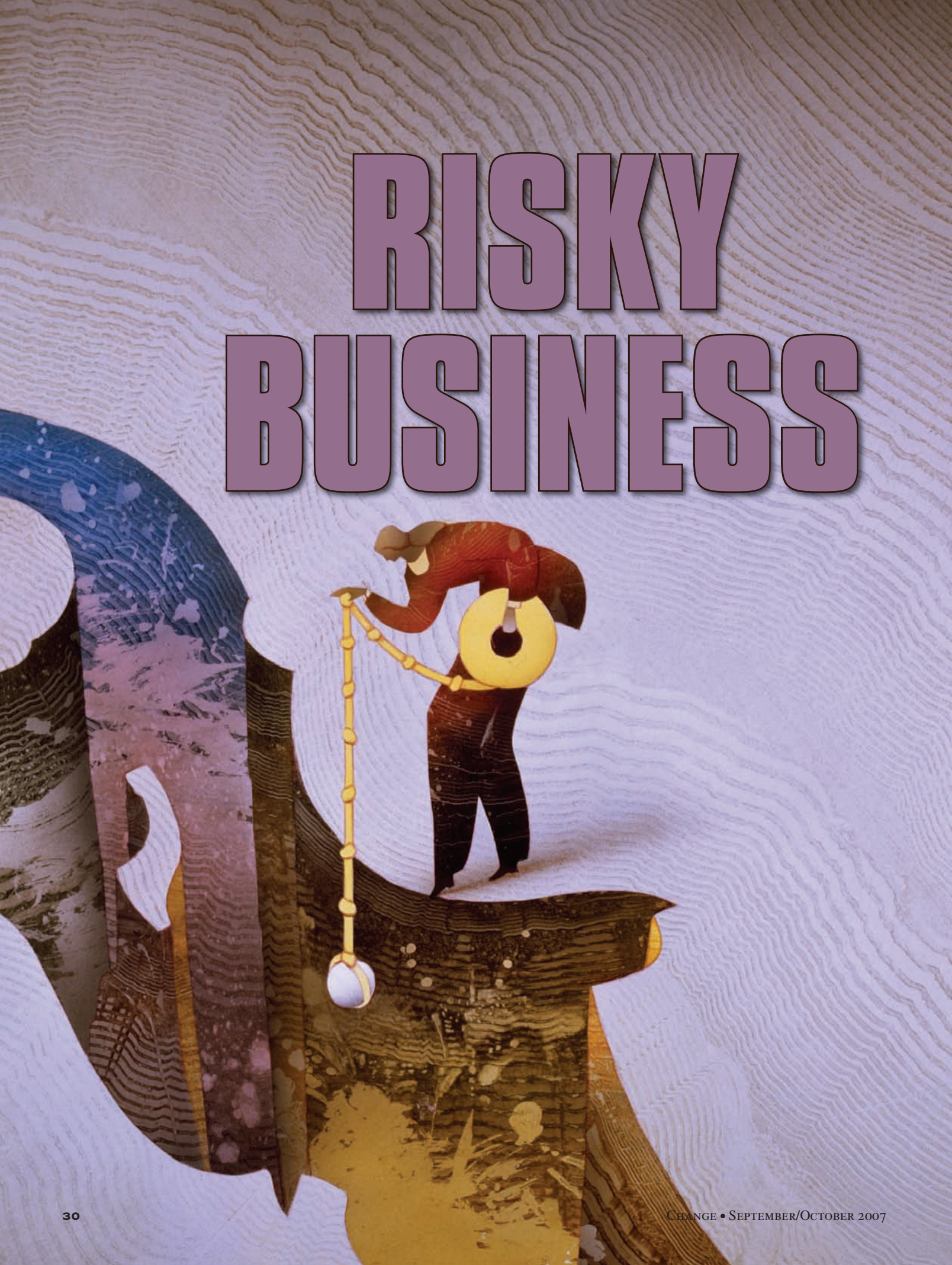
RISKY BUSINESS



Promises and Pitfalls of Institutional Transparency

By George D. Kuh

The heat is on. After a year of public hearings and not-so-private debate, the National Commission on the Future of Higher Education last year proposed six sweeping recommendations to improve "the less than inspiring realities of postsecondary education" in the United States (see *A Test of Leadership,* Resources). One key recommendation was creation of a "consumer-friendly information database" on issues such as cost, price, and student success, to enable prospective students to compare colleges and universities in order to make informed decisions about where to attend college. According to Secretary of Education Margaret Spellings, who formed the commission, the goal is to provide answers to the kinds of questions typically asked when consumers make major purchases: How much does it cost? What are the financing options? How does the "product" perform compared to others like it?

