

In Pursuit of Number ONE

by John V. Lombardi

Competition is hardwired into all of us, and the essence of competition is keeping score. From commerce to sports, from celebrity popularity to national preeminence, we create scoring systems to permit the celebration of the brightest, the richest, the biggest, the highest, in short, the best. Finding the best in any domain offers endless challenges and opportunities for interpretation. Even if our scoring system identifies a relatively unambiguous best, we sometimes have difficulty recognizing the value of the distinction. Does the best economy deliver the most prosperity? Do the champions of today remain champions tomorrow? Is it better to be very good for a long time or the best in the world for a short time?

The rhetoric of competition derives in large part from the world of sports. Sports are particularly good for this sort of thing because the score keeping tends to be clearer than in other domains of human activity. The time of the 100-meter dash is a reasonably specific marker, generally recognized by all with an interest in the event. The score of the baseball game gives us a winner, and while we may consider many things to have gone wrong in the game, within the context of its rules, the score reasonably marks the best. This model, displayed in national contests and periodically in the international arena of the widely observed Olympic games, colors and influences all of our thinking about keeping score in other much less easily defined areas.

Our sports metaphor for competition is so clear that we often extend its meaning well beyond reason. We imagine that the cumulative scores on events at the Olympics indicate national success, international significance, and the continued vitality of nation-states. Sports contests, of course, identify the most transient of accomplishments, awarding a transitory title for the best that lasts only until the next contest. World records in the high jump persist only until the next high jump contest. While everyone understands this temporal difficulty of the sports-model score keeping, it nonetheless remains the standard methodology for measuring many phenomenon of much less transitory and much less definitive nature.

Academics are as competitive as the participants in any basketball or football league. They, too, want to win. They want to be recognized as the best in class. Unfortunately, the contests among academics are poorly structured, indifferently scored, and unstable in rules. Where all American college football fields are of exactly the same dimensions, the teams have exactly the same number of players, and the rules are exactly the same for the competitions of all who participate, the American academic competition is vague in its organization, highly differentiated in its fields of play,

and manifestly unclear in its scoring rules. Who is the best scholar, a historian or a chemist, a psychologist or an engineer? When we ask whether the good chemist is better than the good historian, we enter a world of arcane if not imaginary qualities.

Still the sports model tempts us with the power of its simplifying paradigms. Sports we believe provide a model, a representation of life, an opportunity to create surrogate competitions that by simplifying and clarifying life's real conflicts deliver satisfying contests that separate the winners from the losers. Unlike the life they represent, sports permit endless repetition and offer an unending hope for future, eventual triumph.

Academics in search of competitive satisfaction can look to their universities as substitutes for the sports teams they follow so enthusiastically. If Mid-State University appears to have more of the right academic qualities than Upper-State University, then we have a winner and a loser. If we can take all these university academic teams and through a virtual competition score their achievements on an annual basis, we can produce a championship winner each year in the academic world. Winning and losing, that satisfying binary result, are insufficient for leagues where we want to know not only the best but also how the rest of the completion ranks against the best. We want a rank ordered league table that tells us not only who is Number One, but also who is number 25, 30, or 50. Following our sports metaphor, we want our academic league tables to reflect a season's achievement, an annual tally that gives the satisfaction of closing the cycle of competition each year and restarting it anew the next.

This notion of opening and closing the season of competition, so familiar to the sports fans among us, serves many purposes. It encapsulates our performance, puts an end to what might have been a bad season, and creates hope for a better result in the next cycle. In sports, unlike academics, this charming conceit has a base in reality. We have constructed sports seasons to be self-contained competitions that begin with everyone equal and end with a hierarchy of winning and losing. Then, the teams readjust personnel, leadership, organization, practice, financing, and other characteristics of their competitive contexts and prepare for a new season that will begin again with everyone unbeaten.

For universities, this model of competition, however attractively simplistic it may be, does not fit. We do not have seasons, for universities are in continuous operation, constantly producing their goods and services, perennially revising and renewing their faculty, staff, facilities, and

students in a permanent, on-going process. Universities never start or stop their academic performance. When we attempt to compare our universities for the quality of their goods and services, we pick a moment in time, a place on the constant timeline of progress, change, growth, and decline that represents academic life. There are no seasons; we select an arbitrary moment for measurement.

Sometimes, to match our sports metaphor, we frame the measurement in terms of academic accomplishments in a particular time: one year, three years, whatever seems useful. These time frames give us the false security of allowing us to believe that the time frame is relevant to an institution's success, and we often forget the artificiality of the construct.

Perhaps the worst misuse of the sports metaphor is the annual competitions run by commercial publications in an attempt to display an accurate sense of the ebb and flow of academic achievement. Year one they may publish some list that marks Big-Rich University as number one and then in year two they publish a list that marks Not-So Big-but-Richer University as number one while the previous year's winner falls to number 3. Such a result, which captures headlines equivalent to the national championship in college football, means nothing of substance in the world of universities. Universities have histories of hundreds of years. Their faculty, the players on the university team, have academic careers of 20 to 40 years. Their work takes place over multiple years and produces various products on different time cycles. The notion that a season of work (a year) equivalent to the football season can identify major relative change in academic quality among significant research universities is, of course, an exaggeration.

Yet, we in the academic community, so desperate to distinguish our more or less similar institutions, the one from the other, participate in these yearly league tables and imagine that the changes we see represented in the annual publications reflect substantial improvement or decline. We often place good rankings constructed with suspect methodologies in prominent positions on our web pages and in our alumni magazines (assuming we rank high or have improved). When our ranking declines, however, the information disappears from view and we give sober interviews on the methodological failings of popular rankings.

