Accountability, Learning Assessment and the Future of Higher Education

Geri H. Malandra
Vice Chancellor for Strategic Management
The University of Texas System

7th Annual Assessment Conference
Texas A&M University
February 23, 2007

Introduction

When I had the opportunity to speak with Texas institutional researchers in College Station three years ago, I predicted that accountability was going to be with us for some time to come. I also challenged everyone to take the initiative to create, redesign, and demonstrate the use of accountability ourselves – rather than wait for it to be done to us. But, I'm not sure anyone could have predicted the headlines we've been seeing lately:

> "Should Government Take a Yardstick to College?"
> "U.S. in a Testing Frenzy"
> "Colleges Enter Testing Debate"
> "A Warning on Measuring Learning Outcomes"
> "Governor Calls for Exit Testing"

And, now, perhaps most alarming given your speaker this afternoon,

> "UT System Applauds Standardized Testing"

What is going on?

It used to be that it was hard to get anyone to notice, let alone take the time to understand accountability and assessment. Now it seems to be the "new new thing." Are you feeling, like me, that there might be just a little TOO much love out there these days? Is anyone else tempted to sing along with the Animals: "I'm just a soul whose intentions are good. O Lord, please don't let me be misunderstood?"

As assessment professionals, you are on the front lines – there is a lot of ferment in our field right now and, I hope, forward movement. So, I am very grateful to Paul Meyer, Lorraine Philips, and the conference planning committee for this opportunity to reflect with you on:

- Some of the challenging policy questions about assessment and testing;
- Some key directions we are heading in nationally and at the state level; and
- Conclusions and – risky as it may be – a few new predictions for the future.

My focus will be primarily on policy issues and trends more than on the specifics of the kinds of assessment plans you each develop and implement.

Policy Questions

Across Texas and the nation, every institution that seeks regional accreditation has been required for some time to develop methods to assess and report student learning outcomes. Higher education policy groups and colleges and universities – including all of yours – have developed a number of significant instruments and initiatives to help respond to these requirements. These efforts vary considerably from institution to institution and among the various regional accrediting associations.

Recently, the national dialogue on the future of higher education initiated by U.S. Department of Education Secretary Margaret Spellings, has focused national attention on learning outcomes. The Commission report, *A Test of Leadership: Charting the Future of U.S. Higher Education*,[1] calls for a much stronger emphasis on assessing student learning, and higher education associations, systems, institutions, and others have already begun to anticipate and develop a response.

These activities raise important policy questions:

First, why is it – despite the mountains of evidence, including conferences like this one and 20 years of documentation through accreditation reviews – that the academy's ability to measure and demonstrate its outcomes is still being called into question?

Second, assuming a different focus on accountability is needed, what should the units of analysis be? Student? Program? Institution? State? Country? What is it we don't know? Is it the process, or the results, or the use of the information that needs improvement? Isn't it enough if we can show through census data, for example, that a B.A. confers nearly double the earning power of a H.S. diploma?

(When our daughter was in kindergarten, at our first parent-teacher conference we were told that Emily had achieved a "comfortable place for herself in her class." I can tell you, that was not enough information for us – and I wasn't even into accountability back then. When we pushed, the teacher finally gave in and told us she ate slower than the rest of the little five-year-olds. We definitely wanted a different kind of metric analysis!)

Third, what should be the role of testing in assessing learning outcomes within this new accountability framework? Should tests at the postsecondary level ever be mandated? How should they be designed and implemented? How should the results be used? The idea of testing, *per se*, is hard to defend as aberrant. We test all of the time. We're naturally competitive – and anyone with pets and kids knows this is not just a human trait. The ancient Greeks had the Marathon. The ancient Chinese had civil service exams that were, indeed, high stakes for government careers. In the United States, testing has become a kind of fetish. Perhaps some of you remember the *New Yorker* cartoon, where someone hoping to pass through the doors of heaven exclaims to Saint Peter, "No kidding! You count SATs?"

