they like to enable?

- What are these stakeholders afraid of?

Policymakers

- State and Federal Officials
  - NGOs like the OECD
  - Private foundations
- Data to inform decisions about policy related to the
  - Distribution and development of human capital
  - Optimization of educational resources from a societal perspective
    - Which are essentially political decisions
    - And will certainly be perceived as political
- Money
  - They are, for better or worse, the stakeholders who have the money
    - Efficient
    - Cheap
  - There is clearly a tension between the desire for the very best possible data and the desire not to spend lot of money

Administrators Within Institutions of Higher Education

- But a caveat: not all universities are alike
  - Probably better to think in terms of kinds of universities

- Interested in
  - benchmark data, for improvement or bragging rights
    - But for certain institutions the status quo is just fine
  - data that will inform internal resource allocation discussions
  - data that will support the institution in achieving its mission
    - i.e. not (necessarily) in data that will drive the institution to be like other institutions

- Fearful of
  - Additional costs
  - Administrative burdens
  - Having to fight battles with the faculty
  - Additional oversight from state or federal policymakers
  - Reductive approaches that ignore the complexity of specific institutions

University Faculty

- To speak of university faculty as a single category is probably silly.

- Faculty value their discipline and their intellectual independence.

- Faculty are loyal to their institution, but for many of them their true loyalty is to their discipline.

- Faculty are fearful of any assessment scheme that tells them what to teach and how to teach it.

  - They are particularly concerned about assessment schemes that will over-simplify domain and reduce it to a series of facts.

Students

- This is in many respects the most problematic group; they represent the biggest barrier to the successful implementation of standardized measures of learning outcomes.

- From a student perspective the quality of their education should be the highest concern.
  - (A question for those of you who were once students: do you remember that as high on your list of concerns?)

- But it is the quality of their education now—not the quality that their institution provides in years to come—that is their concern.

Measuring the Power of Learning.™

Students (continued)

- If outcomes are to be measured by an assessment, students have to suffer through it.

- Student motivation is the biggest impediment to high quality measures of learning outcomes.

- Motivation depends on culture, institution, individual, and task.

- What can be done to deal with low motivation?
  - Technical solutions which screen out unmotivated test takers
  - Make the task interesting
  - Make the task short
  - Make the assessment results useful to the test takers
    - Typically implies longer test length
  - Make the task consequential
    - Can be problematic
  - Provide incentives
    - Will be expensive and, perhaps, tend to invalidate results

Parents

- To the extent that parents pay for and are interested in the outcomes of their children's education, they may have an interest in the value and efficiency of that education.

Employers

- To the extent that college and university credentials are no longer valued or trusted, employers could be interested in a supplemental credential (a certificate or badge) that certifies some specific competencies, either generic (problem solving, teamwork) or subject specific (principles of accounting).

- But a "market" for these badges or certificates would have to develop
  - And if it did the student motivation problem might disappear.

- But this could be a highly problematic development from the perspective of the universities
  - Universities are something more than trade schools or the training department of large corporations.

Measuring the Power of Learning.™

Assessment Designers and Providers

- Have an obligation to provide technically sound assessments that are appropriate for the uses that have been agreed to with the resources that are available.

- They are presented with a problem of design within constraints
  - Time
  - Money
  - Consequences

Time

- Testing Time
  - Generally, the longer the better
    - But very few of the stakeholders want to give up the time
  - Higher stakes = longer testing time
  - Often a "political" question

- Development Time
  - Generally,
    - the more innovative
    - the more original
    - the more complicated
  - The longer it takes to build
  - The more it costs

Money

- In practical terms, money is usually the ultimate constraint.

- The normal process:
  - We try to imagine what we want, and then we compromise and try to get to what we can afford.

    or

  - We know how much we have and maximize within that budget.

- We need to have the courage to not do something if all we can afford is inadequate in terms of certain key stakeholders.
  - Probably not worth doing if we can't do it well.

Is It Possible to Find a Solution That Will Look Good to Every Perspective?

- Probably not

- But . . .

- Important to be clear and transparent about who is getting short-changed and why.

- Important to be clear and transparent and specific about who will get what benefits from the assessment.

Lessons from AHELO

- Certainly possible from a technical perspective
  - Delivery, scoring, etc.
- AHELO was not successful because it was a vision that only satisfied one stakeholder—the policymakers who wanted something cheap.
- Doing this well will take (testing) time and money.
- Buy-in from faculty members is the most critical and most important hurdle.
  - If they are not convinced, the effort is doomed to fail.
  - Any assessment scheme needs to support the work of faculty.
- The problem of student participation needs to be solved.
- One size does *not* fit all.
- We need a well-articulated theory of action for any such scheme.

A Possible Path Forward?

- Measuring and Comparing Achievements of Learning Outcomes in Higher Education in Europe (CALOHEE)
- An Initiative of the International Tuning Academy (TUNING)

  - European Commission funded initiative
    - Focused on Europe
  - Based on the work of Tuning Initiative
  - A "bottom up" approach
  - Support of disciplinary associations
  - Support of various European higher education associations

A Possible Path Forward?

- Will attempt to identify the key learning outcomes for five disciplines

    - History
    - Nursing
    - Education, including Teacher Training
    - Civil Engineering
    - Physics

CALOHEE Challenges

- Can the faculty agree on the most important learning outcomes?
  - And can these learning outcomes be measured in a economically feasible way?
  - Will it be possible to create a sustainable assessment program?
- Can they agree on the precise purpose of the assessment?
  - Formative, for program improvement?
  - Benchmarking?
  - Extra curricula certification of competencies?
- Will students participate?

Conclusion

- So what should be the perspective of researchers in this field?

- Researchers need to be open, respectful, and attentive to the perspective of various stakeholders.

- Researchers need to have a clear "theory of action" that specifies the mechanism by which the assessment will produce specific desirable outcomes *and* avoid unintended negative consequences for stakeholders.

- And researchers have to be able to say "no" if they can't come up with a satisfactory solution.

Vielen Dank!

- Questions? Comments?

Contact:

Thomas Van Essen

tvanessen@ets.org

Measuring the Power of Learning.™