AACRAO®
1910

C&U

COLLEGE *and* UNIVERSITY
Educating the Modern Higher Education Administration Professional

Rating (Not Ranking) the Undergraduate Experience
Principles from a National Discussion

Rehabbing the Rankings
Fool's Errand or the Lord's Work?

Toward Devising Measures of Quality and Effectiveness Across All Institutions

In Search of a Better Mousetrap
A Look at Higher Education Ranking Systems

forum

In SEARCH of a BETTER MOUSETRAP

A LOOK AT Higher Education Ranking Systems

College rankings create much talk and discussion in the higher education arena. This love/hate relationship has not necessarily resulted in better rankings, but rather, more rankings. This paper looks at some of the measures and pitfalls of the current rankings systems, and proposes areas for improvement through a better focus on teaching and learning and workforce outcomes for graduates.

By Watson Scott Swail
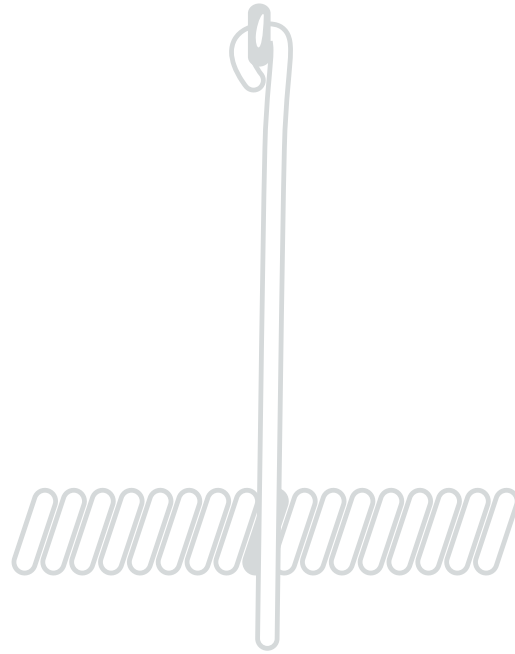
Domestic and international rankings of higher education continue to evolve and garner greater interest from educators, administrators, policymakers, students, and parents. With each annual release of rankings from *U.S. News & World Report* (*U.S. News*), *MacLean's* (Canada), Times Higher Education, QS World University Rankings, and Shanghai Jiao Tong University's Academic Ranking of World Universities comes increased media awareness and scrutiny on the meaning and impact of these data.

The information used to produce these ranking instruments is hamstrung by relatively limited availability of accessible, reliable, and comparable information. As Kuh (2011) describes in a recent paper, ranking systems focus primarily on inputs (*e.g.*, freshman SAT) rather than outputs (*e.g.*, job placement). Measuring the quality of an institution via what goes in rather than what comes out certainly limits the efficacy of such analysis. But until additional data on student outcomes are made available, the true utility of these rankings will remain suspect.

Of course, while these methodological issues are important to academics and researchers, the true consumers of this information—most notably students, parents, and policymakers—find these methodological details simply to be ignorable background noise; consumers are most interested in knowing which schools are the best. In a complex web of higher education systems, ranking in a hierarchical manner is seemingly of limited utility. However, to consumers of these goods, rankings are rich data that impact decisions, money, and policy.

I argue that the institutional ranking process is better seen as an analytical game than as a tool of great utility for public policy and/or college choice. Regardless, institutional rankings are not likely to disappear any time soon. In fact, it is more prudent to suggest that rankings will grow in use and importance over time. Given that reality, the purpose of this paper is to provide reflection on current ranking systems and to serve as a foundation for discussing how to possibly improve rankings and ensure greater validity, reliability, and therefore utility.

WHAT ARE UNIVERSITY RANKINGS?

University rankings, or league tables as they are often referred to internationally, are mechanisms that use available information to rank order institutions of higher educa-

tion based on criteria defined by the ranking organization. The purpose of ranking systems is to quantify—down to a single number—the relative quality of institutions. The process of reducing institutions to one number makes most of researchers and academics cringe, because we fully understand that the complexities of institutions of higher education cannot be boiled down to a single 2- or 3-digit number. Even the Times Higher Education Supplement, producer of the *World University Rankings*, admits that higher education institutions are "extraordinarily complex organisations" and that it is "rather crude to reduce universities to a single number." (Times Higher Education 2010).[1] As a result, institutional rankings have become contentious and oft-debated in the higher education arena over the course of the last quarter century: first in the U.S. and Canada, and now encapsulating a global audience.