Fourth, how can and should the decentralized and highly segmented "system" of postsecondary education in this country be affirmed – a system that may be regulated but is not run by state or federal governments? According to David Brooks, the essential quality of this "system" is its

---

[1] U.S. Department of Education, Washington, D.C., 2006. http://www.ed.gov/about/bdscomm/list/hiedfuture/index.html.

inherent competitiveness, inherited through its "expansionist genes."[2] Higher education has never been state-dominated or state-run, as is the case in most of the rest of the world. "The competitive American universities not only became the best in the world – 8 out of the 10 top universities are American – they also remained ambitious and dynamic. They are much more responsive to community needs." Our strength is that the state has supported autonomy and private creativity.

Fifth, and most critical, once these questions are resolved and changes have been made, how will higher education be different – for students, faculty, graduates, governing boards, policy makers, employers, citizens?

There are no simple, final answers, but I invite you to consider these questions in the context of work on accountability, accreditation, assessment, and testing at the national and state levels that I'd like to discuss now.

<u>National Trends</u>

The Spellings Commission on the Future of Higher Education recommended a broad series of goals on student preparation and access, cost and financial aid, transparency and accountability, innovation, life-long learning, and global leadership. These are now being addressed through a 30-item action agenda in which accreditation and assessment loom large.

If we drill down in the Commission report to the statements on learning assessment, it recommended to:

- Measure learning with quality-assessment data from tests like the Collegiate Learning Assessment and Measure of Academic Proficiency and Progress;
- Provide incentives to states and others for interoperable outcomes-focused accountability systems – Texas is one of the states that has a head start on this through the data sets collected by the Texas Higher Education Coordinating Board, that align with IPEDS data;
- Collect data in ways that allows meaningful state comparisons;
- Ensure that faculty are at forefront in defining educational objectives and developing measures;
- Make learning assessment results, including value-added measures, available to students and public;
- Administer National Assessment of Adult Literacy more frequently (every five years) and broaden the sample to include graduates of 2- and 4-year colleges and universities.

The "hook" to get much of this done is provided in the report's accreditation recommendations: agencies should make performance outcomes, including completion rates and student learning, the core of their assessment. The accreditation framework should:

- Allow comparisons among institutions on learning outcomes
- Encourage innovation and continuous improvement
- Require institutions/programs to move toward world-class quality
- Report measurable progress in relationship to national and international peers
- Make the accreditation process more open and accessible to public

[2] David Brooks, "Our World Cup Edge," AGB *The Presidency*, special issue 2006, p. 21.

The Department of Education has been actively pursuing this aspect of the action agenda. The Secretary hosted a national forum on accreditation in late November 2006.[3] And, it is one theme to be discussed at a summit on higher education to be convened on March 22, 2007, on the action agenda, followed by regional meetings. Possible adjustments to accreditation regulations are also being considered through negotiated rule making, which began in February 2007.[4]

The Commission probably did succeed in building more awareness of these issues. But, even before the Commission was formed, higher education associations had been working on new approaches to assessing learning, in response to accreditation requirements for structures, process, and evidence of more student learning assessment. This has been a serious concern for the past 20 years and more.

But, what policy makers seem to remember is that "our sector was dragged into outcomes measurement kicking and screaming. No one wanted to do it, no one knew how to do it," as Elise Scanlon, executive director of the Accrediting Commission of Career Schools and Colleges of Technology reminded the January 2007 Council for Higher Education Accountability (CHEA) annual meeting.[5] Many policy makers believe we are still kicking and screaming, despite the evidence of meetings likes this, the thousands of accreditation reviews each year, and more. There is a big perception and communication gap, and this begs the question of whether – even if the policy makers believed we were doing it – would what we are doing satisfy them?

The mantra has been, "grades are not enough." "Tests are not enough." Other evidence – learning portfolios, capstone courses, other embedded activities – were to be included. The American Association of Higher Education partnered with the North Central Association, running national seminars and working with institutions, all on a voluntary basis. The focus was on individual institutions or programs, that each would design its own goals and measures.

