HAMPSHIRE AND WABASH ASSESSMENT COLLABORATIVES REVIEW

Charles Blaich and Kathy Wise
Center of Inquiry in the Liberal Arts
Wabash College

www.liberalarts.wabash.edu

Introduction

The following is a report on the progress of two assessment collaboratives, the Hampshire Teagle Consortium and the Wabash National Study of Liberal Arts Education, toward improving student learning at their member institutions.

The two consortia were formed in 2005 in the midst of rising national pressure for colleges and universities to demonstrate that they create the kinds of intellectual and personal growth they claim. During this time, the United States Department of Education was publicly exploring the possibility of creating standardized assessment measures for higher education. The Department of Education was also pushing regional accreditors to hold colleges and universities accountable for demonstrating their impact on student learning.

Partly in response to these growing demands, the Teagle Foundation created a number of grant initiatives that called for private colleges and universities to collaborate to create "faculty-friendly" forms of assessment.[1] The leaders of the Teagle Foundation framed these grants both in terms of the legitimate responsibility that colleges and universities had to assess student learning and the importance for institutions of assuming this responsibility before the government did so by fiat. In essence, when it came to assessment and accountability, those of us in higher education should "do it to ourselves before the government does it to us."

The Hampshire and Wabash consortia were founded on the idea that one way of increasing the formative impact of assessment at institutions was to have faculty, staff, and administrators across institutions collaborate to gather and share assessment data. The leaders of both consortia believed that inquiry-driven, cross-institutional collaboration would reduce the likelihood that staff and especially faculty would see assessment as bureaucratic or externally mandated. They hoped that approaching assessment as a process of inquiry about student learning would help faculty understand assessment as a natural extension of the kind of intellectual work that first brought them into the academy.

---

[1] For background on the Teagle Foundation's Faculty-Driven Value-Added Assessment Collaboratives grant initiative, see the report on Teagle's September 2004 Listening on Value-Added Assessment at http://teaglefoundation.org/learning/pdf/struck.pdf. Also, go to the Outcomes and Assessment section of http://teaglefoundation.org/grantmaking/rfp.aspx for links to the RFPs and lists of funded projects for different rounds of the assessment collaboratives initiative.

What follows is a first evaluation of what happened in the Hampshire and Wabash consortia. It is based on interactions we have had with the six institutions in the Hampshire consortium, up to and including presentations at the April 2009 Mohonk conference. It also includes our work to date with other institutions in the Wabash National Study.

The review will begin with reflections on the nature of assessment and how to gauge its success, turn to a discussion of things that did and didn't work in these two collaboratives, and end with some questions about the extent to which collaboratives and "faculty-friendly" assessment work in general.

What is assessment?

Both the Hampshire and Wabash consortia began as "assessment collaboratives." In doing so, both were created assuming that assessment refers to the process of inquiring about student learning and using the results of those inquiries to inform changes that demonstrably improve student learning. Yet, the term "assessment" has commonly been used to describe two distinct processes. One use is consistent with the way that the consortia understood "assessment," as a process that includes both gathering evidence and using it to promote improvements. For example, Palomba and Banta (1999) state:

> Assessment is more than the collection of data. To make assessment work, educators must be purposeful about information they collect. As a basis for data gathering, they must clarify their goals and objectives for learning and be aware of where these goals and objectives are addressed in the curriculum. After data are gathered, educators must examine and use assessment results to improve educational programs. (p. 4)

This is similar to the way that the Teagle Foundation defines assessment:

> The Teagle Foundation's Outcomes and Assessment program is based on the belief that nothing has the potential to affect students' educational experience as much as the systematic assessment of what they learn, along with the use of such assessment to frame discussions on learning and to drive continuous improvements in teaching practices. (Teagle Foundation, n.d.)