Expansive Popularity of League Tables

So enamored are we of the league table, the ordered list of the great and near great, that it has become an international specialty. Not only do we have these tables for the United States, but also we have them for the world. World league tables ranking university quality are a particularly interesting phenomenon. While the techniques for constructing

these tables constitute a fascinating field in itself, the underlying premise of these tables is worth a review.

In the dynamic economy of our time, nations in search of global significance have fastened on the notion that scientific knowledge is one of if not the key differentiator between those nations that dominate world trade and take a leadership in global affairs and those relegated to second tier status. Science and scientific knowledge, produced by major research universities, appear in this narrative as the magical touchstones of progress, prosperity, and power. The nations with strong competitive science-based universities have the ability to create the future in their own image. They will produce the element (which for lack of a better and more precise term we call knowledge) that transforms poor nations into rich ones, emerging nations into international powers, and rapidly developing countries into societies with sustained economic achievement. The older established countries, whose historic preeminence is assumed to be based in large part on the power of their science-based university establishments, worry that their previously unchallenged excellence in academic performance may find worthy competitors in the increasingly dynamic research universities of Australia, Japan, China, Malaysia, Hong Kong, Singapore, and Brazil. To chart this competition we find the international league tables.

If the challenge of measuring university performance and quality in one country, the United States with all its institutional mechanisms for evaluating research, is difficult, imagine the complexity of international comparisons. In the end, there appears to be only one coin of the realm for international research comparisons: the publication, and in most cases, the publication in a significant peer reviewed scientific journal. This coin is of somewhat non-standard composition however. Not every publication is equal to every other publication, and so the league table needs a method of assessing the quality of this common coin. That method, implemented in complicated ways, involves a citation index. For those not immersed in the theory and implementation of academic league tables, the citation index not only counts the number of articles published by each individual, but it counts the number of times other articles cite the article in question. If my article is significant, we assume that many people will cite it when writing their article. When they do, the value of my article goes up. Then, we add up the number of articles, we weight the number of articles by the number of citations, and by that method, we achieve a score. We add up the scores for all faculty employed by my institution to get the annual score for my university. Your university gets a similarly constructed score. If my university's score is higher than yours is, then my institution beat your institution in the virtual competition for higher standing in the annual league table competition.

All this would be an interesting exercise if it were not taken

so seriously. Conditioned as we are to the immense rewards provided successful major sports teams that win their league championships, we imagine that winning the virtual annual academic university league championship (however artificially constructed) deserves a major reward too. National governments follow these league tables and rearrange, fund, and manipulate the work of their universities to score higher. Newspaper stories highlight the number of U.S. winners in the league tables compared to the number of other country winners, much as we count gold, silver, and bronze medals in the Olympic Games, as symbols of national success.

Underneath all the hoopla, self-promotion, and money-making publications associated with many league table activities, the real issues of academic research performance remain, sometimes enhanced by all the rankings but often distorted out of all recognition by the various rankings' peculiar characteristics.

Academic Research and University Competition

Academic research has a very simple definition as the intellectual work that creates and publishes new knowledge. The newness of the knowledge distinguishes research from many other intellectual pursuits of high value. Undergraduate teaching, for example, is a core activity of almost all universities, but its purpose is to convey and teach critical thinking about knowledge and to teach the process for creating and evaluating knowledge, sometimes through research-focused course work. Students learn the current state of the art, the skills and content associated with a wide range of disciplines and methodologies, and the general skills of an educated and engaged citizen. Teaching, however important for the prosperity and competitiveness of the nation, resists effective and comparative quality evaluation. Much is made of various test scores, but the methodological problems associated with these measurements, especially when used in comparative contexts across societies that organize their educational systems in fundamentally different ways, make teaching, and the learning acquired from it, much less susceptible to comparative measurement than research.

Research as codified in league tables of various kinds poses a different challenge. The producers of research are individual or groups of individual faculty, staff, graduate students, and post-doctoral employees. The focus of research is through the principal investigators, or the individual faculty or groups of faculty who are responsible for developing and executing a research plan that can produce new knowledge for publication. The league table, following the model of sports competition, ranks the academic equivalent of the sports team, the university, although the team ranking is the

sum of the rankings of all the faculty in the institution.

The university is the organization responsible for the context and funding that makes the work of most individual researchers possible and provides the bureaucratic structure that holds the assets of research and deploys them in the competition for the ranking. Universities hire the best faculty, build the necessary facilities, provide the legal and bureaucratic frameworks that support and protect research activity, and by doing these things well, enhance the productivity of the individual researchers in their midst. The more of these winning academic research teams within a country the stronger the country's research base and the more likely its current and future success in the global competition for wealth and power.

Throughout the world, universities operate in significantly different legal, organizational, bureaucratic, and political spaces. Some countries have highly centralized university policies related to funding, mission, student qualifications, research support, and appropriate size. Other countries, especially the United States, have a mixed system. Private universities compete individually within the context provided by their resources, their historical missions, and the interests and support of their private ownership boards. Public universities—almost all owned by their respective states not the national government—operate within widely varying bureaucratic and political contexts. Some state political structures tightly control institutions while other states may regulate but not control their universities. Some state universities receive high subsidies from the public tax base while others have much lower levels of taxpayer support.

In the United States, the federal government provides a wide range of subsidies to educational institutions and imposes a variety of regulations in addition to those imposed by the states. As a result, the context for any individual public university will vary significantly from the context of another. Although the national government does not control the operation of universities in the United States, it creates the primary marketplace for research competition. Significant investment by federal research agencies creates a pool of funds in support of research that public and private university researchers compete for. In this model, the competitors are individual or small groups of researchers, sponsored, sheltered, subsidized for sure by their universities, but competing as individuals or groups of individuals for the federal funding that serves as the primary token of academic research quality. They may also make alliances with groups of faculty from other institutions to enhance the competitiveness of a research proposal.