Vincent Tinto, professor and chair of the Higher Education Program at Syracuse University, said at the November 2006 National Symposium on Postsecondary Study Success, "with all due respect to my colleagues, one might argue that we already have sufficient research on student success….what is missing in our view is the ability to transform the knowledge that we have into practical knowledge."[6]

Margaret Miller, director of the Center for the Study of Higher Education at the University of Virginia's Curry School of Education, said at the January meeting of CHEA that much of this work was good but failed to ask the question, evidence "compared to what?" She has been persuaded that after two decades of trying to assess college (as contrasted with student) performance, it is not sufficient to gauge a college's success based only on information about itself. "The answer to the

---

[3] http://www.ed.gov/news/pressreleases/2006/11/11292006.html
[4] http://www.ed.gov/policy/highered/reg/hearulemaking/2007/nr.html
[5] Doug Lederman and Elizabeth Redden, "Accountability and Comparability," *Inside Higher Education*, January 31, 2007. http://www.insidehighereducation.com/news/2007/01/31/compare
[6] Elizabeth Redden, "Out of Step," Inside Higher Education, November 2, 2006. http://www.insidehighereducation.com/news/2006/11/02/success; http://nces.ed.gov/npec/pdf/symp_agenda_FINAL_10_06.pdf

question, ‘How well are we doing?’ really depends on an answer to a prior question: Compared to what? Compared to whom?”[7]

Some groups have been anticipating this shift. For example, the Council for Aid to Education partnered with higher education associations and institutions to develop and pilot test a “value-added” approach to learning assessment, trying to get at the general cognitive skills that college graduates should improve. The result of this effort is the Collegiate Learning Assessment (CLA), which has been administered for the past three years.[8]

But, following the Spellings Commission, activity has increased. The American Council on Education (ACE) sent a letter to members in September 2006, on behalf of a consortium of associations representing most of postsecondary education in the United States, essentially accepting the challenges posed by the Commission and proposing to address the renewed need for change. Its purpose was to reaffirm independence of approaches – to do all of this voluntarily but with more coordination and communication.[9]

The National Association of State Universities and Land Grant Colleges began, just around the time the commission was formed, to develop its own Voluntary System of Accountability (VSA). Now supported in part by the Lumina Foundation, the VSA project has grown into a national effort in partnership with the American Association of State Colleges and Universities (AASCU), with five working groups on student/parent information; core educational outcomes; student engagement; technical work on learning outcomes; and technical work on student growth outcomes. The product will be recommendations, including <u>specific tests and measures</u>, by the end of summer 2007. For example, the core educational outcomes task force charge is ambitious: to focus on how the academy can develop useful and credible evidence of the value that higher education adds for its students in terms of core educational learning outcomes and student growth outcomes. The purpose is 1) to convince the current skeptical audiences that what is measured really amounts to the addition of important value for the student; 2) the evidence must facilitate comparison among institutions; 3) evidence should be reliably useful in helping higher education institutions go about making programmatic improvement.[10]

In order to talk more specifically about learning assessment and testing, it is necessary to address accreditation. Accreditation is considered by many academics to be one of the most expensive and least value-added processes that universities are required to engage in. It may be largely invisible to prospective and enrolled students most of the time, but it looms as a large burden for many university and college administrators and faculty involved in assessment work.

For many programs and institutions, the return on investment in accreditation is the quality seal of approval that accreditation confers. Yet, accreditation, whether regional, national, or professionally specialized, is perceived as expensive and schizophrenic – focused simultaneously on regulatory compliance (to prevent fraud in use of federal financial aid) and on quality assurance and improvement.

---

[7] Lederman and Redden, Jan. 31, 2007.
[8] http://www.cae.org/content/pro_collegiate.htm
[9] http://www.acenet.edu/AM/Template.cfm?Section=Search&template=/CM/HTMLDisplay.cfm&ContentID=19857
[10] http://www.nasulgc.org/public%20affairs/pressreleases/nasulgc-aascu-luminarelease.pdf

What is broken? Various groups point out the following concerns:

- Lack of comprehensible standards;
- Lack of consistency in applying standards within accreditation associations and at the federal level;
- Bureaucratization and politicization of the accreditation process;
- Undercapitalization – scarce resources of accreditors and institutions to improve the focus on student learning;
- Lack of public information about performance;
- Internal focus on individual institutions – no comparability or ability to examine progress "compared to what?"
- Lack of objective data (of learning, of valuable skill sets, that the learning was obtained);
- Focus on inputs (resources and processes) more than outcomes of higher education.