However, other scholars and assessment experts refer to assessment as the process of systematically gathering and interpreting evidence about student learning—not the process of using that evidence to make changes. Schuh, Upcraft, and Associates (2001) describe assessment as "any effort to gather, analyze, and interpret evidence which describes institutional, divisional, or agency effectiveness" (pp. 3–4). They go on to distinguish assessment from evaluation:

> Assessment must be contrasted with but linked to evaluation. . . . We assert that evaluation is any effort to use assessment evidence to improve institutional, departmental, divisional, or agency effectiveness. In other words, assessment describes effectiveness; evaluation uses these descriptions to improve effectiveness, however that might be defined by the institution. (p. 4)

Choosing one of these definitions is critical for evaluating the success of the Hampshire and Wabash assessment collaboratives. If they are assessment collaboratives in the sense that Schuh, Upcraft, and Associates describe, then gathering and reviewing evidence about student learning constitutes success. If, however, assessment is taken to mean gathering, reviewing, and using evidence to change practice, then the criterion for success is improved student learning.

We think it is clear that the richer "gather-and-use" standard should be applied when evaluating the success of the Hampshire and Wabash collaboratives. This higher standard was the intent of the Teagle Foundation, as well as the leaders of the two collaboratives. However, it will be important to return to these two understandings of assessment as we consider practices that did and did not work, and how framing assessment as inquiry implicitly favors one of these two understandings of assessment.

Were the collaboratives successful?

Since the goal of both collaboratives is to improve student learning, the obvious starting point is to ask whether the 22 institutions in the two collaboratives have palpably improved student learning. [2] A gentler way of asking this question is to ask how many of these institutions have a)

---

[2] Nineteen institutions from the original round of the Wabash Study plus Hobart and William Smith Colleges, Allegheny College, and Vassar College.

introduced any changes based on the assessment data they gathered during this project, and b) assessed the impact of those changes.

It is important to insert one caveat here. Although we have visited many of these campuses, participated in a number of meetings with representatives from these institutions, and had frequent phone and email conversations with campus representatives, it is still very difficult to find out exactly how institutions are working with their assessment data. Institutions, even small ones, are decentralized organizations, and so it is possible that we are unaware of ways in which campuses have engaged with their assessment data. Therefore, the instances we cite below represent a "lower bounds" or conservative estimate of how institutions in the collaboratives are responding to evidence about student learning.

We believe that at least three institutions have succeeded in introducing changes based on assessment data they gathered in the project and assessing the impact of those changes. Two institutions have introduced and assessed changes that were designed to increase the level of academic challenge, specifically focusing on the amount of work students are doing and the intellectual challenge of course content and assignments. Another institution introduced and assessed a student-faculty lunch program that they developed to increase the frequency of interactions between students and faculty.

Although the remaining institutions have not reached this standard, nine institutions have in some way engaged with the evidence that they have collected. These institutions are actively involved in a variety of activities including hosting site visits from Teagle Assessment Scholars, reviewing assessment evidence in various committees, attending meetings on how to better make sense of and use the evidence they've gathered, and gathering more assessment evidence to clarify the information they've initially collected. These responses are serious and public, and they move the institutions well beyond the kinds of activities in which they engaged at the start of the collaboratives. Indeed, a number of the institutions are working seriously on deep institutional issues, such as fundamental problems with the curriculum and minority student persistence and success, that have percolated for many years without being addressed.

Overall, three years into these collaboratives about half of the institutions in the two groups are actively engaging their assessment data. It is also important to note that the Wabash collaborative gathers data over a four-year period, so new institutions may join this list in the

next year. For example, three Wabash collaborative institutions recently contacted the Center of Inquiry for the first time to express interest in working with us to use their assessment data.

Does the level of activity to date count as success? It depends on the standard we apply. Comparatively speaking, there's little information on how many assessment projects ultimately produce improvements in the quality of student learning. Most conferences and publications highlight successes rather than failures, making it hard to develop a sharp sense of the proportion of projects that successfully improve student learning, or at least promote active campus engagement with assessment evidence. One recent article suggests that the rate at which campuses attain either of those levels of success is very low. To collect information for a new book, Trudy Banta, Elizabeth Jones, and Karen Black invited over 1,000 people at colleges and universities to submit profiles of good assessment practice at their institutions (Banta, 2009). Banta, Jones, and Black received 149 profiles (15% return rate), and only six of the 149 profiles contained evidence that student learning had improved (Banta, 2009).