Reasonable people may disagree about whether something akin to a J.D. Powers automobile "rating" system can do justice to the multidimensional performance of postsecondary institutions and their students. But what is all but certain is that some form of a common reporting template will be coming soon to an institution near you.

Public reporting about various aspects of institutional performance is long overdue. But as with any new initiative, the technology can get ahead of the public's capacity to use it responsibly and productively. To maximize the benefits and minimize the possible mischief of making public heretofore unreported information about student and institutional performance, I offer some observations about the appropriate and inappropriate uses of a common reporting template that makes institutional-performance data public.

George D. Kuh is Chancellor's Professor and director of the Center for Postsecondary Research at Indiana University Bloomington. The center administers the National Survey of Student Engagement (NSSE) and related projects.

How We Got Here

Some brief background may be helpful. The National Commission on the Future of Higher Education met six times between October 2005 and August 2006, conducted two public hearings, commissioned several papers, and invited testimony on selected topics. One recurring theme in its deliberations was higher education's aversion to transparency and accountability.

Anticipating some stern words from the commission, the National Association of State Universities and Land-Grant Colleges (NASULGC) released a discussion paper at the commission's April 2006 public hearing entitled "Improving Student Learning in Higher Education Through Better Accountability and Assessment." The NASULGC paper set forth some principles for a "Voluntary System of Accountability" (VSA)—an effort to provide timely, accurate, institution-specific information about costs, learning, and other educational outcomes via a common reporting format—which the American Association of State Colleges and Universities (AASCU) subsequently endorsed. (For a description of the VSA, go to http://www.nasulgc.org/vsa-8-31-06.pdf) A presidential advisory committee and six task forces and working groups made up of representatives from NASULGC and AASCU member institutions immediately began to design the accountability system.

Meanwhile, the National Association of Independent College and Universities is also developing a template that will allow its member schools to present comparable information—as is the Association of American Universities, which announced that its members will identify performance indicators to be reported publicly at some future point. While the Education Department's National Center for Education Statistics considered expanding its College Opportunities Online Locator (COOL) Web site to include institutional-performance indicators, the associations' timely initiatives make this less likely.

Now that colleges and universities are on the verge of becoming more transparent about their results, what's the worry? Well, to balance the demands of public interest and institutional autonomy, we need to determine the legitimate applications of common-reporting templates and ameliorate to the extent possible any problematic aspects of making data public. It is especially important to discourage unacceptable uses of information.

Purposes, Promises, and Pitfalls

A common reporting template is intended to serve three general purposes:

• *Improvement*. If students are not acquiring the knowledge, skills, and competencies the institution values and that other institutions serving similar students have been able to impart, how can the institution change its practices to make that happen?

• *Transparency*. Is information about costs, financial aid, student performance, and so on accessible to those who have an interest in it, and do the data enable them to compare institutions with similar missions and that serve similar students?

• *Accountability*. Does the institution provide policymakers with assessment findings of students' learning and experiences as well as other information about institutional performance? For instance, are students well prepared for 21st-century challenges and opportunities, and are they graduating in a timely fashion?

The indicators used in the template should be sensitive to the institution's mission and educational purposes and point to actions faculty and staff can take to improve teaching, learning, and student engagement in educationally purposeful activities. In addition, the information should make it possible to judge whether resources are used appropriately and efficaciously.

Toward these ends, common reporting templates will likely employ institution-level data in the form of raw or adjusted average scores, depending on the measure. For example, institutions may report the proportions of undergraduate students who complete baccalaureate study in four, five, or six or more years, or they may report average scores for students completing tests such as the Collegiate Learning Assessment (CLA) or the Measure of Academic Proficiency and Progress (MAPP). Whether certain performance indicators are appropriate, though, will depend on context and the quality of the data used.

Be Careful What You Wish For

While improvement, transparency, and accountability are desirable ends, data can always be misused and misinterpreted. There are two categories of untoward possibilities: the problematic and the unacceptable.

Problematic possibilities. One advantage of using the same measures and reporting the results in a common format is to permit institutional comparisons. But simple displays based on institutional averages can be misleading if used by prospective students to "pick the right school." This is because—as Pascarella and Terenzini (1991, 2005) concluded—individual student performance typically varies much more *within* institutions than average performance does *between* institutions. This is true at every level of education.

