

NAEPing Higher Education? Decidedly Not!

Clifford Adelman, Senior Associate, Institute for Higher Education Policy¹

To date, NAEP has conducted studies keyed to 12th-grade preparation for higher education and other postsecondary dependent variable configurations, and with occasional probes of student performance in first year college courses. As I understand it, the proposal or question on the table pushes NAEP further into higher education, all the way to the receipt of a bachelor's degree, i.e. that NAGB wants to make quantitative statements on the summative cognitive attainments of undergraduate students no matter how old they were when they entered higher education, whether they were from non-English-dominant households (in a country in which 170 languages are spoken), what types of institutions or combinations of institutions they attended and where, what other life events or settings (employment, military service, parenting) they experienced as adults (which they all are), which of 4000 major fields they focused their education (including those who experienced depth study of 2 or 3 fields) no matter how disparate those fields are, and on and on. This is a far more complex territory—and a full field of intervening variables—than pre-collegiate schooling. Even if I supported NAGB's intention to offer yet another set of metrics demonstrating how dumb college graduates are and why they never should have been awarded degrees (whether one of the myriad types of bachelor's degrees and/or one of the myriad types of associate's degrees available), I would not enter this swamp under any conditions. But I do not endorse the intention. Why?

1) Let's start with the rest of the world, much of which has developed Qualification Frameworks that spell out degree-culminating proficiencies that

¹Formerly Senior Research Analyst, U.S. Department of Education, where he built transcript-grounded data bases for three national grade cohort longitudinal studies, and published a dozen research monographs based on those data sets, the best known of which were *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment* (1999) and *The Toolbox Revisited: Paths to Degree Completion from High School Through College* (2006). Since leaving the Department in 2006, he has published three studies of Europe's Bologna Process (*The Bologna Club* [2008], *Learning Accountability from Bologna* [2008], and the book-length *The Bologna Process for U.S. Eyes: Re-learning Higher Education in the Age of Convergence* [2009]), along with *The Spaces Between Numbers: Getting International Data on Higher Education Straight* (2009), and *Searching for Our Lost Associate's Degrees* (2013). He is one of four co-authors of *The Degree Qualifications Profile* (2011, 2014), and has been a member of a dozen Technical Review Panels for NCES and the European Commission.

students must demonstrate. There is a range of levels of education addressed in non-US qualifications frameworks, a range of specificity in their proficiency statements, a range in strength of qualifications criteria. That is, the proficiencies are stated as facts to guide instruction and curricular provision, or requirements else degrees will not be awarded, or wish lists. That is, declarative, imperative, and subjunctive voices, if you remember your basic English grammar.²

1.1) We see Qualification Frameworks for higher education in the 47 nations participating in the Bologna Process, from Cork to Vladivostok. We see them in Australia, South Africa, Ontario. And we see them in the U.S. under the moniker of the Degree Qualifications Profile, first issued in January 2011 in an iterative process, and, following exploratory work by roughly 400 institutions of higher education in the U.S., a 2.0 version posted on-line at the Lumina Foundation Web site in January, 2014, with a printed version expected in September. For those of you who know little or nothing about the DQP, log on to

luminafoundation.org/dqp

and read version 2.0 slowly—all of it, including the appendices. It covers 24 qualifying proficiencies at the associate’s level, 28 at the bachelor’s, and 20 at the master’s, all driven by operational verbs that get ratcheted up in terms of challenge as one moves up the degree ladder.

1.2) In none of the QFs in other countries does assessment or testing play a prominent role. It’s almost an afterthought. In the DQP, assessment is replaced in function by “assignments,” logical extensions of verb-driven proficiency statements. Faculty give assignments every week, but in the DQP vision, they are tweaked to match the proficiencies. These assignments are a central, necessary, and constitutive factor in the DQP. External examinations are specifically excluded from this universe on the grounds that they do not belong to the instructional workforce therefore are outside the chain of learning. They are not mentioned by name, but the CLA is certainly at the top of the list of exclusions. We will come back to this.