According to Banta (2009), the vast majority of the profiles described the impact of assessment in terms of "improved academic processes" such as increased campus awareness of the importance of assessment, more efforts to delineate learning outcomes and gather evidence about those outcomes, more efforts to design rubrics and identify quantitative instruments, more use of assessment findings, an institution meeting accreditation requirements, increased resources to support assessment, and increased professional development to support more assessment by faculty and staff. Banta (2009) argued that one practical obstacle to using assessment evidence to promote change is time: "It often takes at least two years to collect sufficient data to convince faculty that a change is needed. Then it may take a year or two to craft and implement an appropriate change and another two years to collect data on the impact of those changes" (p. 4).

Another hint about how often institutions use assessment to create demonstrable improvements in student learning comes from a statement on assessment released by the Consortium on Financing Higher Education (COFHE, n.d.-a) According to the statement, institutions have, "a fundamental responsibility to assess the effectiveness of our programs (COFHE, n.d.-a, para. 1)." COFHE also argued that colleges and universities were already engaged in "effective assessment in ways that are appropriate to our missions and that enhance the quality of the

educations we provide to our students (COFHE, n.d.-a, para. 3)." COFHE listed the 93 institutions that signed the statement, and then, for 73 of the institutions, provided examples of how these institutions use assessment (COFHE, n.d.-b).

Almost all of the examples COFHE provided refer to the process of assessment—forming assessment committees, obtaining grants to fund assessment programs, reviewing student work, or administering surveys. Only 24 of the examples referred to changes that followed from assessment. Even then, the majority of these changes were described in vague terms, such as "The results of this project are being used by faculty and staff at *Institution X* to develop and improve programs to support students, both academically and socially."[3] We did not find a single example that described in any detail how assessment prompted a change that led to demonstrable improvements in student learning.

By the standard set in these two examples, the Hampshire and Wabash collaboratives show exceptional promise. In three and a half years, a relatively short period of time, they have created an environment in which at least half of the institutions are engaged in the kinds of activities that Banta and her co-authors plan to highlight in their book. Furthermore, most of the institutions in these collaboratives will continue working with their assessment evidence for several more years. So the opportunity for institutions to implement successful changes remains.

As well as these collaboratives have done in comparison to the low bar we described in the two examples above, it is important to acknowledge that we should adopt a much higher absolute standard of success, especially when we consider the amount of money, time, and energy that have been devoted to these collaboratives.

What helped to promote institutional engagement with evidence?

In this section we will describe activities, methodologies, and attitudes that promoted campus engagement in these collaboratives.

---

[3] This is a direct quote from one of the examples.

Talking with students to help interpret assessment evidence

One of the most important good practices that we discovered during our site visits was the simple act of talking with students to help us understand survey data. These conversations with small groups of students usually followed one of two paths. Sometimes, we would ask students general questions about their experiences at their institution. We usually framed these questions in pairs, for example, "Think about the best faculty that you've had. What did they do that made them so good?" and "Think about the worst faculty that you've had. What did they do that made them so bad?" We would ask similar pairs of questions about the best and worst projects they had done, their best and worst experiences at college, the things that they would change about their institution and the things that should remain the same, etc. Students usually responded with vague answers. "Professor X is just so cool!" When this happened, we would always push for more specific answers or examples. "What made professor X so cool? Can you give me an example?"

Other times, we would highlight survey data that intrigued or confused us and ask the students to help us make sense of the data. For example, why would students at a diverse institution report having very few meaningful interactions with diverse peers? Or why would students report on one question that their institution emphasized lots of academic work, and then report spending only a modest number of hours each week preparing for class later in the survey?

In most cases, the things students talked about, both good and bad, lined up very neatly with data from the National Survey of Student Engagement (NSSE) or other survey questions about student experiences. Indeed, we learned that small quantitative variations in survey data often pointed to larger problems than we first thought. For example, students at one institution reported experiencing slightly lower levels of teaching clarity and organization than students at other similar institutions. The students with whom we spoke at that institution consistently mentioned a sense of disorganization in some of their courses. We later learned that the senior faculty member who chaired the tenure and promotion committee publicly dismissed the use of course syllabi because they constrained creative pedagogies. We do not know how many faculty followed his advice, but based on comments from students and staff, we believe that at least a few faculty put this advice to practice.