For many, a – if not the – underlying policy question is: who should and will control quality in higher education? Again, quoting Margaret Miller: "If we can… look at ourselves carefully and rigorously, I think there's a very good chance that we will be able to control the terms in which this question is answered. If we can keep this question within our own control, we will do something that K-12 was unable to do, to everybody's great sadness."[11]

At the annual meeting of ACE in February 2006, Undersecretary of Education Sara Martinez Tucker told members that they are the professionals, "probably in the best position to understand what's necessary to ensure that we have more access and better results."[12]

The Secretary of Education's summit on higher education in March 2007 will include a theme on improving the focus on student outcomes through accreditation.

And, through negotiated rule making, which has just begun, the Department of Education is also looking at ways within the existing policy and legal framework to improve – to focus, clarify, align, and make more consistent – the application and implementation of the accreditation standards. For example, a frequent observation about accreditation is that the standards already have a lot that can be used to focus on student learning and to establish benchmarks of performance, but that the standards are not used fully. It remains to be seen what will be left on the table – for consideration in the HEA reauthorization process or elsewhere – when this process is done.

It is clear that throughout these activities at the national level, there is no consideration of federally mandated exit testing, nor high stakes testing. And there is not yet agreement that standardized tests must be used in all student outcome frameworks, although this strategy is being more broadly considered now. No one is advocating a "one-size-fits-all" approach, or comparing everyone to everyone else.

---

[11] Lederman and Redden, Jan. 31, 2007.
[12] Doug Lederman, "Into the Lamb's Den," *Inside Higher Education*, Feb. 14, 2007. http://www.insidehighereducation.com/news/2007/02/14/tucker

But, it is clear that the Secretary of Education wants to measure student learning and may be thinking about incentives for colleges, states, and consortia that are willing to report on it. And, there seems to be growing agreement that academia's current approach to assessing student learning is highly elaborate, labor intensive, and inward looking and doesn't really serve our interests in communicating the value added of institutions to the public.

Somewhat ironically given the mantra that "tests are not enough," much of the national and state debate has focused lately on standardized testing and comparability. Most admissions offices and faculty at most institutions don't have a problem using the SAT, GRE, LSAT, etc., to help evaluate prospective students. But, there is definitely a resistance to using the same types of tests to evaluate what happens once students are in the institution.

The debate is just that – framed as a black-and-white, either-or dichotomy – when it will probably be more productive to think about a suite of linked, complementary assessment approaches and tools

Margaret Miller, in "The Legitimacy of Assessment," distinguishes three overlapping purposes of testing and assessment:[13]

- To certify student achievement – assessing individual student achievement and potential so that the student is accountable and pays, for example, in taking the SAT or the LSAT.
- To improve programs – this focus increased in the 1980s through federal and state accreditation – asking are academic content and quality acceptable? In this framework, faculty and institutions are accountable – this is largely the focus of assessment experts like you; it is largely internally focused, and the institution pays, although external instruments like licensure exam results taken by individual students can be used to inform the analysis.
- To improve and assess institutions as a whole – focusing on holding institutions accountable to constituencies and asking: is the investment paying off? And, the institution pays.

Now we are experiencing a shift to a fourth level, to improve the outcomes of postsecondary education at the state and national level – the accountability is owned by policy makers asking: is progress taking place at the state level? At the national level? Are we competitive globally? At this level, data collected for other purposes may be selected and aggregated, and it requires methodological care to ensure the validity of conclusions. Presumably, costs are covered outside of institutions.

How would testing fit into these frameworks? Richard Shavelson, one of the architects of the Collegiate Learning Assessment, offers a positive approach in "Assessing Student Learning Responsibly: From History to an Audacious Proposal."[14] He emphasizes that there is no single test that can be used to assess all of the important things we hope a student gains during college. And, standardized tests are not a magic bullet; they don't pinpoint problems; at best, they signal where a problem may exist. External assessments must, according to the CLA framework, be used in the

[13] Margaret A. Miller, "The Legitimacy of Assessment," *Chronicle of Higher Education*, Sept. 22, 2006. http://chronicle.com/weekly/v53/i05/05b02401.htm
[14] "Richard J. Shavelson, Assessing Student Learning Responsibly: From History to an Audacious Proposal," *Change*, January/February 2007, pp. 26-33.

context of a range of institution-specific assessments and other processes. At the same time, one of the attractive features of the CLA is its capacity to provide an indicator of value added – what the difference in results from freshman to senior test-takers suggests about the value the campus adds to its learners.