In America as well as in the rest of the world, the most important element in producing good research is the availability of money. New knowledge is expensive to produce and

requires consistent investment over time. One of the characteristics of much new knowledge is that it has no useful application at the moment of its creation. When we identify something new, some characteristic of the physical, biological, or cultural world previously unknown or imperfectly understood, the discovery may appear trivial or bizarre to some observers. The utility of research discoveries often appears much later when others have extended the discovery into related areas of research, broadening the significance, and developing a full understanding that transforms the original invention into products or processes that change the way we live, create substantial economic or social value, or provide a major national competitive advantage. It is this prospect of transforming new knowledge into competitive global products that motivates the international concerns about research competitiveness.

The focus on university performance, which is in effect the team and not the actual producers of knowledge, recognizes that success in research requires consistent management of the research process and the research capable individuals. Success requires a continuous process of quality control and improvement to sustain long-term research productivity. Individual researchers will come and go and display bursts of brilliance, but the organizational structure of the research university offers the best mechanism for sustaining national research success over the long periods required to move from discovery to competitive product delivery.

Constructing League Tables

The league tables theoretically serve as annual markers of institutional success in sustaining a continuous level of competitive research productivity. The difficulty in all this, of course, is that the data on which we can base our league tables is not as accurate as the data from our sports competitions. This circumstance has led many institutions and agencies to focus on the problems of measurement in an attempt to improve the accuracy of the league tables. This effort has greatly improved our understanding of the publication process used to disseminate and validate research results, and it offers considerable promise for improvements in the international research ranking industry.

When constructing league tables, we have to resolve a number of issues that fall into several categories:

Unit of comparison: As mentioned above, we generally compare universities, not the departments of chemistry or biology, although discipline level data may have a much higher likelihood of being comparable. Universities differ significantly in their organization and structure in ways that affect measurements of research productivity of any kind. For example, the presence or absence of a medical school has a significant impact on the research productivity of a

university, especially in the United States with its strong tradition of funding biomedical research preferentially in comparison to other fields of study. Some ranking systems adjust for this by publishing their research information with and without the medical school productivity included [as is the case for *The Top American Research Universities*].

However, while this can provide additional information, it is not always persuasive. Universities without medical schools can and do place considerable emphasis on biomedical research in other science units on campus, and in a university with a medical school, the university may give a priority to the medical school for biomedical research, leaving the non-medical school science units focused on different topics. When we remove the medical school from such a university's research productivity, the remaining departments may not compare appropriately with similar departments in institutions without medical schools.

Less visible effects occur when universities have different historical emphases in their research priorities. Where a university invests heavily in work on the humanities, social sciences, and the professions such as business or education, the research captured by external funding statistics may not accurately reflect the competitive research accomplishments of the institution.

Nonetheless, in the normal academic league tables, the issue is usually not about research commitment, but rather scientific research commitment, and given that preference, the league tables based on scientific indicators are likely to serve some purposes even if the implication of measuring scholarly research productivity of the entire university is inaccurate. University research is also expensive, and some research fields and specialties are more expensive than others giving advantages to the well-funded institution. Similarly, those universities whose research profile most closely matches government priorities for research will also fare better than other institutions in most league-table arrangements.

Scale: In every case, scale is important. All other things being equal, the larger the university faculty dedicated to research, the more research productivity the university will show and the higher it will rank in the league tables. In some instances a university may contain large, non-faculty but highly effective research enterprises, such as occurs with the Applied Physics Laboratory at The Johns Hopkins University, whose work inflates a university research score relative to the faculty-only generated scores of other institutions. This is where the sports model of league standings fails the university competition. In sports, for the most part, every team fields the same number of players and the competition is between teams of equal size. The won-loss record, the ranking criterion of most significance, is the result of competition among units of the same size.

In the case of academic league tables, however, the implication that the ranked units are equivalent and the only difference is in the research productivity is inaccurate. What we measure in a league table for research is the aggregate research produced by the individuals associated with a particular institution, and the more individuals associated, the higher the research score. Often we try to normalize for size by calculating a measure of productivity per faculty member, but the wide variation in institutional definitions of faculty status and the complex practices for determining faculty responsibilities make such calculations unreliable at best.

Because of the scale factor, we pay much attention to the definition of the unit that goes under the name “university.” In the United States, the name university does not distinguish clearly the organizational structure of a higher education entity. Some organizations calling themselves universities are composed of two to ten or more institutions of higher education, many if not all of which support research programs and faculty. To appear more significant in the research competition, these university systems often report the research productivity of all their campuses as if they were one institution. Underlying this discussion is an important element influencing the development of research capability.

In the United States, the focus of university activity has usually been the geographic campus, the physical location of the buildings, the laboratories, the library, and especially the undergraduate student body. Because most American research universities emerged from a base of undergraduate student programs and later developed the advanced study associated with research, the funding, organization, support structure, physical plant, and other features of an academic research enterprise derives from the actions of geographically defined academic institutions. While there is indeed an entity with the name University of California, no one imagines that Berkeley and UCLA are the same institution. Faculty, the key components of the research system, see themselves as members of the faculty of UCLA or of Berkeley, not members of the overall faculty of the University of California. While technically, these campuses are subordinate units of the single University of California, in fact they operate as almost autonomous units for the purpose of developing and sustaining research productivity (as well as for the purpose of managing undergraduate and graduate student programs).

Most evaluation systems in the United States define the unit of interest as the campus, and the National Science Foundation now collects its data in this fashion. *The Top American Research Universities* has always defined its metrics to apply to an individual campus and not to the system of which a campus may be a part. In the United States, this issue is even more significant because public universities are often organized into large bureaucratic

constructs that carry the name university while their private university counterparts, major players in the research competition, are almost always single campus enterprises.