In general, our follow up conversations with students helped us get a sense of how they interpreted generic survey questions and gave us specific examples, including a sense of student voice, that we could use in our conversations with faculty, staff, and administrators. We believe that these conversations also helped to show students that people were listening to and thinking about their responses to surveys. Finally, our conversations with students helped us move beyond beta coefficients, means, and standard deviations to form a richer understanding of the degree to which students were moved by their interactions with faculty and staff, the stark decline in their interest in academic work over the first year of college, and the particular ways that students understand the purpose of college. Attempts to improve student learning cannot succeed without these conversations.

Engaging multiple methods to understand evidence about student learning

Because faculty and staff are trained in academic disciplines, they come to assessment with a strong predilection to apply these methods to their inquiries about student learning. This makes perfect sense in light of the years they have spent learning and refining these methods. For scholars methods are not just skills, they are ways of knowing the world. One practical consequence of this is that faculty and staff from the social and natural sciences are often far more comfortable gathering and interpreting quantitative assessment data, while faculty and staff from the humanities are usually more comfortable with narrative forms of assessment.

Unfortunately, evidence about student learning is not split neatly along the lines of scholarly expertise. It comes from surveys, student papers, blog postings, conversations, and rhetorical and multivariate analyses. The discomfort that faculty and staff experience as they look at information for which they have no interpretive expertise beyond their common sense can cause them to disengage from working outside of the boundaries of their disciplinary training. At its worst, faculty and staff can become methodological purists dismissing any form of evidence that is not gathered via a certain kind of interview or random assignment experiments.

It is not clear to us whether such methodological purity is really about methodology or an attempt to block the process of assessment. What is clear is that the campuses that are responding vigorously to assessment data are using information from a range of sources including interviews, surveys, course grades, quantitative outcomes measures, and student papers. Although faculty and staff at these institutions may be focusing on methods common to

their disciplines, the people who are leading the assessment efforts are not. This willingness to use information from different sources and different methodologies is, we assert, a best practice. Conversely, the willingness of campus assessment leaders to disregard, for example, all quantitative or all qualitative data, marks that institution as being far less likely to engage assessment information.

Of course, the openness to considering a broad range of assessment information comes at a cost for campus assessment leaders—more information means more complexity. This leaves campus assessment leaders with a difficult intellectual task. They have to be able to work through large amounts of disparate information and develop a concise set of themes or findings that they can discuss with their colleagues. Synthesizing and consolidating these different forms of information is hard enough, but then assessment leaders need to identify a small set of potential responses that address the evidence and are politically and economically feasible. This second step of moving from understanding to action takes assessment leaders still further out of their customary academic habits of mind. As we consider the kinds of developmental support we provide for institutional assessment leaders, it will be important to find ways to help them develop their capacity for this new form of intellectual work.

Campus leadership

Strong, supportive, and consistent campus leadership is essential for institutional engagement with assessment. We've identified at least four effective leadership practices from our work with institutions:

1. Protecting staff and faculty from colleagues who are morally opposed to assessment.

2. Providing resources to assessment leaders, faculty, and staff who are gathering, considering, and using evidence on student learning.

3. Providing space within existing campus governance structures in which assessment leaders can bring evidence of student learning to bear on the official work of the institution.

4. Publicly articulating their support for assessment and campus assessment leaders. This public articulation is most effective when it takes an improvement-oriented perspective

and focuses on using assessment to strengthen the institution, rather than the defensive stance of using it to keep accreditors and other outsiders at bay.

We used the adjective "consistent" in our description of effective campus leadership above to emphasize another point—Presidents, Provosts, and Deans at the collaborative institutions have typically had relatively short tenures. The majority of the 19 institutions that joined the Wabash National Study in 2006 have had significant changes in academic leadership since they joined the study. In some cases, this change has increased the engagement of the institutions, while in other cases it has declined. But in every case, changes in leadership had an impact on how the institutions in the collaboratives used assessment. To improve the impact of our assessment programs, we will have to find ways to create a consistent institutional commitment to the kinds of leadership practices we describe above despite changes in campus leadership.