While the CLA approach may not be perfect – it is a sampled approach; it does not necessarily test the same cohorts; it does not directly relate to the college curriculum; and it still needs to be replicated over several cycles to verify trends – it is among the most promising tools we have to get at this change. Shavelson's fundamental conclusion from a review of 35 years of attempts to assess learning is that the "cultures of evidence do not lead to educational improvement if what counts as evidence does not count as education or only counts as part of what we expect from a college education." To address this concern, he concludes by proposing – audaciously – that the CLA approach be combined with assessment of specific academic programs, to tap knowledge and reasoning within a discipline or to get at individual and social responsibility outcomes as content within the critical reading, writing, and problem-solving tasks.

By contrast, Trudy Banta, a doyenne of the assessment research movement, issued "A Warning on Measuring Learning Outcomes," based on her experience over the past 25 years and research going back to the 1960s.[15] She warned:

"A substantial and credible body of measurement research tells us that standardized tests of general intellectual skills cannot furnish meaningful information on the value added by a college education nor can they provide a sound basis for inter-institutional comparisons. In fact, the use of test scores to make comparisons can lead to a number of negative consequences, not the least of which is homogenization of educational experiences and institutions." For Banta, this potential consequence calls into question the greatest strengths of American higher education – its variety of opportunity and lack of central control.

She concludes that, while standardized tests can be helpful in initiating faculty conversations about assessment, her research casts serious doubt on the validity of using standardized tests of general intellectual skills for assessing individual students, then aggregating their scores for the purpose of comparing institutions.

My own "audacious proposal" is that these points of views should not be allowed to cancel one another out; that serious researchers continue to develop and refine instruments; that institutions continue to use them and objectively evaluate how effective they are in answering the policy questions and contributing to institutional improvement; and that accreditors and policy makers create incentives and practice some patience while this work goes on. Perfection will definitely be the enemy of making progress in this debate.

[15] Trudy W. Banta, "A Warning on Measuring Student Outcomes," *Inside Higher Education*, Jan. 26, 2007. http://www.insidehighereducation.com/views/2007/01/26/banta

Texas Trends

State legislatures and policy makers have been addressing issues of higher education accountability for years through the National Governors Association, National Conference of State Legislatures (NCSL), and the State Higher Education Executive Officers (SHEEO) group. Following on recommendations of the Spellings Commission, the NCSL issued a report in November 2006, *Transforming Higher Education: National Imperative – State Responsibility,* recommending that "state legislatures must take a leading role in reforming higher education, to make college more accessible and to better meet state needs in a competitive global economy."[16] Commenting on the report's findings, Texas Representative (and chair of the House committee on higher education) Geanie Morrison, and said:

> "too often, higher education policy is made in reaction to the latest budget crisis rather than on the basis of long-term strategic goals. The Texas legislature is trying to do the latter: set clear expectations and develop data collection systems….And base appropriations on progress toward goals – retention, graduation, increase accessibility and affordability."

Texas Governor Rick Perry put forward a bold proposal to set aside incentive funding for production of degrees, weighted for at-risk students and critical fields. This framework would include performance on exit – but not high stakes – exams that each graduating student would have to take. Institutions would receive incentives based on the overall performance compared to national standards, signaling a high bar for all.

The policy issue is how well particular degree programs prepare students compared to the same degree programs at other schools – it is a mix of individual, program, and institutional accountability. Thus, the focus would be on discipline-specific tests but would also include licensure exams and others that could be designated – perhaps GREs, LSATs, etc. Existing, nationally- or state-normed tests would be used; none would be created for this purpose.
Each student would just have to take one test for this purpose. The institution would pay for the exams. The key is in the details, and much remains to be worked out. Since The University of Texas System has used a form of testing – through sampling only – for three years for institutional accountability, the nervousness about the idea is familiar. The UT System experience also suggests that it is worth having the policy discussion and clarifying the purpose and use of this approach.[17]

Meanwhile, the Council of Public University Presidents and Chancellors in Texas has developed a proposal for a different kind of incentive program. The key operating principles are that, first, base funding, growth, and cost-of-living increases would be covered with new incentive funding above that level. Then, based on a survey of all institutions, the proposal is to consider a small group among standard performance indicators: graduation and persistence rates; enrollment, including diversity; administrative efficiencies; community college transfers; faculty instruction of lower-division students; and research measures.[18] Comparisons would be to prior individual institutional performance – not in comparison to other institutions.