THE HISTORY AND RISE OF RANKING SYSTEMS

Ranking systems are not a new phenomena. In the United States, rankings have been around in some form since the 1800s (Kuh 2011), but it was the rankings developed by the *U.S. News & World Report* (*U.S. News*) in the early 1980s that truly stoked the rankings fire. Salmi and Bassett (2009) suggest that rankings grew out of an apparent need for transparency and greater levels of accountability. In truth, the creation of rankings in the United States was initially fueled by an insatiable appetite for higher education by the baby boomer generation. The growth of rankings has been further fueled by the massification of the U.S. higher education system in the mid-1900s and the emergence of baby boomers as the "helicopter parent," hovering over their children's educational achievements and future. By the early 1980s, *U.S. News* was able to capitalize on the demand for more information about colleges and universities.

Ranking systems are only able to rank institutions based on widely available data from institutions and governments. Perhaps the main reason the United States began ranking institutions early on is the vast information collected by the federal government. The U.S. Department of Education's Integrated Postsecondary Education Data System[2] (IPEDS) collects information on almost ev-

ery college and university in the nation. Now, on the international stage, the flat and global knowledge economy has produced an increased appetite for higher education rankings. In the prior era, the world traveled to the United States and Britain for higher education. However, the flattening of our new world has pushed the expansion of higher education in most industrial countries.

This global massification of higher education has resulted in a new "arms race" in the post-Cold War era: the battle for higher education supremacy. As the world began to catch up to the United States in other areas of commerce, such as manufacturing, communications, and engineering, they began to see the need for better systems of postsecondary education to create the type of workforce that could compete with the US. Despite much criticism of the United States from abroad on a variety of issues, it is a widely held belief that the system of higher education in the United States is the best in the world. Foreign governments simply put 2+2 together: a great higher education system must be linked to a great economy. And with that, the arms race began.

WHY USE RANKINGS?

I argue that the first and most important consumer of ranking information is not the student, as many suggest. In fact, data from UCLA's Higher Education Resource Institute (HERI) found that only 18 percent of students said that college rankings were important in the college choice process (Hurtado and Pryor 2011). Rather, it is the parent, in large part because they often are the monetary source for their child's education and are the driving force behind the collection of information and the weighing of variables. It is widely believe that many parents push *their* children to attend institution of *their* choice, based on *their* beliefs and knowledge, with students only tangentially involved in the decision-making process.

Of course, this does not hold true for all parents. Some parents, especially those who never went to college, are simply happy to see their child make the great leap to the postsecondary world. To them, rankings matter little. The rankings of this discussion serve the needs of a select group of parents that have (a) gone to college, (b) are more likely working in professional fields, and (c) have enough disposable income or available financial resources to pay for colleges in the elite area of the college rankings.

---

[1] *See* <www.timeshighereducation.co.uk/world-university-rankings/2010-2011/analysis-methodology.html>.

[2] *See* <http://nces.ed.gov/ipeds>.

The second stakeholder with a viable interest in rankings is university leadership. While most administrators are quick to denounce the importance of *U.S. News* and other rankings (particularly when the rankings are poor or declining), they are also the first to send out an alumni fundraising letter with the announcement of their rank on *U.S. News* when they have experienced increased or high rankings. In fact, advertising a ranking is the most broadly used method of fundraising for those schools which happen to be in the top 25 or 50 institutions of a category. I have personally held conversations with CEOs and other high-level administrators at institutions that are intently focused on raising their institution's rank in *U.S. News*. If they are 27th, they want to be in the top 25; and if they are 17, they want to be 16. To them, rankings matter because they are directly correlated with perception of excellence, which in turn correlates with increase student enrollment, offering further opportunity to raise tuition and fees, develop additional research capacity and attract resources, and garner additional government funds. In the end, rankings are about money and little else.

The third stakeholder is the policymaker. Policymakers are interested in the rankings for many reasons. Higher education is a market chip for economic growth and is a valuable commodity for research dollars and investment. There is a vested interest in having state or regional universities rank high, as they tend to spur additional technological development, corporate investment, and federal support. Although there are limited data on this issue, it is likely that rankings have had a significant impact on higher education in many states due to competition generated by the rankings.