Using student work for assessment

The Hampshire collaborative gathered and evaluated student work. Although the mechanics of accumulating papers and teaching faculty and staff to use rubrics to evaluate them is challenging, our sense is that incorporating student work into campus assessment projects has two critical benefits. First, faculty and staff more readily connect with student work than they do with surveys or standardized outcome measures. Not only does student work seem inherently more relevant than most standardized outcome measures and student surveys, but it is also easier for faculty and staff to move from evaluating student work to identifying possible changes in classes and programs. This leads to the second benefit—the evaluation of student work readily translates into grist for faculty development. Our experience is that individuals responsible for faculty development can translate assessment based on student work into workshops, micro-teaching, Scholarship of Teaching and Learning, and other faculty development activities more easily than they can utilize standardized assessment measures.

Mixed blessings

In this section, we will review some of the recurring phrases that have come up at the Mohonk meeting, in our work with the Hampshire and Wabash collaboratives, and in our work with institutions outside of the collaboratives. Generally, these phrases are meant to be encouraging, but we believe that they carry more complicated implications for the success of assessment

projects. We will describe below our sense of the positive and negative implications of these phrases for campus engagement with assessment.

"The most important outcome of this assessment project has been to create conversations"

>   Positive implication – Rich, formative, inviting conversations are a necessary part of good assessment projects. Good conversations build community and trust, and they can develop support for implementing responses to assessment data in challenging political and resource-limited environments.

>   Negative implication – Sometimes satisfaction with good conversations is really an excuse that keeps institutions from moving on to responding to their data. When it comes to making changes to our institutions, good conversations are a means, not an end. We set the bar too low when we congratulate ourselves for merely talking about student learning.

"Our campus/student body/curriculum/department is unique"

>   Positive implication – Developing a thicker understanding of the culture of an institution is important both for developing better interpretations of assessment evidence and identifying politically and culturally plausible responses. Taking the special qualities of our institutions into account is a good assessment practice.

>   Negative implication – The less-favorable take on "uniqueness" sounds like this, "The unique qualities of our campus/students/curriculum/department means that fifty years of research on the practices and conditions that promote student learning don't apply to our institution. Student confusion about assignments or projects is a sign that our program is appropriately challenging; syllabi are restrictive pedagogical tools and student confusion about the structure of their classes is formative; our students learn from lectures because our faculty are master lecturers; data indicating that 40% of our students don't write drafts of their papers is meaningless because they actually use continuous revision; etc." This "we are unique" argument can all too easily be used to dismiss any challenging news that emerges from assessment.

"We need more assessment data before we can act"

> Positive implication – Student learning is complicated. It is often helpful to dig into data to get a better sense of its implications for an institution. For example, it is always important to use focus groups to follow up on survey data to build a better understanding of the specific institutional and classroom practices behind students' survey responses.

> Negative implication – Assessment becomes the ongoing process of gathering and analyzing data without any action toward the improvement of student learning. There are two reasons this can happen. First, assessment evidence often challenges practices and structures that faculty, staff, and administrators value by suggesting that these practices and structures do not promote student learning. Sometimes the easiest way to keep doing assessment without threatening those cherished structures is to continually gather more data. In essence, doing more assessment becomes a defense mechanism. When this defense mechanism kicks in, gathering and filing data safely away without using it becomes a ritual institutional act. One institution in the collaboratives has surveyed incoming students for over 30 years without ever, at least as far as we can tell, using the data to inform a single action. Second, it is important to remember that scholarship is built around the process of gathering and reflecting on evidence and information for the sake of gathering and reflecting on more evidence. It will seem both natural and satisfying to academics for assessment to evolve into scholarly research. But scholarly research, especially in the liberal arts, eschews application. Its success is based on the fact that it generates more questions and more publications, not on whether it creates practical changes outside of the world of scholarship.

"Assessment evidence must be kept separate from faculty evaluation"

> Positive implication – For assessment to flourish, it has to be formative, not punitive. Faculty and staff will engage in assessment only if it helps them create the kind of learning environment that they value. Once we connect assessment evidence with faculty evaluation, assessment will be rejected or "gamed" by faculty to protect their jobs.

Negative implication – Ineffective teachers will get tenure and be promoted at liberal arts colleges. Faculty teaching prowess will continue to be evaluated on the basis of whether faculty are popular among students, whether they get good course evaluations, and whether or not their colleagues are engaged by their classroom practices, all of which may or may not be effective for promoting student learning.