To balance the demands of public interest and institutional autonomy, we need to determine the legitimate applications of common-reporting templates and ameliorate ... any problematic aspects.

To illustrate: Results from the National Survey of Student Engagement (NSSE) shown in Table 1 reveal that for all but one of 14 NSSE scales for both first-year and senior students, less than 10 percent of the total variance in student engagement is between institutions. The remaining variance—in several instances more than 95 percent—exists *at the student level within* a college or university.

These results mean the average institutional scores on NSSE scales—and likely as well on tests of critical think-

ing, problem-solving, and other learning measures—may say little about how an individual student with certain characteristics will perform.

In this regard, Shulman's questions about measuring higher-order intellectual functioning are instructive (*Change*, January/February, 2007, p. 22): "Is this an institution where students of one particular ethnic background score well across the categories while others do well only in knowledge acquisition but not in the higher-order achievements? Or is this a college where those majoring in the sciences flourish while those studying the humanities flounder?" To draw inferences about student performance as Shulman recommends requires having assessment data from enough students to reliably disaggregate the results and calculate scores for different groups—for example for women and men majoring in science, technology, engineering or mathematical fields; for African-American men and women living on and off campus; or for first-year students with different levels of pre-college academic achievement, among others.

Another potential problem is the simplification of student experience and institutional performance by comparing schools on only one or two indicators. For example, emphasizing graduation rates while ignoring measures of student engagement or learning may result in misleading conclusions about institutional performance. Is it acceptable for a school to have a high graduation rate but low engagement and educational outcomes scores? Or are individual and public interests both better served by institutions where students are academically challenged and demonstrate they have developed skills and competencies at a high level, even if fewer make it to graduation? Strong performance on engagement, achievement, and graduation measures are certainly not mutually exclusive, but each says something different about institutional performance and student development.

Another example: Many students who leave one institution ultimately complete their degrees elsewhere. Traditional determinations of graduation rates thus often produce warped, incomplete, and indefensible data when used to evaluate an individual institution's performance. To illustrate, 70 percent of the baccalaureate-degree recipients in the 2005-2006 academic year at the University of Texas-El Paso were not counted in the institution's official graduation statistics, which are based only on the first-time, full-time-enrolled student cohort. Thus, neither part-time students who started there nor transfer students are counted. Multiple indicators that more accurately reflect the behavior of today's students are needed to yield a balanced picture of student and institutional performance.

Unacceptable possibilities. Information from a common reporting template must not be used to rank institutions. Ranking systems currently privilege institutional reputation and resources, which are unrelated to what students do in, or gain from, college. Now some argue that if indicators such as student engagement, persistence, graduation rates, and learning outcomes were incorporated into rankings, this major problem would be solved. But others remain.

First, rankings *de facto* reduce rich, complicated, and often distinctive patterns of student experiences and outcomes to a single number, masking much of what is informative about variations in student and institutional performance. Second, rankings tend to exaggerate differences between institutions. This is the case because for many dimensions of performance, the actual difference between a school ranked in the top 10 and top 50 is trivial. Third, rankings divert institutional energy away from improving teaching and learning and redirect it to artificially improving the institution's position— for example, by inflating the number of applications for admission so as to appear more selective. Ranking schools based on data contained in a common reporting template could further intensify their efforts to attract "the best [i.e., the most privileged] students," who are more likely to perform better on various measures.

Another unacceptable application is to use publicly reported data to justify funding cuts. Institutions need several years to understand and determine the proximate causes of their performance, to design and implement initiatives to

Table 1. Percentage of Variance Between Institutions on National Survey of Student Engagement Scales

NSSE Measurement Scales	First-Year	Senior
Level of Academic Challenge	8.71%	5.94%
Active & Collaborative Learning	6.47%	4.53%
Student Faculty Interaction	5.00%	8.37%
Enriching Educational Experiences	8.10%	17.67%
Supportive Campus Environment	6.17%	6.72%
Deep Learning	5.65%	5.00%
Higher Order Thinking	3.75%	2.84%
Integrative Learning	5.23%	5.12%
Reflective Learning	4.14%	3.38%
Overall Satisfaction	5.77%	5.90%
Quality of Relationships	6.20%	6.63%
Gains in Practical Competence	4.15%	3.82%
Gains in Personal and Social Development	6.29%	7.71%
Gains in General Education	4.10%	5.83%

Source: Center for Postsecondary Research, Indiana University Bloomington

improve that performance, and to assess the impact of their efforts. Penalizing schools may also contribute to efforts to "game the system," one of the untoward responses to rankings noted earlier, and this can disadvantage institutions that are willing to examine their performance with an eye toward improvement.