The impact of rankings on public policy has been more dramatic outside of the U.S. China, for instance, has been very specific and open about its intentions to challenge the United States in higher education. One of their primary, stated goals is to increase the number of Chinese institutions in Shanghai Jiao Tong University's Academic Ranking of World Universities. Saudi Arabia, United Arab Emirates, India, and South Korea also represent a host of nations looking toward the rankings as a lever for economic stimulus and international competitiveness. To them, rankings matter because they want to be perceived as the purveyor of the best higher education in the world.

Recent activities suggest that they are well on their way toward this goal.

WHAT DO RANKING SYSTEMS REALLY MEASURE?

As previously mentioned, IPEDS collects information on almost every postsecondary institution in the US. In fact, in order to participate in the federal student aid system (*i.e.*, be able to provide federal grants and loans to students), institutions *must* complete the IPEDS series of surveys each year or they risk losing their ability to provide federally-sponsored student aid. This possibility is the death-knell of almost any postsecondary institution.

Other organizations that survey institutions, specifically the College Board, *U.S. News*, and Peterson's, formed the Common Data Set (CDS) Initiative in order to streamline data collection efforts and simplify the submission process for institutions. These organizations, working in concert with the U.S. Department of Education and IPEDS, share their information in order to reduce the potential burden on institutions from multiple surveys. Even the competitors in the U.S. play well together.

Although we are blessed, to a degree, with rich datasets, the Achilles heel of rankings in the United States and beyond is the sophistication of available data. This is where most criticism of rankings fall. As Kuh (2011) states, most rankings indicators are input-level data rather than output. Vedder (2008), in denouncing *U.S. News*'s ranking system, said that "They're roughly equivalent to evaluating a chef based on the ingredients he or she uses."

A brief analysis of what goes in to some of the major ranking systems helps us understand the strengths and weaknesses of these efforts. *U.S. News* creates several types of university rankings, from undergraduate, to graduate, to international. Regarding their basic undergraduate analysis, here are the categories and weights used to create their numerical index:[3]

- **Undergraduate Academic Reputation** (22.5%). Peer and professional surveys are administered to solicit feedback on the reputation of the institution. This is a viable and appropriate measure, but is also based entirely on subjective data.

---

[3] For brevity, only weights for "National Universities" are provided. *See* <http://www.usnews.com/articles/education/best-colleges/2010/08/17/how-us-news-calculates-the-college-rankings.html?PageNr=4 for additional information>.

- **Graduation and Freshman Retention** (20%). Uses freshman retention rate (fall-to-fall) and six-year graduation rates, via IPEDS data. Arguably one of the better indicators used in *U.S. News*, but hamstrung by not providing any consideration of student academic ability.
- **Graduation Rate Performance** (7.5%; for National Universities and National Liberal Arts Colleges only). This is an interesting calculation developed to measure whether an institution does "better" than expected by comparing predicted versus actual graduation. An important measure to balance the gross data provided through IPEDS because of the variance of institutions and students.
- **Faculty Resources** (20%). Indicators include class size, faculty pay, and faculty degree status (terminal). Because so many factors impact these areas, this is a weak indicator. As stated previously, there is limited research supporting class size, but using it as a proxy for how much human resource is available per student is of some interest.
- **Student Selectivity** (15%). Completely input-based data on ACT and SAT test scores, high school class ranking, and acceptance/admit rates. This indicator illustrates how attractive an institution is by the apparent "quality" of the student who attends.
- **Financial Resources** (10%). A calculation of spending per student, which again is a proxy for the level of service provided to students. Meaningful, but input based.
- **Alumni Giving Rate** (5%). This indicator is meant to serve as a proxy for student satisfaction by the percentage of alumni that give back to their alma mater. In lieu of the limited information available, this is interesting but extraordinarily weak. Many institutions have figured out how to game this indicator by automatically creating alumni contributions through special fees. Thus, it becomes of less utility for the rankings.