2) The world has also engaged in a process called Tuning, under which faculty teams in specific disciplines first develop “reference points” of content in their

²This observation is elaborated in Adelman, C., “Use and Problems in the Language of Discipline-based Qualification Statements: Tuning and Its Analogues,” under journal review.

disciplines, then student learning outcomes pegged to those reference points. In the process, both employers and recent graduates in those disciplinary fields are active contributors to the shaping of a Tuning statement.

(2.1) Tuning started in Europe just after the signing of the Bologna Process with 9 disciplines in 15 countries, and now stands at something like 28 disciplines in 37 countries. It came to Latin America in 2005, with 12 disciplines, 188 universities in 18 countries. It came to the U.S. with Tuning USA in 2009, and since then, we've seen 6 state higher education systems (Indiana, Kentucky, Minnesota, Montana, Texas, and Utah), one interstate compact (the Midwest Higher Education Consortium), and one national scholarly association, the American Historical Association, develop Tuning templates. Australia ran a pilot in 2012-2013; China a test in 2013; and we've now got Tuning enterprises in Africa (organized by region), 4 countries in Central Asia, and, most recently a 12 university consortium Tuning project in Japan. There has to be something compelling going on here, or one would not see the breadth of this activity on 6 continents.

In all discipline-oriented Tuning projects (parallel to the more generic QFs and the DQP), assessment is relegated to the disciplinary faculty responsible for executing the template in local contexts. What Tuning yields is a convergence of coverage and delivery, not standardization. The business programs at Fribourg in Germany and Coimbra in Portugal share the same reference points, e.g. the firm as a value-chain from procurement to customer service, but not necessarily the same learning outcomes or weighting of the components of that value-chain. Everyone at least sings in the same key, but not the same tune. That's a big achievement with considerable import for cross-border student mobility (and, in the U.S., where community colleges were included in every project, for vertical mobility as well).

If you want to see how well we, in the U.S., wrote learning outcome/competency/proficiency statements under Tuning, I recommend the Texas HECB board production in 4 engineering fields³, the Midwest Higher

³The easiest way to do this is to Google "Texas Tuning," and, at the lead page, scroll down to the engineering fields list and pick whichever branch of engineering in which you are interested, and click for the pdf report. For lay readers, Civil Engineering is probably the most accessible. THECB chose to focus on the transfer function and course proxies, which is not the way other Tuning USA projects worked, but for our purposes the way they wrote learning outcome statements is the point.

Education Consortium's template for Marketing⁴, and the AHA's template as published in the November 2013 issue of their bi-monthly, *Perspectives on History*, pp. 21-23. Dozens of colleges and community colleges participated in these undertakings, demonstrating that, contrary to the going mythology, higher education is not asleep in setting consequential benchmarks for student learning.

It should be noted that Tuning's greatest challenge, no matter where it is located and no matter what disciplines are involved, is critical mass. The question always arises as to how many faculty in how many departments in how many universities have either participated in, endorsed, or carried out a disciplinary template composed of reference points and learning outcomes. The Europeans were able to capitalize on what the European Commission calls "Thematic Networks" to expand participation,⁵ and there is no doubt that the Thematic Networks have had multiplier effects. But the challenge of participation remains.

3) Both our native DQP and Tuning USA were sponsored by the Lumina Foundation for Education, and with no government involvement whatsoever. That was not true in other countries. Lumina got into the game as a bank-shot from my research and analysis of the Bologna Process (which Lumina sponsored—and if you want to learn about Bologna and the epiphanies it generates for U.S. higher education, download

ihp.org/assets/files/EYESFINAL.pdf,

and under the mantra we inherited from macroeconomic historians: nations that learn from other nations grow; those that don't learn, don't grow. The fact that four regional accrediting bodies, three national higher education organizations, and one state system have worked deeply with the DQP to date (all under Lumina sponsorship), along with all the participants in Tuning USA, says something about how widespread and serious these student-centered learning outcome projects are. Again, though, the only one of these to deal in depth with assessment has been the DQP, and the vehicle, in that case, is local ownership of assignments. DQP 2.0

⁴The MHEC Tuning report for Marketing is a fold out glossy that cannot really be downloaded. To get a copy, log onto mhec.org/programs/tuning, then "contact us" and submit a request for the Marketing document.