Open questions

Is it best to start assessment projects by focusing on evidence that institutions already have or by gathering new evidence?

We've raised this question because it is clear that most institutions, including those in the two collaboratives, already have a great deal of information that they could use for formative assessment. The question cuts in different directions for the Hampshire and Wabash collaboratives. The leadership of the Hampshire collaborative originally intended to pool survey data that institutions were regularly collecting. These surveys included the Higher Education Research Institute's (HERI) Freshman Survey, the Higher Education Data Sharing Consortium's (HEDS) Senior Survey, and the Indiana University Center for Postsecondary Research's NSSE. Although the consortial effort to gather student papers was successful, the attempt to pool commonly used surveys was not. Even though the Hampshire consortium institutions used many of the same surveys, and most institutions in the consortium administered those surveys on a regular basis, differences in when and how they administered the surveys made it impossible to create a data pool with common survey responses. Ultimately, the collaborative solved this problem by participating in the Wabash Study and sharing the data that was gathered in the study.

The institutions in the Wabash consortium all had common survey data by virtue of their participation in the study. But campuses' participation in the study often ended up highlighting the amount of useful data they already had when they entered the study. The Wabash Study includes questions from two surveys, the HERI Freshman Survey and the NSSE. Since these two instruments are among the most widely used surveys in higher education, it was not surprising that most of the institutions in the Wabash Study had used them prior to joining the study. It also was not surprising that students' responses to these surveys in the Wabash Study were similar to their responses in previous administrations of the surveys on their campuses. Students' responses to most of these surveys are remarkably stable across different cohorts.

Despite this fact, our experience is that people on campus were often surprised by our reports. At one Wabash Study campus, we described survey results indicating that students did not spend much time preparing for class. The faculty responded that they couldn't assign more work because most of their students were working outside of college to pay their bills. It turns out that one of the surveys we administer in the study, the NSSE, includes a number of questions about how students spend their time, including how much time students work for pay. We checked and found that only a tiny fraction of the students at this institution reported working for pay. Faculty and staff were surprised by this finding even though this institution has been administering NSSE for years and getting the same data about the number of students who work. Thus, despite years of evidence to the contrary, faculty at this institution were structuring their classes based on a completely erroneous assumption about their students.

This happens all the time. At most institutions assessment data are not widely known on campus. Why not? In some cases it is because the data are controversial or, as one Provost put it, "too depressing." In others, assessment data is used by a small group of people for accreditation but not for ongoing institutional improvement. Sometimes assessment reports are too complex, and faculty and staff aren't sure why they should pay attention. Finally, many administrators are unsure about how they can profitably share the information. The end result is that most campuses have actionable evidence about student learning that they are not using. In fact, in almost every case in which data from the Wabash National Study prompted fruitful conversations on campus, that data merely replicated information that campuses already had. Ultimately, one of the important benefits of the Wabash Study was to surface evidence that institutions already had.

Turning back to the question of whether campuses should use what they already have or gather more data, we conclude that the positive impact of these two collaboratives to date does not result directly from the assessment data they've gathered, *but from activities that each collaborative created to support gathering and using new data.* The public act of gathering new data in a consortial setting and committing to discuss and share that data are essential steps towards using evidence to improve student learning. It is, to invoke an old assessment phrase, a step towards creating a culture of evidence—or at least a culture that doesn't bury evidence.

While replicating previously gathered information is wasteful, there is an ironic efficiency underlying this waste—by bringing old data to light as they gather new data, campuses end up considering a broader array of evidence. The question for the future is whether we can create the kinds of environments and activities that support reflection on data an institution already has without adding an expensive layer of additional data collection.

What is a reasonable time frame for change?

In reviewing the work of the two collaboratives, we believe that there are two timelines we should adopt for future work. First, it should take no longer than three years for campuses to move from engaging in no public assessment activity to having campus-wide discussions about assessment, making sustained efforts to delineate learning outcomes across departments and programs, experimenting with assessment measures and techniques, and reallocating resources to support sustained assessment work[4]—in essence, the kind of activity that about half of the institutions in the two collaboratives demonstrated. We suggest that in the future we identify a developmental set of "assessment benchmarks" that define activities that institutions need to meet in the first, second, and third years following the creation of an assessment initiative.