On the international level, *U.S. News*, Times Higher Education, and other ranking systems utilize similar indicators as posted above. As described, the U.S. analysis benefits greatly from the availability of data from IPEDS data and the Common Data Set. The Canadian rankings, conducted by *MacLean's*, is severely handicapped in comparison to *U.S. News* because of the lack of similar data. In Canada, the federal government does not collect information like IPEDS. International rankings, by comparison, are even more limited by data since the common denominator for analysis is reduced to only those data universally available at the institutional level.

U.S. News, in its *World University Rankings* analysis, for instance, focuses on data that measure the following (with the subsequent weights):
- Academic Peer Review (40%)
- Employer Review (10%)
- Student-to-Faculty-Ratio (20%)
- Citations per Faculty Member (20%)
- International Faculty (5%)
- International Students (5%)

Similarly, the Times Higher Education Supplement's ranking system gives one-third of its rankings weight to published citations of faculty, 30 percent on research indicators, and 30 percent on teaching.

These indicators beg several questions: Is it clear that having international faculty necessarily makes a school a better place to learn? Does having a higher percentage of institutional (foreign) students improve the outcomes of students? Does having a lower student-to-faculty ratio illustrate a better learning environment, even though there exists no significant research suggesting that to be the case? What do peer and employer reviews really tell us about an institution? And finally, do the number of citations per faculty member provide an accurate measure of institutional quality, or just how much focus faculty spend on publication rather than teaching?

Current ranking systems utilize mainly input measures such as institutional resources (*i.e.*, faculty salaries, library resources, number of faculty with terminal degrees), but with the exception of graduation rates and, in the case of *Money m*agazine, first-to-second-year persistence rate (Kuh 2011), very few ranking systems include indicators of student performance and learning.

If the primary purpose of the university is to provide an educational vehicle for students, shouldn't the education of those students be the primary indicator of institutional quality?

BUILDING A BETTER MOUSETRAP

U.S. News takes considerable abuse for what they do. I argue that most of this is undeserved. The magazine is not the "bad apple." Rather, *U.S. News* has simply cre-

ated instruments and information based on data that are both universal and available. In 2010, I attended the annual Council of Higher Education Accreditation (CHEA) conference in Washington, DC, where Bob Morse of *U.S. News* received significant criticism for their rankings system. Morse shot back, "If you can make a better system, make it better." To be fair, *U.S. News* has been very inclusive in the development of their system. They are misrepresented as the Wizard behind the curtain, hiding all their cards from the public, which just isn't so.

Unless the type of data collected changes significantly, the evolution of college rankings will be stagnant. "In fairness, in order to include meaningful measures of desired learning outcomes in their algorithms, ranking outfits need valid, reliable data from large numbers of colleges and universities that have the same or comparable measures," states Kuh (2011, p. 16). Without an injection of new information, there is very little that can improve the ranking systems.

If we want a better ranking system, what type of information do we need? If we wish to move to an output-based ranking of higher education, what type of data will provide us with more valid indicators of institutional excellence and success? I suggest two major areas for consideration and exploration.

Quality of Teaching and Learning

Currently, "quality" in *U.S. News* is quantified via surveys of peers and professionals, which are, to a degree, useful indicators. But there are no indicators on the absolute quality of how teachers teach and how students learn. However, there have been several recent efforts to collect data domestically and internationally to rectify this omission.

On the domestic side, the *Collegiate Learning Assessment* (CLA), developed by the Council for Aid to Education (CAE), which, at the time, was a subsidiary of the RAND Corporation, is an effort to quantify learning on campus. The CLA is essentially a student-level inventory to measure the "critical thinking, analytic reasoning, problem solving, and writing skills of college and university students" (CLA 2010). The purpose is so that schools can see how their students, as a group, compare to students at other schools. CLA also builds in professional development and support activities to help institutions and departments improve their teaching practices. To date,

over 400 institutions have worked with the CLA. Lumina Foundation for Education recently funded a longitudinal study of the CLA,[4] and there is also an instrument being developed for community colleges.

Other domestic data collection efforts include ACT's *Collegiate Assessment of Academic Proficiency*[5] (CAAP), which measures student academic achievement on a nationally normed basis, and the National Science Foundation's *Critical Thinking Assessment Test*[6] (CAT).