Elements of research distinction: Although there is general agreement that research distinction and productivity are the key issues in constructing the rankings in league tables for universities, the definition of measures to use for this purpose is more elusive. Many elements of distinction define the research university, although not all apply equally to the university environment in every country. In the United States, research universities share a number of characteristics. They have high quality undergraduate student bodies and faculty who win national and international awards for distinction and recognition by their peers. They produce a large number of doctoral graduates as part of their research commitment and of course, they produce large volumes of quality research. They generally have engaged alumni and private supporters who contribute annually and whose gifts build endowments.

Identifying the indicators for each of these, and other elements of quality we might find, proves difficult. If the goal is to create a league table focused only on publications, the data for such a metric is accessible if not altogether easy to use effectively. This element, as mentioned above, looks primarily at publications and citations to determine how much research each university’s staff produced and published, and how significant the community believes this research to be. The benefit of this measure is that it identifies an internationally accessible metric.

The disadvantage of this measure, besides some technical issues with the construction of the measure itself, is that it focuses on what individuals do, without additional indicators that would speak to the institutional context within which they do it. The publication/citation measure may represent a good proxy for the effectiveness of the institutions that house the individuals who publish, and for a simple won-loss record, this could be true. However, for a fuller understanding of the context that produces high quality research, it helps to have additional indicators. While it is certainly correct to assume that what matters for high quality universities are high quality faculty, the ability of a university to attract and retain internationally competitive personnel requires an elaborate and expensive institutional support system.

The United States has another, perhaps more direct method of sorting research productive faculty and programs in American universities. This occurs through the federal review process for grant awards. The panels assembled for these reviews include expert faculty from all over the country. The individuals chosen are recognized research faculty, and they must do their reviews in accord with specific guidelines with the results of the scoring made available

to the applicants. These measures limit if not eliminate the opportunity for favoritism to influence the judgments. At the end of the process, each year, the federal agency makes awards to deserving research proposals, either to continue ongoing work or start new work. The bulk of the awards are for science-based projects although considerable opportunities exist for social science and some humanities research.

However, as all participants in this process know, the application-reward cycle does not easily match the productivity cycle of research and as a result the Federal government also collects data on the amount of federal research dollars spent by each university each year, evening out the award data that may reflect projects with durations from one to five years. The expenditure data show the funds spent to do actual research (and are audited and therefore reasonably reliable). This indicator of annual research expenditures from federal funds serves as one of the most reliable indicators of research activity. Federal agencies will generally not renew grants without publications that demonstrate achievement or award new grants to scholars who do not publish, and consequently the annual expenditure of federal funds is a stable, reliable indicator.

An additional benefit of the federal funds indicator is that it reflects not only the ability of the researcher to design a worthy project but also the ability of the university to support this research. Every successful application includes indications of the institution's contributions, the resources of the institution available in support of the project (equipment, space, collaborative activities, instructional programs related to the research), and similar items.

This perspective illustrates that academic research is usually not an isolated individual product that stems from the independent work of a creative research scholar but is instead an institutionally supported enterprise built around the creative researchers. In this model, it becomes clear that even before a proposal reaches a funding agency, the sponsoring institution has made many choices to sort out the high quality productive researcher from the rest of the faculty. This institutional sorting is one of the major contributions of the university to the development of a national research capability.

The sorting occurs at various points in the career of an individual faculty member. At the point of hiring, the university first expresses its standards relative to research productivity by employing only those who show significant promise and past productivity in research. The promotion and especially the tenure process at American research universities also establish the standards for performance by keeping those who can perform at a competitive level and discouraging or dismissing those who cannot. An additional level of sorting occurs in the distribution of internal university funds and

assets in support of research. Not all aspiring faculty members can have the labs they want, the support staff they need, the time for research work required. When the university chooses to provide internal funds and resources to support the work of one faculty member more than the work of another, it makes a bet on the research productivity and competitiveness of the faculty. The best research universities make the best bets on future performance and they invest to make sure their bet is a winner.

Finally, in the American university, all research, however well funded, operates at a financial loss to the institution. If a faculty member receives a grant for \$100,000 to perform a research project, it will almost certainly cost the university at least \$150,000 to fulfill the research obligations contracted when accepting the grant. The external funding covers only certain direct and indirect costs of performing the research and fails to account for many other costs. The university picks up the other expenses using funds from other sources. As a result, the amount of money the university has available to invest in research is a key competitive advantage in achieving high levels of research performance.

The other indicators presented in *The Top American Research University* annual reports serve to help universities understand the context within which high performing research universities function. At the top of the list, universities tend to perform well on all quality measures, however constructed. It appears that high quality undergraduate student bodies, for example, provide an important context that helps attract and retain the best researchers, or at least create environments in which these people choose to work. Other elements help identify portions of the support base for institutional achievement through private giving and endowment, through the support of graduate students and doctoral programs, and through the presence of distinguished colleagues recognized with national awards and honors. An extensive discussion of the measures used in *The Top American Research Universities* appears in the first volume and has remained stable throughout the ten years of the publication's existence.

Identification of research product: Publication is the principal evidence of research accomplishment. Unpublished research serves no one and cannot be subject to the review and evaluation of other experts. Publication puts the results of research in the public domain, available for all to see, enhance, critique, replicate, and serve as the basis for advances in all areas of knowledge. Publication comes in many forms from articles to books, anthologies, and conference proceedings and these items appear in many venues. In an ideal world, we would have complete bibliographic control over all forms of research publication and could then attribute the publications back to the institutions that supported the authors to provide a reliable institution-based publication count.

It is not as easy as one might expect to achieve this kind of bibliographic control, although continuous work in the field of bibliometrics has greatly improved the accuracy and completeness of the process. The explosion of activity on the Internet and the gradual emergence of alternative virtual publication venues have complicated the measurement issues although the core requirement for scholarly publication whether issued in virtual or physical form remains peer review. Traditionally, physical publication through journals and scholarly books has rested on a gatekeeper function managed by the publication. Editors and reviewers sort and select from among all possible manuscripts submitted for publication, choosing those with the greatest likelihood of significance to the research field in question.