Of course, this three-year period only marks the first stage in which institutions use assessment to identify areas that they want to improve. We believe that Banta's five- to six-year timeline for producing documented changes in these areas is reasonable. The very kinds of successful activities that these collaboratives emphasize, such as sustained faculty engagement, collaborative meetings, and focused follow-up data gathering, are all processes that take considerable time. But we believe that they are the kinds of processes that lead to deeper cultural changes that will support a sustained commitment to using evidence to improve student learning. Changing an institution's culture and values takes time.

---

[4] Please note the use of the word "public" here. Many campuses have sustained assessment efforts that are implemented by a small group of people. In our view, if assessment results are not widely publicized and shared across campus, it's as though that effort hasn't taken place. Campuses cannot create sustained, evidence-driven changes in student learning when only a small group of people knows about the evidence.

What is a good batting average—in other words, what percentage of institutions should improve in order for us to call our work successful?

We answer this question from the perspective of an organization that is learning at the same time as it is implementing programs. Our practice is to work incrementally and establish goals that represent an improvement from our current work. Based on what we've seen in our review, we believe that our goal for the next cohort of Wabash National Study schools should be that 90% of the institutions are engaged in sustained assessment activity in the third year of the project, and that half of the institutions would have demonstrable improvements in student learning, or in the conditions and practices that promote student learning, in five years. Although this may be disappointing to some, we believe that it is ambitious in light of the extraordinary leadership turnover at many liberal arts institutions. With so many institutional leaders moving into and out of positions, and the shifts in institutional and programmatic priorities those moves bring, it is difficult for us to imagine sustained institutional focus on using assessment to improve student learning for a higher proportion of schools.

Final thoughts

We've spoken about the high impact of collaborative activities, but we haven't spoken about one of the critical mechanisms that supported many of those activities—the Teagle Assessment Scholars. The Teagle Scholars have helped to facilitate meetings, visit campuses, give presentations, build and implement workshops, and create the very kind of face-to-face formative interactions that have supported the good work of many of the institutions in the collaboratives. Assessment advocates have spoken blithely about the goal of creating cultures of evidence without specifying quite how this might happen. The Teagle Scholars have engaged in the work of collaborating with people at institutions to create the first glimmers of this kind of culture. Looking back, it is remarkable that it is the ordinary business of Teagle Scholars to help institutions that often compete with their own for students. We are not asserting that Teagle Scholars are the key ingredient in the successful examples of these collaboratives. We do believe, however, that given stable institutional leadership, faculty who are interested in using assessment data for change, and assessment leaders who are willing to consider multiple forms of evidence, Teagle Assessment Scholars can provide the critical "activation energy" necessary to move those institutions from having the potential to use assessment formatively to actually doing so.

Furthermore, we believe that one factor that is critical to the success of Teagle Assessment Scholars is that they are almost all former faculty from the arts and sciences. They understand faculty values and can frame assessment as a process that is connected with deeper faculty values. But we have also become concerned that there is a risk to wedding assessment so intimately with faculty values.

We began this review by talking about two different conceptions of assessment. Both of these conceptions included the idea of investigating when, how much, and under what circumstances students are learning. But only one added the idea of using the results of that investigation to make improvements. By the easier measure, half of the institutions in these two collaboratives have been successful—by the more stringent, only a few.

One of the organizing principles of these two collaboratives was the idea that assessment, like scholarship, is driven by intellectual inquiry, and that helping faculty understand this point would make assessment more acceptable. But what are the risks of tying assessment to the idea of inquiry, as it is understood by liberal arts faculty? The risks are connected to the difference between the two competing definitions of assessment with which we started—is assessment about inquiry, or is it about inquiry and action?

It is important to acknowledge that faculty across arts and sciences disciplines have established, and cherished, ways of both developing and sharing knowledge. Faculty have conferences, write articles, and review one another's work. They are passionate about creating a deeper understanding of complex phenomena and processes and about sharing that understanding with colleagues. The social and intellectual way of life that many of us first learned as undergraduates, honed in our graduate studies, and have continued to refine over the course of our career is dedicated to learning, talking, arguing, contesting, and writing about what we learn. However, for most of us in the liberal arts, it is not dedicated to creating practical changes in the world outside of our scholarly exchanges. More than being accidently cloistered away from the world of practical concerns, many of us actively disdain the possibility of engaging in "applied" research.