A recent development generating much discussion is OECD's AHELO Project. AHELO (Assessment of Higher Education Learning Outcomes), supported in part by Lumina Foundation for Education, is being designed to measure student learning to inform universities, students, policymakers, and employers about quality of teaching and learning. The instrument to measure student learning will include emphasis on generic skills (*e.g.*, critical thinking, analytical reasoning, problem solving, written communication); discipline-specific skills (in economics and engineering); and contextual information (*e.g.*, institutional indicators, such as equipment and facilities, research, etc.).

The development of the AHELO metrics is currently underway, with a pilot of 150 institutions in 15 countries slated to begin in July 2011. If the pilot is successful, OECD will consider what they call a "full-scale AHELO." The American Council on Education (ACE), the American Association of Colleges and Universities (AACU), and the Council for Higher Education Accreditation (CHEA) are all sponsors of AHELO.

Whether any or all of these developments impact college ranking systems remains to be seen, but these are steps down the appropriate path. If we want to truly measure institutional quality, we need to measure teaching and learning.

Workplace Indicators

For the most part, workplace indicators, such as earnings and employment status of former students, are not part of any ranking efforts. But if we want to measure the ultimate output of higher education via success in the workforce, we need to add these types of indicators to the analysis.

[4] *See* this article summarizing the longitudinal findings: <www.collegiatelearningassessment.org/files/CLA_Lumina_Longitudinal_Study_Summary_Findings.pdf>.

[5] *See* <www.act.org/caap>.

[6] *See* <www.tntech.edu/cat/home/>.

U.S. News does provide some level of information in their Best Graduate Schools rankings, depending on the discipline. For instance, in their analysis of business schools, *U.S. News* is able to collect average starting salary and employment rates. Similarly, their law school analysis uses employment rates of graduates and bar passage rates. However, they have no similar indicators for graduate schools in other schools such as education or engineering.

Ultimately, we need indicators such as those used in the business school analysis. It would be helpful to know the percentage of students who gain employment after graduating from a school, the type of employment (*e.g.*, full- or part-time), and also whether it is in a field relative to the individual's degree.

CONCLUDING THOUGHTS

There are surely many other areas that we could focus on in developing new and better rankings of higher education institutions, but our focus should stay on the development of better outcome indicators so we can use input variables only for clarifying analysis. Further development and collection of data that enhances our understanding of the learning process at an institution and what students do post-graduation are important for all consumers of rankings data.

The remaining challenge is providing greater utility of rankings. The major rankings systems are all static. That is, they are represented by a number in a list. The next generation of rankings needs to be more flexible, allowing students, parents, and others to manipulate data based on their interests and needs. Canada's Globe and Mail newspaper created the "Campus Navigator," which al-

lowed students and parents to compare institutions based on criteria important to them. This type of flexibility that provides more power to the user is important to explore.

In the end, the development and refinement of rankings systems depends on who the user is. For students and parents, it needs to provide enough information in a user-friendly manner to help with their college choice. For the administrators, it needs to provide factual, comparative information to help them improve education—rather than focusing on simply gaining market advantage. And for policymakers, the better mousetrap needs to provide details that can help craft and maneuver public policy to improve higher education for all.

REFERENCES

Kuh, G. 2011. Rehabbing the rankings: Fool's errand or the Lord's work? *College and University*. 86(4): 8–19.

Pratt, S. 2010. Robust, transparent and sophisticated. *Times Higher Education*. Retrieved from: <www.timeshighereducation.co.uk/world-university-rankings/2010-2011/analysis-methodology.html>.

Salmi, J., and R.M. Bassett. 2009. *Impact of the Financial Crisis.* Washington, D.C.: The World Bank, September 9.

Hurtado, S., and J. Pryor. 2011. Toward devising measures of quality and effectiveness across all institutions. *College and University*. 86(4): 20–27.

Vedder, R. 2008. How to choose a college. *Forbes Magazine.* May 19. Retrieved from: <www.forbes.com/forbes/2008/0519/030.html>.

About the Author

WATSON SCOTT SWAIL is the President and Chief Executive Officer of The Educational Policy Institute, a non-profit, non-governmental organization dedicated to policy-based research on educational opportunity for all students. Dr. Swail also serves as CEO of EPI International, a for-profit research institute aligned with the Educational Policy Institute.

Swail earned his doctorate in educational policy from George Washington University. He earned a Master's of Science from Old Dominion University in Norfolk, Virginia, and a Bachelor's of Education from the University of Manitoba, Winnipeg, Manitoba.