⁵The best example is probably the European Chemistry Thematic Network, with membership of 120 universities in 30 countries. To see who is involved and how, log onto ectn-assoc.cpe.fr/network/.

includes examples of such assignments designed to elicit student behaviors to allow faculty judgment of whether generic proficiencies have been achieved, and the project is also developing a library of such assignment examples through a charette method. How large a portfolio will be assembled is an open question, but as soon as faculty see these concrete examples in the hammocks of proficiency statements, they have an “Ah-hah! So that’s what they mean!” moment.

As for faculty judgment of how well students respond to these assignment prods, the DQP rests the authority of judgment with faculty. At the same time, higher education has introduced rubrics methods and benchmarks through the work of the Association of American Colleges and Universities. The DQP takes no stand on rubrics.

Standardized testing, particularly that coming in from parties that have nothing to do with instruction or with integrating prods with curriculum is vigorously objected to in this undertaking—and justly so. After all, too, the DQP proficiency statements include blocks for ethical reasoning, differential perspective, use of information resources, and learning applications including collaborative variations, and there is no one examination that has ever addressed such a portfolio—nor could there be. Degrees are complex phenomena; high school diplomas are not.

4) Typically, OECD tried to get into this act with its cross-border and multi-language AHELO project, half of which was a version of the CLA and the other half based on Tuning templates in economics and engineering. With a tortuous history inside OECD since 2003, AHELO finally saw a test run last year, which the technical advisory committee has judged to be a large disappointment. The CLA ate up 80% of the budget for this project, and could not produce acceptable cross-language prods or cross-national scoring. The Tuning-based assessments were judged a bit more acceptable. Where AHELO goes from here is anyone’s guess, but the U.S. will not participate.

5) Now, I will put good money on the table that the vast majority of people in this room have, at best, but a peripheral consciousness of any of this—not Bologna, not NQFs, not the DQP, not Tuning, not AHELO. I will also put good money on the table that the vast majority of people in this room would have difficulty articulating what a “competence” is, or why the DQP dropped the term in favor of the summative judgment of “proficiency,” and that the governors of the NAEP have never looked at the products of OECD’s DeSeCo project that form probably the

best set of considerations for those entranced by “competencies.” Your ignorance is not wrong; it’s just a fact, but if one is to explore the potential of externally-generated and imposed assessments in higher education, all of this knowledge is indispensable.

6) What we have seen of standardized testing in higher education has nothing to do with improving the enterprise, and results in information best described as waste. To follow the ways in which the CLA is used, to say that my effect size is bigger than your effect size, particularly when you pay student volunteers to take the examinations in question (something all research has shown to have unreliable effects), has no specificity, no impact on either instruction or student learning, and no use except in the numbers pundits and half-conscious state legislators like to quote, even though they have no idea of what those numbers represent. To present the results of such examinations claiming they were based on a representative sample of 200 of Arizona State University’s 51,000 students requires statistical gymnastics worthy of an ignoble Olympic medal. The assignments construct of the DQP, by contrast, affects *all* students (not just sample) and *all* faculty responsible for courses and learning experiences in which proficiencies are embedded. The minute a government assessment —particularly when what the most the federal government understands about higher education are graduation rates and financial aid—crosses the threshold of higher education, all DQP-related efforts to clarify the meaning of degrees, reconstitute curriculum mapping, and placing the assignment-variation at the core of the enterprise will vanish. You will have killed the best chances we have ever had for student-centered transformational higher education reform. The shadow of a government test---and that is what NAEP is---inserting itself in non-compulsory education where a messy mass of adults seek learning and advancement at different stages of life, in different types and combinations of institutions, on different time schedules, at different levels of credentials, and in over 4000 different fields and disciplines, is enough to freeze the soul, let alone elementary statistical sanity. Stay out of it!