Because physical publication is expensive it creates a requirement for selection, and in that selection resides the possibility of quality control. The gatekeeper function creates its own controversies as rejected authors identify possible inappropriate preferences for certain fields or subjects and for certain authors or for authors from preferred institutions. The Internet created a free zone for self-publication or for the creation of unrefereed or lightly refereed publication in many fields. Publications, not fully certified by one of the various academic guilds, are not likely to be significant in the score keeping of the league tables as the most reputable efforts to measure research will almost certainly continue to focus on the peer reviewed journals, whether virtual or physical in presentation.

Even within the domain of physical or at least organized and refereed publication venues (sometimes published online) quality appears varied. Some publications see many submissions and accept relatively few. Other journals have a much higher acceptance rate. Although we could assume that the rejection rate is an indicator of quality, a low rejection rate may also signify a rare and lightly populated subfield of considerable significance to the advancement of knowledge.

Additionally, we have the challenge of disciplinary variation in the patterns of publication. Some academic guilds expect their members to publish papers constantly and with multiple authors. Others anticipate single-author papers of greater length published at a slower rate. The humanities and some social sciences value books more than they value journal articles. Other fields almost never see a book as a major contribution, relying instead on journal articles. For some disciplines, conference papers have significance while for others these serve as no more than announcements of significant research in process that may appear published in the future when complete.

While publication remains the coin of the realm for research achievement, the details of measuring publications remain complex and difficult to resolve. A simple publication

count, even if we could identify them all, would not provide a reliable comparative measure of research productivity.

The challenge of identifying an internationally comparable publication count leads to various possible solutions. Citation indexes, mentioned above, are perhaps the best known, and most developed mechanism to assign differential value to publications. A citation index attempts to identify the impact of a particular publication by counting the number of times other authors refer to it. The more other scholars cite a publication, the more significant it is and the more it should count in constructing a league table. Although this has considerable attractiveness, it is not without its own difficulties. Citations are, of course, time lagged from the time of the original publication, and a league table that is designed to reflect the annual performance of universities must determine the time frame for counting both publications and the citations that determine their significance. The time lags also differ by discipline, with some scientific disciplines cycling citations to new work within months while others may take a year or more for other researchers to receive and reference their own work.

Additionally, citations in prestigious journals may be more significant for identifying relative importance than citations in less prestigious journals, and in recognition of that circumstance, some counting systems weight the citations by the prestige of the journal. Obviously, the identification of prestigious or important journals is itself a challenging and controversial task.

The league tables that use citations as a marker for research performance resolve these questions in different ways. Of course, depending on the methodologies used, they will get different results. Among the groups working on these critical issues, The Center for Measuring University Performance is collaborating with the Global Alliance for Measuring University Performance sponsored by the United Nations University's International Institute for Software Technology (UNU-IIST) and Elsevier under the project name of the Global Research Benchmarking System (GRBS).

Reputational Surveys: Among the least reliable but still frequently used ranking measurements involves reputation surveys. These questionnaires go to presumably expert reviewers who then provide a reputation score for each institution included in the ranking table. The ranking method usually aggregates these scores to get a rank order of reputation, which the magazine can then weight and combine with other measures to produce a final league score. The problems with reputation surveys are many.

If the purpose of a ranking is to determine prestige or quality then perceptions of quality confuse the issue. Either the data and indicators measure quality, which is reported

in the rankings and can inform our opinions of quality or we already know what quality is and by surveying the experts, we achieve the right ranking without needing a reference to data on performance. We might want to do a study to see if the quality ranking based on data matches the quality ranking based on opinion, but to combine them creates a circular evaluation. The National Research Council's recent ranking publication does report reputation and various performance data separately, but that evaluation project has not been as successful as anticipated.

Even if by some means we could avoid the circular nature of including opinion surveys of quality into the ranking based on performance measures that would justify our opinions of quality, we have an additional dilemma. Expert opinion of university quality is notoriously inaccurate. No academic knows enough about more than a very few universities to make accurate judgments about their academic, research, student, or other quality.

When we ask many experts to rate university quality even within one country like the United States what we get are a wide range of halo effects. Reviewers rank Ivy League institutions highly because they are well known by name, if not so well known by the details of their programs. Institutions with prominent and highly televised sports programs rank more highly than institutions without such programs. Institutions in states with a reputation for educational investment rank more highly than institutions in states with poor reputations for educational quality. Sometimes these prejudices accidentally match performance data, but often they do not.

If reputation among a certain group of reviewers who will not have the time or opportunity to gain detailed information about our institution is an important criterion, universities will need to increase their visibility. When the reputation survey arrives, the higher visibility will prompt the respondents to associate an institutional name with the attributes of quality. This naturally leads to advertising and institutional self-promotion. The pages of *Inside Higher Ed*, the *Chronicle of Higher Education*, and other publications oriented toward academic administrators, often the respondents for the reputation surveys, carry many advertisements promoting the unique attributes of institutions seeking higher placement in the rankings through higher visibility.

Reminiscent of the publicity blitzes that precede the ranking that produces Academy Awards, the effort to create what some have called celebrity universities is a consequence of the ranking popularity contests that are a part of some highly promoted league tables such as the *U.S. News* and *Times Higher Education* ranking. Indeed, the *Times Higher Education* ranking business has found such commercial success with reputation ranking that they publish

an especially unreliable ranking based only on the opinions of presumed experts. Most serious students of higher education quality see the inclusion of reputation surveys as a clear indicator of a suspect league table.

Feedback Effects of League Tables

However we construct a league table, it will have a variety of consequences for the institutions included (as well as for those not included). Because the league table purports to identify the best and those that are less than the best in descending order, institutions attempt to modify their behavior to improve their comparative position, and some attempt to modify their data to look better in comparison to those above them on the list. Not all of these effects are pernicious, although many are.