It should not be surprising then, that "faculty-friendly assessment" can easily become a process that captures not just the beneficial qualities of scholarly inquiry, but also comes to mirror the qualities of faculty intellectual life. "Successful assessment" will then be something that

generates knowledge, creates conversations, allows for meetings and publications, sparks constructive disputations, but is short on action when it comes to improving student learning.

This is, of course, not a new tension. There is a difference between knowing about and knowing how. The movie *The Matrix* pivots on this point. As the film's hero Neo struggles to move from understanding to acting on the fact that everything he's known and experienced is part of a computer-constructed reality, his tutor Morpheus says, "Neo, sooner or later you're going to realize just as I did that there's a difference between knowing the path and walking the path." Drawing from Gilbert Ryle and Jerome Bruner, Brown & Duguid (2000) make a similar distinction between "learning about" and "learning how" saying, "Learning about involves the accumulation of 'know that': principally data, facts, or information. Learning about does not, however, produce the ability to put 'know that' into use. . . . We learn how . . . by practice (p. 128)."

Improving the rate at which institutions move from gathering data and reflecting on assessment evidence to acting to make demonstrable improvements in student learning will be contingent on developing explicit strategies to help convert the new-found faculty interest in inquiry about how and when their students are learning into a willingness to navigate the occasionally painful and always messy process of changing the courses, programs, and other institutional processes that assessment evidence challenges. Faculty, staff, and administrative appetite for working with their colleagues on inquiry is far greater than their appetite for working with colleagues on change. Since the formation of these two collaboratives, we have learned a great deal about traveling the path of inquiry about student learning. We now have to develop the practical wisdom with which to help institutions take the path of applying that knowledge to improve.

Some suggestions for new assessment collaboratives

In closing, we will suggest several tactics for new assessment collaboratives, based on our experiences working with the Hampshire and Wabash consortia, as well as with other institutions:

1. Adopt a two-phase, six-year timeline for working with institutions—three years to actively engage, discuss, and review assessment data in order to identify possible changes, and

three years to institute and assess those changes. The activities in collaboratives should be designed to help institutions move in small steps through these two phases.

2. Adopt specific targets to evaluate the success of collaboratives. For example, in the next round of the Wabash National Study we would aim for 90% of the institutions to be engaged in sustained assessment activity by the third year of the project and for half of the institutions to have demonstrable improvements in student learning, or in the conditions and practices that promote student learning, in five years. It may also be useful to make funding and support contingent on the level of activity in which institutions engage during the course of the project. This would allow limited resources to be directed at institutions that show the most potential for change.

3. Devise mechanisms for increasing the level of administrative commitment and support for institutions that enter new assessment collaboratives.

4. Create a process for helping institutions audit the assessment information they have in place at the start of a collaborative and then engage that evidence in subsequent work on the assessment/change project.

5. Develop more support mechanisms for helping institutions move from gathering to using data and for preventing them from entering the recursive cycle of "gathering more data without acting."

Center of Inquiry
WABASH COLLEGE

References

Banta, T. W. (2009, March/April). *Assessment Update*, *21*(2), 3–4.

Brown, J. S., & Duguid, P. (2000). *The Social Life of Information.* Boston: Harvard Business Press.

Consortium on Financing Higher Education. (n.d.-a). *Assessment: A fundamental responsibility*. Retrieved July 28, 2009, from http://www.assessmentstatement.org/index_files/Page717.htm.

Consortium on Financing Higher Education. (n.d.-b). *Schools endorsing the statement*. Retrieved July 28, 2009, from http://www.assessmentstatement.org/index_files/Page1402.htm.

Palomba, C. A., Banta, T. W., & Associates. (1999). *Assessment Essentials*. San Francisco: Jossey-Bass.

Schuh, J. H., Lee Upcraft, M., & Associates. (2001). *Assessment Practice in Student Affairs: An Application Manual*. San Francisco: Jossey-Bass.

Teagle Foundation. (n.d.) *Outcomes and assessment*. Retrieved July 28, 2009, from http://www.teaglefoundation.org/grantmaking/education.aspx.