---

[16] Karin Fischer, "Panel Urges State Lawmakers to Take the Lead in Higher-Education Reform," *The Chronicle of Higher Education*, December 8, 2006. http://chronicle.com/weekly/v53/i16/16a02202.htm
[17] Ralph K. M. Haurwitz, " UT System Applauds Standardized Testing," *Austin American Statesman*, Feb. 9, 2007. http://www.statesman.com/news/content/news/stories/local/02/09/9test.html
[18] Private correspondence, CPUPC.

Given these policy directions in Texas, it is worth noting the experience in Florida with college-level standardized testing. The College Level Academic Skills Test was designed to test readiness for transfer from two-year to four-year programs. A proposal was made earlier this year for high-stakes college-level tests, i.e., required for graduation. People involved in this plan have commented that it was hard for Florida to set aside sufficient funding for incentives; ending up even, not budget increases, was considered an achievement.[19]

At the institutional and System level in Texas, the UT System includes a set of multiple measures of student outcomes for each institution in its accountability reports.[20] This is, of course, in addition to the individual Quality Enhancement Plans required by SACS for regional accreditation or other approaches to assessment that each institution conducts, and these two approaches are not yet directly linked.

The UT System decided three years ago to try to include some information about outcomes assessment tests in the accountability framework. It is not easy or free, but it can be done. This student outcomes framework currently includes:

- 1st year persistence.
- 4-, 5-, and 6- year graduation rates – in addition the Board gets quarterly updates on progress, and a separate, in-depth annual analysis of the progress (or not).
- Graduation rates of community college students who enter UT Sysytem institutions with at least 30 credits to transfer.
- Composite persistence and graduation rates for students who matriculate at a UT System institution but go elsewhere in Texas to complete their education.
- Licensure exam pass rates for key fields (teaching, engineering, etc.).
- National Survey of Student Engagement (NSSE) data – this is reported to the UT System Board of Regents separately and in depth. For the accountability report, measures are extracted regarding satisfaction with advising, the students' overall assessment of their college experience, and whether they would "attend again" if given the chance.
- Results of Collegiate Learning Assessment – this is administered to small random samples of freshmen and seniors. Results are summarized and displayed in a fair amount of detail and analyzed in depth – the Board gets a separate, in-depth report on these, too.
- Data on postgraduation experience – employment or enrollment in a graduate/professional program after graduating, provided by the Texas Higher Education Coordinating Board.
- The accountability report is posted on the Web and sent to legislators and hundreds of other people in Texas and nationally. The data are completely public. The institutions use the data in their internal planning and improvement efforts. The UT System looks at the trends when it evaluates presidents and when proposals are considered by the Administration and the Board about specific investments in the campuses. In other words, the information is used for strategic decision making and institutional improvement.

The UT System is aware that this framework is layered over the complex, internal assessment plans that SACS requires each institution to develop and implement – and this is the level at which faculty can and must become engaged. A policy question for the future is to what extent should

[19] R.E. LeMon, personal correspondence, Feb. 13 2007.

[20] *The University of Texas System 2006-07 Accountability & Performance Report.* http://www.utsystem.edu/osm/accountability/2006/studentaccess.pdf

assessments be incorporated into the regular curriculum, within capstone courses, on transcripts, etc.? These activities do not directly align with the Governor's or CPUPC's work, but the expectation is to keep working with all policy makers on the details.

Concluding Questions and Observations

What does the future hold for accountability and learning assessment? What will be different about postsecondary education if the changes being considered now take hold? What will be the intended consequences? What about UN-intended consequences? Will people be dancing in the streets – or rolling in their graves?

Will Texas and the national be more competitive? For example, in its October 2006 survey of global higher education, *The Economist* predicted that the meritocracy gap will grow, not get smaller, between those with good educations and those without.[21] Where should the bar be set, in terms of state-level and national policy? Should we aim to get more, or even most, people to some mean standard? Or should we aim to prepare more people to excel? In this same survey, *The Economist* reported that "most people in America will accept inequalities if they are coupled with equality of opportunity." Do we agree? Is this the right policy? *The Economist* concluded that "the rise of a global meritocracy offers all sorts of benefits from higher growth in productivity to faster scientific progress." And, if policy makers don't think they can afford both equity and excellence – and neither of these is free – how do they choose? What is the best return on the public investment? Only if those choices are understood, can we decide what would be the most critical and useful assessment and outcome goals and measures. If, instead, we just keep assessing and testing because it seems to be a good idea generally or a convenient policy tactic, I predict continued dissatisfaction with the process and its utility.