When parents use a league table to make short-term decisions about the suitability of an institution for an undergraduate student or a governing board uses the lists to evaluate the annual effectiveness of institutional administrations, the impact is most damaging and least useful. Research university performance is a long-term phenomenon, and small changes in a league table are often the result of statistical errors, subsequent corrections, or one-time special events rather than reflections of actual improvement or decline.

If a league table encourages universities to perform better in terms of research, instruction, or other elements that define high quality within their context, then the league table encourages useful activity. If, however, the university emphasizes behaviors that influence the rankings elements but do not necessarily improve the performance of the institution, then the impact of the league table is negative. Even more significantly, if the importance of the league table in sorting institutions by reputation has an impact on access to resources or the imposition of government policies and controls, the influence of the ranking can easily become pernicious.

Universities and other academic units ranked in these contests may also work to manipulate their data by strategies that give them a comparative advantage in the rankings unrelated to their actual performance. Other than simply submitting fraudulent data (which can happen and which is why good rankings are always transparent and provide all the data and calculations for public review), universities can pursue other strategies. If high scores on entrance exams figure significantly in the calculation of an important ranking, universities can divert funds to increase the scholarships used to entice students with exemplary entrance exam results to enroll. By buying students with better examination scores, the institution's ranking will increase although the institution itself may not have become better.

Also in the United States, institutions can admit low quality students in the Spring and more qualified students in the Fall, since the measurement of student quality in some league tables only refers to the Fall semester. Universities can inflate their first-time student scores by admitting larger percentages of their students as transfers from other institutions after the first year. Some law schools in the United States, for example, found that by hiring their own graduates they could inflate an employment measurement used in the *U.S. News* ranking of law schools. These techniques are but a few examples of how institutions will manipulate the operation of their programs to match league table criteria in ways unrelated to improving academic performance.

These simple examples of manipulating the rankings system have many permutations and alternatives depending on the particular methodology of any individual ranking scheme. When, as is the case with *U.S. News*, part or all of the data and calculations on which the ranking are based are considered proprietary and unavailable for independent review, manipulation is made an even more attractive activity.

In some instances, league tables can affect large-scale revenue distribution when a country with highly centralized funding methodologies rewards institutions with high rankings by redistributing educational funds. When there are significant funding issues involved, the methodology and accuracy of the ranking system come under great scrutiny and, inevitably, controversy. In the United States, for example, some observers believe that the *U.S. News* ranking influences the decisions of students and their parents to choose one college or university over another for undergraduate work, one law school over another for professional study, and so on. The scale of this effect is difficult to establish since there is also some evidence that parents and students may use the rankings after the fact to justify a decision made on other grounds such as finances, admissibility, or geographical location.

Sustaining Effective Research Universities

The Top American Research Universities approached all these issues related to ranking as we began our project early in the rankings cycle in 2000 and the following principles have guided our work over the decade that followed.

- The first principle we established is that the purpose of our work was not to find the best research university but rather to identify the characteristics of the most successful research universities.
- The second principle we established was to only use data available from public sources. While much university

data are self-reported by the institution, we focused on data difficult to report inaccurately because of their open availability and because they would be checked, usually for other reasons, by other agencies.

- The third principle recognized that data have many interpretations, and while we have our own way of presenting university performance, we should make the data easily available for our colleagues to calculate or analyze using other perspectives or criteria.
- The fourth principle was to try to maintain stable definitions of institutions and data so that as much as possible we could compare performance over time.
- The fifth principle required us to correct errors identified by our colleagues and issue corrected tables and data in subsequent editions of *The Top American Research Universities* and its accompanying website.

The Center for Measuring University Performance (MUP) produces these tables not to create a ranking, although of course any such effort does create rank ordered lists, but rather to analyze the characteristics that identify successful research universities and identify benchmarks for university improvement. In our competitive world, universities need good data to help them recognize their own institution's strengths and expectations for improvement. They need to know how well the best institutions perform, and they need to consider what actions and investments will lead to better performance within their own context.

While the MUP center provides data that can be used for ranking institutions on a variety of indicators, our *Top American Research Universities* publication offers a different, and perhaps less satisfying, method of presenting university performance. We believe that university performance does not fall along a linear, one-dimensional ranking hierarchy. Instead, universities are complex, multi-product enterprises whose success involves competition on a number of dimensions. No single weighting system can produce a realistic and useful calibration using all relevant measures to establish the best of the best.

Rather, we find that research universities fall into groups, with similar if not identical performance on a variety of dimensions. If an institution is among the top 25 on all of the measures we identify as reliable and appropriate, then these institutions are all first rate even if they may have different levels of performance on the various measures. We do not believe identifying a calculated single best score within this group helps us understand how good universities deliver high levels of performance. We then cluster the institutions in groups depending on how many of our measures are in the top 25 or the second 25, illustrating that some institutions can compete at the top level in some

dimensions but perhaps not in others. In this fashion, we create groups of institutions of more or less comparable performance.

This technique frustrates many because they want us to mark a university as Number One and show how the rest fall on the scale beneath that top performance. This we refuse to do because it distorts the fundamental nature of university development and performance. Universities, as mentioned above, are not sports teams with seasons and short-temporal cycles. They are long cycle institutions whose work changes over time, whose qualities improve or decline in uneven cycles, some qualities exhibiting different rates of change. Our data appear to indicate that a university's relative performance within the larger marketplace of research universities changes only modestly over short time frames. The closer university performance indicators are, one to the other, the more year-to-year change in rank occurs but the less importance these changes represent. If a university improves its research from \$50M to \$52M it may move up in rank, but that move may be the result of its nearest competitor declining from \$53M to \$51M. These differences may produce a change in position on a linear scale, but they do not represent anything of significance. If, however, over a ten-year period, one university moves from \$50M to \$70M and the median of its group moves from \$50M to \$60M, then we may be observing an institution with consistently improving performance.