For example, it would be counterproductive if an institution's funding were cut because of comparatively poor performance (accountability) the first time it released certain data, since the school might well become reluctant to report further data (transparency), which could in turn have a dampening effect on collecting the information needed to enhance aspects of teaching and learning (improvement). Great care must be taken to make sure institutions are encouraged and rewarded for using performance data in ways that enlarge their capacity to understand, learn from, and make public ways in which their performance is distinctive, as well as acknowledging aspects that may need improvement.

Responsible Use of Common-Reporting Templates

Adhering to three principles will help maximize the benefit and minimize the potential mischief of making performance information public.

1. *Focus on educationally meaningful indicators that are linked to student success in the context of the institution's mission.*

The indicators an institution selects will determine how judgments about its performance will be made, and hence they will almost certainly signal which activities will receive additional resources and attention. So careful thought must be given to what the institution is trying to accomplish, the desired outcomes for students, and whether the indicators chosen accurately represent student and institutional performance.

Toward this end, common reporting templates should feature measures of student and institutional performance that are demonstrably linked to the desired outcomes of college (such as those distilled by Chickering and Gamson), to graduation rates, or to other measures of student performance. These will differ from college to college, depending upon mission, educational purposes, and student characteristics. For example, engineering colleges, women's colleges, and art-and-design institutes may appropriately opt for different engagement and outcomes measures, given their varied educational purposes.

When institutions are grouped for comparison, the selection of the schools that comprise the groups must be explained so that users can draw their own conclusions about the merits of the comparisons. To help users determine if apparent differences between schools are statistically significant and practically meaningful, means and standard deviations for institution-level scores must be included, including confidence bands based on population estimates for samples of students. Such statistical measures and their significance must be carefully explained in laymen's language.

The present
moment provides
an opportunity to
educate the public
about what higher
education does,
as well as
about *how well*
it does those things.

As my colleagues and I explained in *Piecing Together the Student Success Puzzle*, higher education needs additional indicators of student performance today because of the increasing diversity of college students and their multiple, winding pathways to a baccalaureate degree. These indicators should include attainment of individual students' goals, course completion, and transfer rates. Other indicators of students' success could include performance in course work in the major, fall-to-fall persistence, degree/certificate completion, student satisfaction, student personal and professional development, student involvement and citizenship, and learning outcomes such as writing proficiency and mastery of major field content. Also relevant are post-collegiate outcomes such as post-baccalaureate degree completion, suitable employment, job performance, and a capacity for lifelong learning. Measures of success in subsequent coursework are especially important for students who have been historically underrepresented in specific majors and for institutions that provide remedial education.

2. *Evaluate the quality of the data on which performance indicators are based.*

A performance indicator is only as good as the data on which it is based. To assure the public that what is being reported about student and institutional performance is trustworthy, we must expect transparency and accountability not just from colleges and universities but also from the organizations that provide them with measurement tools and data. Because the stakes are higher than ever, it would be helpful to have neutral, third-party evaluations of the validity, reliability, and utility of various measures, akin to the analyses produced by the National Center for Higher Educational Management Systems in the early 1990s (Ewell & Jones, 1996). The necessary first step is for testing companies to make their instruments available for such evaluations.

The indicators also should be presented in clear, understandable formats with sufficient detail and explanations of the types of students, behaviors, and institutional characteristics and actions the indicators represent and what they *do not* represent, as well as what can and cannot be concluded from them. Institutions should justify any statistical adjustments made for student and institutional characteristics that they consider to be beyond the direct influence of the institution, such as students' background characteristics, institutional selectivity, the percentage of students attending full-time and living on campus, and so forth.

3. *Use performance indicators appropriately.*

The present moment provides an opportunity to educate the public about what higher education does, as well as about *how well* it does those things. Especially useful toward this end is the

elegant, informative statement from the Association of American Colleges and Universities (2007) about what the college graduate in the 21st century needs to know and be able to do, based on information collected from business, governmental, and educational leaders, and recent college graduates.

Some indicators may be of greater interest to certain groups than others or be used in ways different from their intended purpose. Thus it would be helpful to clarify the audience(s) for which the indicator is intended: prospective students and their families, current students, alumni, government agencies, accreditors, and/or others.

