



Assessment Comes Around Again

**The challenge of documenting
student learning**

Rules of Engagement

- **If you have a question, interrupt at any time**
- **But please save the speeches for our interaction period**
- **You have some handouts to which I will refer, but which are for illustrative purposes in the story line---and later reading**
- **This collection of slides will be on the CAIR Web site.**
 - Cliff Adelman, Senior Associate,
Institute for Higher Education Policy**

What we're going to do today

- **Review the provenance and short history of the assessment movement in U.S. higher education**
- **Ask what assessment means and where it fits in current debates about accountability**
- **Bullet potential sources of information**
- **Consider some alternatives to what the Spellings Commission suggests we do**



This is not a data-essay

It's a form of "public writing"

The GRE History Area Test Question on the Battle of Lepanto

Which of the following was not a consequence of the Battle. . .

- a) The Turkish fleet ceased to be a threat to the Mediterranean Christian states**
- b) Plans developed by The League to invade the Turkish Empire became more plausible**
- c) The Turkish Empire was so weakened as to put it on the road to collapse**
- d) Admiral Don John of Austria became a great hero among both sailors and emissaries to The League**

Historical markers

- **Competency-based experimental degrees of the 1970s**
- ***Careering After College*, the grounds of the Alverno model (1977-1983)**
- ***Involvement in Learning*, report of the last ED “commission” (1984)**
- **Performance and portfolios: the early years of the AAHE Assessment Forum (1987-1992)**
- **Hijacked by TQM: the middle years of the AAHE Assessment Forum (1993-1998)**
- **Assessment disappears: replaced by GRS**

Filling in between the markers

- **The ACGE (grandmother of the CLA) and its mass-AASCU try-out (1975-80)**
- **Value-added, its testing vehicles (COMP), performance funding in Tennessee, and the total-assessment university (N.E. Missouri), 1980-1986**
- ***The Standardized Test Scores of College Graduates, 1964-1982 (1985)***
- **High Stakes: Ability-to-Benefit 1989-95**
- **Early NPEC: exploration of a national assessment (1992-1994)**

And along the way, the literature explored

- **External examiner models**
- **Model indicators of summative learning in the major**
- **The validity of student self-assessment**
- **Classic psychometric questions, e.g. cut scores, in new contexts**
- **Experimental measures for the study of creativity**
- **Uses of technology in testing**

And along the way, the states got into the act

- **Tennessee, 1979—COMP required**
- **Florida, 1982---CLAST, a high stakes**
- **Colorado, 1985---legislative mandate**
- **Missouri, 1987---canonizing NE Mo**
- **New Jersey, 1988—college academic outcome, a homegrown**
- **Virginia, 1987---SCHEV mandate**

Session topics in first 6 AAHE Assessment Forums

Topic	#	Trend
In the discipline/major	33	Accel from yr 3
General education	30	Fade from year 5
Student development	26	Start in year 2
Methodology/Technical	23	Plateau in year 5
Classroom-based	22	Accel from yr 3
External accountability	16	Start in year 2
Institutional	15	Partner with QI
Quality improvement	14	Start in year 4
Value added	8	Fade from year 2

Where were we by the early 1990s?

- **Confused about the difference between assessment of student learning and institutional performance**
- **Mixing up assessment, testing, and evaluation**
- **Dealing with competing claims of a raft of commercial testing products (over 400 in the ETS annotated bibliography)**
- **Located principally in 2nd and 3rd rank institutions**

Avoidance behavior

- It became a hallmark of the assessment movement to avoid the tension inherent in the judgment of individuals and full census reporting
- Instead, it embraced both the institution or the program as subject, and samples of performers representing the subject

In an age of accountability, what kind of problems does this preference raise?

Bright spots of learning and their limitations

- Have you ever read an Alverno transcript? It *is* about student learning
- The ACGE: the virtues and difficulties of constructed response
- **Performance assessment: the problem of task sampling variability**
- **Assessment as learning & the self-reflective instructor: heavy on form, light on content (exception: *The Journal of Engineering Education*)**

We were trying!

But then were overtaken by external currents . . .

- **TQM, CQI, Baldrige Awards, and the culture of organizational performance**
 - **SRK and the totem of graduation rates**
 - **National Education Goals Panel**
 - **The NGA's Education Summit and the momentum of "America 2000"**
- . . .and backed off. Assessment became an accountability movement, not a normative one.

Accountable v. normative: GRE content representativeness

- **Current curriculum v. Ideal curriculum v. tested curriculum in computer sci**
- **Software systems and methodology**
- **Computer organization and architecture**
- **Theory**
- **Computational mathematics**
- **Special topics, e.g. AI, graphics, data communication**

Fast forward to the Spellings Commission and its discontents

- **Complains college graduates are illiterate, and cites NAAL data**
- **Cites second-hand reports of employer complaints about communication and problem-solving skills of recent college grad hires**
- **Cites complaints of *Measuring Up* that states have no systematic warrantee of the learning of college graduates**
- **So, recommends use of NAAL, CLA, NAEP and whatever else crossed the radar screen to at least provide value-added measures**

Slouching toward the Spellings Commission: the lead-ins, #1

- ***Measuring Up on College-Level Learning* (2005), a.k.a the battle of the states, with an index composed of:**
 - **Statewide NAAL 25%**
 - **Licensure/teacher certification pass rates plus “nationally competitive scores” on GRE/GMAT etc. 25%**
 - **CLA for a sample of 4-yr students and Work Keys for a sample of 2-yr students 50%**
- This one wins the statistical gymnastics prize!**

Slouching. . . #2

- National Survey of American College Students (Jan., 2006), using NAAL on graduating 4yr and 2yr students, found:
 - Both had higher scores than all adults
 - Higher prose and document literacy scores than adults with similar education
 - 4-yr scored higher than 2-yr across the board
 - No differences by 4-yr type or selectivity
 - Standard differences by family income and parental education
- So what else is new?