In any case, we do not see absolute rank as the issue. Instead, it is the competitive issue of how any individual university performs against the whole group of top American research universities. If federal research volume grows by 5%, an increase of 2% by any individual university means that university is losing ground against the marketplace of research institutions. It may be that the 2% institution increased its rank because those immediately above them performed even worse, but it does not mean that the 2% institution is performing better than the group as a whole. Moreover, if a university grows by 10% but the median growth of those above it in performance grows by 15%, the 10% university is still losing ground, even though it may be improving at a rate higher than those below it.

Part of the strategy of *The Top American Research Universities* publication is to provide data in a careful, organized, and structured way that nonetheless makes simple generalizations about individual research university performance difficult. There is no Number One in our publication, but there are data to tell how well any individual institution is doing over time relative to the group or any subgroup on any one or all of the reliable indicators assembled. This reduces the celebrity value of our publication, but we believe it increases its utility for university people interested in institutional improvement.

On Ranking and League Tables

This eclectic selection out of a large volume of material available on college and university rankings introduces their virtues and defects, their techniques, and their purposes. A search through Google, MUSE, J-Store, EBSCO, or other repositories will identify a large number of additional items on many related topics. Internet locations listed below often require subscription access.

University Rankings: Lists and Overview

“Bibliometrics,” *National Institutes of Health Library, Research Tools*, 2011.

[<http://nihlibrary.nih.gov/ResearchTools/Pages/bibliometrics.aspx>]

University of Illinois at Champaign-Urbana, University Library, Education and Social Science Library, College and University Rankings website

[<http://www.library.illinois.edu/edx/rankings/index.html>]

Wikipedia College and university rankings, Wikipedia on University Ranking: From Wikipedia, the free encyclopedia [http://en.wikipedia.org/wiki/College_and_university_rankings]

World Rankings Examples

Academic Ranking of World Universities (ARWU) Center for World-Class Universities and the Institute of Higher Education of Shanghai Jiao Tong University, China 2003-2009 Shanghai Ranking Consultancy, 2009-. [<http://www.arwu.org/index.jsp>]

Global Research Benchmarking System (GRBS). Global Alliance for Measuring University Performance. United Nations University, International Institute for Software Technology (UNU-IIST) and Elsevier.

[http://www.un.org/News/briefings/docs/2010/101020_UNU.doc.htm]

QS World University Rankings. QS Quacquarelli Symonds Limited

[<http://www.topuniversities.com/university-rankings/world-university-rankings/>]

Times Higher Education, World University Rankings and Reputation Rankings, 2010-2011,

[<http://www.timeshighereducation.co.uk/world-university-rankings/>]

World Rankings: Comment and Critique

Adelman, Clifford. "The Propaganda of International Comparisons," *Inside Higher Ed*, December 15, 2008. [http://www.insidehighered.com/views/2008/12/15/adelman]

Labbe, Cyril. "Ike Antkare One of the Great Stars in the Scientific Firmament," *Laboratoire d'Informatique de Grenoble*, September 2010. [http://rr.liglab.fr/research_report/RR-LIG-008.pdf]

Arnold, Douglas N. and Kristine K. Fowler. "Nefarious Numbers," arXiv:1010.0278v4 [math.HO] 17 Nov 2010. [http://arxiv.org/PS_cache/arxiv/pdf/1010/1010.0278v4.pdf]

Bougnol, M.-L. and J.H. Dulá. "Validating DEA as a Ranking Tool: An Application of DEA to Assess Performance in Higher Education," *Annals of Operations Research* (145:1, 2006). [http://www.springerlink.com/content/m827t62w9420kj24/fulltext.pdf]

Campbell, David, Michelle Picard-Aitken, Grégoire Côté, et al. "Bibliometrics as a Performance Measurement Tool for Research Evaluation: The Case of Research Funded by the National Cancer Institute of Canada," *American Journal of Evaluation* (31:1, 2010). [http://aje.sagepub.com/content/31/1/66]

Cohen, Sande. "Publication, Knowledge, Merit On: Some Politics of Editing," *Cultural Critique* (75, 2010). [http://muse.jhu.edu/journals/cultural_critique/v075/75.cohen.html]

Danchin, Antoine. "A Note on Bibliometrics," *Scientific Indicators and Game Theory*, nd. [http://www.normalesup.org/~adanchin/bibliography/bibliometrics.html]

De La Baume, Maia. "French University Rankings Draw Praise and Criticism," *The New York Times*, November 15, 2010. [http://www.nytimes.com/2010/11/15/world/europe/15iht-educSide15.html]

Gladwell, Malcolm, "The Order of Things: What college rankings really tell us," *The New Yorker*, Feb, 14, 2011. [http://www.newyorker.com/reporting/2011/02/14/110214fa_fact_gladwell]

Gould, Thomas H.P. "Scholar as E-Publisher: The Future Role of [Anonymous] Peer Review within Online Publishing," *Journal of Scholarly Publishing* (41:4, 2010). [http://muse.jhu.edu/journals/journal_of_scholarly_publishing/v041/41.4.gould.html]

Guttenplan, D.D. "Questionable Science Behind Academic Rankings," *The New York Times*, November 15, 2010. [http://www.nytimes.com/2010/11/15/education/15iht-educLede15.html]

Halfman, Willem and Loet Leydesdorff. "Is Inequality Among Universities Increasing? Gini Coefficients and the Elusive Rise of Elite Universities," *Minerva* (48, 2010). [http://arxiv.org/abs/1001.2921v1]