Also, if indicators are to be used to evaluate institutional performance, it is essential to establish sound bases for such comparisons. For example, to determine if a college has improved its graduation rates requires a longitudinal data base, while other intra-institutional comparisons might use the school's "best performance" on certain indicators (e.g., persistence or student-faculty interaction) as the benchmark. Comparisons among schools might fairly be based on average scores on a certain metric for a given year or "industry standards" set by an accreditor or other group.

Common reporting templates make information about institutional progress toward desired levels of performance more convenient and accessible. For this reason, it makes sense to feature indicators that institutions can influence. Most colleges and universities have little direct control over the types of students who matriculate, short of changing their mission or increasing the number and amounts of merit-aid awards. But any institution can enhance student engagement through the use of effective educational practices. Every institution should explain its performance—not to make excuses but to take its fair share of responsibility for the outcomes it reports—as well as the ways it intends to improve aspects of its performance that it can control to some degree.

Finally, institutions—or better, a consortium, state system, or other entity representing a set of institutions with similar missions and educational programs—should provide a succinct, easily digestible guide to the responsible use of the information contained in a common reporting template. That guide should include a strong, unequivocal statement that eschews the use of institutional-performance indicators in a ranking system. It would be unfortunate for the media and others to draw conclusions about educational quality and the meaning of various performance indicators without having a context for making such judgments.

Final Thoughts

Common to the 20 colleges that we studied several years ago that performed unusually well on the National Survey of Student Engagement was an ethic of positive restlessness, or a bent toward continuous improvement. Like many others, these schools did self-studies and benchmarked their performance against one or more sets of peer institutions. But two things set these schools apart. First, they focused on student learning and on facilitating the conditions for it. Second, they used the information they collected to guide policy and improvement, recognizing that collecting and reporting information is a hollow exercise otherwise.

Under the leadership of NASULGC and AASCU, along with encouragement and support from other groups, many more institutions are amassing the kind of information that can become a road map for improvement. If used properly, this information can be instructive for prospective and current students, policymakers, government officials, and other constituents of higher education.

But balancing the multiple purposes of a common reporting template will be challenging. Many institutional leaders have little experience talking publicly about data that represents the core of their school's performance—about what actually happens to students in classrooms, laboratories, studios, practice fields, and beyond. But with practice and patience, we will all get better at deciding what to measure, how to measure it, and using what we learn to improve the quality of the undergraduate experience and other aspects of institutional performance.

Efforts such as the Voluntary System of Accountability and similar initiatives comprise a grand experiment. Therefore, above all, we need to try different approaches, do no harm, and refuse to allow pundits to declare winners and losers. There is too much at stake.

Resources

- *A Test of Leadership*, http://www.ed.gov/about/bdscomm/list/hiedfuture/reports.html
- Association of American Colleges and Universities. (2007). *College learning for the new global century: A report from the National Leadership Council for Liberal Education and America's Promise*. Washington, DC: Association of American Colleges and Universities.
- Ewell, P.T. (1999). Linking performance measures to resource allocation: exploring unmapped terrain. *Quality in Higher Education*, 5(3), 191-209.
- Frank, K.A. (1998). The social context of schooling: Quantitative methods. *Review of Research in Education 23*, 171-216.
- Kuh, G.D., Kinzie, J., Buckley, J., Bridges, B., & Hayek, J.C. (2007). *Piecing together the student success puzzle: Research, propositions, and recommendations*. ASHE Higher Education Report, 32(5). San Francisco: Jossey-Bass.
- Kuh, G.D., Kinzie, J., Schuh, J.H., Whitt, E.J., & Associates (2005). *Student success in college: Creating conditions that matter*. San Francisco: Jossey-Bass.
- Kuh, G.D., & Pascarella, E.T. (2004). What does institutional selectivity tell us about educational quality? *Change*, 36(5), 52-5.
- McPherson, P., & Shulenburger, D. (2006, August). Toward a public universities and colleges Voluntary System of Accountability for undergraduate education (VSA): A NASULGC and AASCU discussion draft. Washington, DC: NASULGC. http://www.nasulgc.org/vsa-8-31-06%20_7_%20_2_.pdf
- Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco: Jossey-Bass.
- Shulman, L.S. (2007). Counting and recounting: Assessment and the quest for accountability. *Change*, 39(1), 20-25.