Pause: The NAAL has been rendered a core benchmark. So what's in it?

- **Prose literacy, e.g. interpretation of brochures**
- **Document literacy, e.g. filling out a job application**
- **Quantitative literacy, e.g. completing an order form**

In other words, life situation tasks in which general learned abilities are applied.

To what extent is this a valid measure of college student learning?

Slouching toward. . . #3

- ***A Culture of Evidence: Postsecondary Assessment and Learning Outcomes* (ETS, June 2006) proposes 4 dimensions of college assessment:**
 - 1) **Workplace readiness and general skills, i.e. what the CLA et al measure**
 - 2) **Domain-specific knowledge & skills**
 - 3) **“Soft skills” (team work, creativity)**
 - 4) **Student engagement with learning****and smartly does not include NAAL.**

Our New Romance: The CLA, Part I

- **Constructed responses to more complex prompts than ACGE or COMP**
- **More sustained time-on-task than its predecessors**
- **Part grounded in the GRE essay section: make/break an argument, computer scored**
- **Part grounded in the performance section of the typical bar exam: integrate information from diverse sources; prepare a memo analyzing problem; faculty team-trained scoring**
- **The provenance, on both groundings, is persuasive**

The CLA, Part 2

- **Is it a good test? For what it does, yes.**
- **Does it measure what college graduates learn? No, and it doesn't claim any more than reasoning & writing skills.**
- **No retired items and scoring criteria yet, so we have to withhold judgment on technicals**
- **Is it designed for individual and full census assessment? No, like its predecessors, it is for institutions and samples.**

The CLA, Part 3

- **When you have volunteers, you don't have high stakes**
- **An assessment with no “incentives to students to participate meaningfully” risks threats to its validity (ETS 2006)**
- **Even \$25 is not an “incentive to participate meaningfully”**
- **The CLA recommended design is not unique in this regard**

The CLA, Part 4: Value-Added is Back!

- **Test 100 freshmen, 100 seniors**
- **By one formula, just control for SAT/ACT scores, and you have it, right?**
- **ACT suggested a similar approach, the concordance methodology, with COMP**
- **With enough institutions participating, peers can compete: “We add more value than you do!”**

Pardon my skepticism, but what would you rather do:

- Offer a criterion-referenced statement of performance for 100% of your graduating students (or even a formative statement for 100%) or
- A value-added domain statement for 100 of your students? Even 3 value-added domain statements by matrix sampling of 150?
- Which one communicates more transparently to governance authorities?
- Which can be better integrated into other institutional analytical and planning frameworks?

Examples of criterion-referenced statements of summative learning

- **93% of our chemistry graduates identified a ferro-liquid utilizing X, Y, and Z in a one-hour performance lab**
- **81% of our history graduates assembled sufficient archival information to build a schematic of corporate relationships in the New Haven Railroad bankruptcy of 1908**
- **89% of our AAS degree recipients in Allied Health/Medical Tech solved 20 simulated tasks using the *Physician's Desk Reference***

Do we need a test? Consider unobtrusive transcript data

- **For writing attainment: 66% of our graduates completed a writing course beyond English Comp (technical, creative, journalism, writing for media)**
- **For quantitative literacy: 73% of our graduates completed *more than one* course in college-level math**

And we have something to learn from the new European Diploma Supplements

Bullets for a Portuguese student completing a degree in environmental design:

- 1) Passed certification exam in computer graphics
- 2) Wrote paper for university facilities planning committee
- 3) 1 term at Univ of Karlsruhe; German assessed at 3rd Stufe
- 4) Team project (nesting behavior in public parks) in Ethology written up in local newspaper
- 5) Short description of final project on design of public plazas

The Diploma Supplement can be a portfolio statement

- **It's about individual attainment**
- **The discrete portfolio statements can be aggregated by program**
- **There is nothing “voluntary” about it**
- **The documentation is produced in the natural course of a student's academic career**
- **It is subsequently combined with a traditional c.v. and a “language portfolio” on an electronic Europass, a pathway to employers on a borderless continent**

A sidebar on Diploma Supplements

- They provide the “what” of student learning, but not necessarily the “how well”
- To date, they have been imperfectly utilized by European universities adjusting to the Bologna Process reforms
- They also include program qualification descriptions and transcribed data, but the latter are less likely to be read

We've covered a lot of territory; it's time to call some questions

- How compatible are assessment and contemporary accountability demands?
- Do criterion referenced performance statements have a place in accountability frames?
- How much do you trust unobtrusive transcript data versus external exams?
- Is there a place for Diploma Supplements in the U.S. scheme of things?

And when we answer these questions, remember:

- **Assessments roll along in the economy and society beyond higher education.**
- **Judgments of quality performance will continue to be passed on individuals by an armada of licensing authorities, funding agencies, and employers.**
- **We can contribute to improving those judgments or wait for the armada to find us. .**
- **The rest, as they say, will be history.**