What about testing? Just about anything can be measured and tested. A recent story described a thoughtful, serious assessment framework of Association for Biblical Higher Education for measuring a student's spiritual growth since that is core to the mission of the schools this organization accredits.[22] If regional and specialized accreditation continue along current directions, these approaches are likely to remain useful on some level. But, we might expect to see more radical changes in the business of testing and assessment – and this is a big business. For example, the National Center on Education and the Economy predicted in its 2006 report *Tough Choices or Tough Times* that moving from how we test now to capture other qualities (creativity, innovation, facility with ideas and abstractions, self-discipline, organization to manage one's work, ability to work in teams, etc.) "will entail a major overhaul of the American testing industry…with revised curricula, syllabi, teacher training….**But it all starts with the standards and assessment"** [emphasis mine].[23] Note that this report did not emphasize the importance of assessing students' knowledge of specific disciplines.

---

[21] Wooldridge, Adrian. "The Battle for Brainpower: A Survey of Talent." *The Economist*, Oct. 7, 2006. www.economist.com/surveys

[22] Elizabeth Redden, "Spiritual Accountability," *Inside Higher Education*, Feb. 1, 2007. http://www.insidehighereducation.com/layout/set/print/news/2007/02/01/assessment

[23] National Center on Education and the Economy, *Tough Choices or Tough Times*, report of the *New* Commission on the Skills of the American Workforce, December 2006. http://www.skillscommission.org/executive.htm

So, it also depends on who is the customer – is it the student, the institution, policy makers, the public? What kinds of assessments and tests can serve more than one purpose and more than one customer group? The emphasis on benchmarking and broad-scale comparisons will grow. The focus on institutional, as contrasted with individual, accountability will increase. We might see development of more interrelated tests along the continuum of milestone events in postsecondary education: Two-year college ready; need for remediation; four-year college ready; transfer ready; employment ready; graduate/professional school ready; profession-change ready…

Can we benefit from our experiences and successes? Policy makers and educators should focus on gathering and improving the best practices, based on the twenty years of experience in assessment. My recommendations align in many ways with those that Lee Shulman wrote about in *Change* recently:[24]

Resist the urge to leap to a single, common measure or instrument – there is no "magic bullet" test; there is no "sleep number" that will give a single good housekeeping seal on an institution. Multiple measures are critical, but let's figure out how to combine them in a simple framework. Measuring everything in a disaggregated way is not much more useful than measuring nothing.

Resist the urge for perfection – the "next generation" of accountability, assessment, and testing must focus on analysis and improvement, not punishment, and not high stakes! Use the information and imbed it at many levels – in academic programs, institutional planning, accountability reports. The more they are used, the more the methodologies, instruments, and reporting formats will be improved.

Resist the urge to be overly critical, to focus on what won't work and why not – embrace opportunities to figure out what we already have that can be improved and what might be added or changed to produce useful, meaningful information at every level of analysis – for the student, the program, the institution, the state, the country.

And now that the words accountability and assessment are nearly household terms, as experts, I hope you will embrace the opportunities to be transparent, to share and publish data, take credit for success, and be forthright about addressing areas of weakness. As one university president said in a speech to his faculty assembly, "the public has every right to expect results. Our challenge is to meet these expectations head on and do a much better job explaining our value."[25]

Finally, please participate actively in the national dialogue, not defensively. There is so much to build on, as you have demonstrated at this conference…the future of accountability and assessment **is** in our hands.

---

[24] Lee S. Shulman, "Counting and Recounting: Assessment and the Quest for Accountability," *Change*, January-February 2007, pp. 20-25.

[25] James D. Spaniolo, "Demonstrating Achievement Through Assessment and Accountability," The University of Texas at Arlington, University Faculty and Associates Meeting, October 12, 2006. http://www.uta.edu/president/about/speeches/20061012.php