Hazelkorn, Ellen. "Rankings and the Battle for World-Class Excellence: Institutional Strategies and Policy Choices," *Higher Education Management and Policy* (21:1, 2009). [http://www.oecd.org/dataoecd/59/48/41203634.pdf]

Himanen, Laura, Otto Auranen, Hanna-Mari Puuska, and Mika Nieminen. "Influence of Research Funding and Science Policy on University Research Performance: A Comparison of Five Countries," *Science and Public Policy*, (36:6, 2009). [http://www.ingentaconnect.com/content/beechnet/spp/2009/00000036/00000006/art00001]

Hongcai, Wang. "University Rankings: Status Quo, Dilemmas, and Prospects," *Chinese Education and Society* (42:1, 2009). [http://mesharpe.metapress.com/app/home/contribution.asp?referrer=parent&backto=issue,5,10;journal,10,192;linkingpublicationresults,1:110902,1]

Institute for Higher Education Policy, ed. *College and University Ranking Systems: Global Perspectives and American Challenges*. Institute for Higher Education Policy, Washington DC, April 2007. [http://www.ihep.org/assets/files/publications/a-f/CollegeRankingSystems.pdf]

"Report on the Pilot Exercise to Develop Bibliometric Indicators for the Research Excellence Framework," Higher Education Funding Council for England, September 2009. [http://www.hefce.ac.uk/pubs/hefce/2009/09_39/09_39.pdf]

Ismail, Sharif, Edward Nason, Sonja Marjanovic, Jonathan Grant. "Bibliometrics as a Tool for Supporting Prospective R&D Decision-Making in the Health Sciences: Strengths, Weaknesses and Options for Future Development," Rand Corporation (Europe), 2009. [http://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR685.pdf]

"Leagues Apart," *Economist* 3/25/2010. [http://www.economist.com/node/15770798]

Marginson, Simon. "Open Source Knowledge and University Rankings," *Thesis Eleven*, (96, 2009). [http://the.sagepub.com/content/96/1/9.full.pdf+html]

Pendlebury, David A.. "White Paper: Using Bibliometrics in Evaluating Research," Thomson Reuters 2010. [http://researchanalytics.thomsonreuters.com/m/pdfs/Pendlebury_Paper.pdf]

"Research Front Maps, Research Front Map Site-Wide Index Listed by Field," Thomson Reuters, 2011. [<http://sciencewatch.com/dr/rfm>]

"Global Map of Science 2010," Thomson Reuters, 2010. [<http://sciencewatch.com/dr/rfm/mos/10maymosGLOBAL/>]

Rothblatt, Sheldon. "Global Branding and the Celebrity University," *Liberal Education* (Fall 2008). [http://www.aacu.org/liberaleducation/le-fa08/documents/le-fa08_Rothblatt.pdf]

Gorraiz, Juan, Henk Moed, and Edgar Schiebel, guest eds. "Special Issue on Performance Evaluation," *Research Evaluation* (18:3, 2009). [<http://www.ingentaconnect.com/content/beechnet/2009/00000018/00000003>]

Santiago, Rui, Teresa Carvalho, and Rosário Relva. "Research and the Universities' Image," *European Journal of Education* (43:4, 2008). [<http://onlinelibrary.wiley.com/doi/10.1111/j.1465-3435.2008.00370.x/full>]

Tappera, Ted and Ourania Filippakou. "The World-Class League Tables and the Sustaining of International Reputations in Higher Education," *Journal of Higher Education Policy and Management* (31:1, 2009). [<http://www.informaworld.com/smpp/content~db=all~content=a909187874>]

Tatavarti, R., N. Sridevi, and D. P. Kothari. "Assessing the Quality of University Research--RT Factor," *Current Science*, (98:8, 2010). [<http://www.ias.ac.in/currensci/25apr2010/1015.pdf>]

Tijssen, Robert J. W., Thed N van Leeuwen, and Erik van Wijk. "Benchmarking University-Industry Research Cooperation Worldwide: Performance Measurements and Indicators Based on Co-Authorship Data for the World's Largest Universities," *Research Evaluation*, (18:1, 2009). [http://www.cwts.nl/pdf/UIC_R_and_B_paper.pdf]

Usher, Alex and Massimo Savino. *A World of Difference: A Global Survey of University League Tables*, Educational Policy Institute, 2006. [<http://www.educationalpolicy.org/pdf/world-of-difference-200602162.pdf>]

West, Peter W.A. "A Faustian Bargain? Institutional Responses to National and International Rankings," *Higher Education Management and Policy* (21:1, 2009). [<http://www.oecd-ilibrary.org/docserver/download/full-text/8909011ec001.pdf?expires=1300296660&id=0000&acname=ocid57016313&checksum=54F7FD53EE342C3ABC2844C5F0956CBB>]

Withey, Lynne, et al. *Sustaining Scholarly Publishing: New Business Models for University Presses: A Report of the AAUP Task Force on Economic Models for Scholarly Publishing* (Washington DC, Association of American University Presses, March 2011). [http://aaupnet.org/resources/reports/business_models/aaupbusinessmodels2011.pdf]

Ying, Yu and Zhang Jingao. "An Empirical Study on Credibility of China's University Rankings: A Case Study of Three Rankings," *Chinese Education and Society* (42:1, 2009). [<http://www.mesharpe.com/mall/results1.asp?acr=ced>]

United States Rankings: Comment and Critique

Adelman, Clifford. "The Propaganda of Numbers," *The Chronicle of Higher Education*, October 13, 2006. [<http://chronicle.com/article/The-Propaganda-of-Numbers/2883>]

Ali, Mir M., Partha Bhattacharyya, and Anthony J. Olejniczak. "The Effects of Scholarly Productivity and Institutional Characteristics on the Distribution of Federal Research Grants," *The Journal of Higher Education* (81:2, 2010). [http://muse.jhu.edu/journals/journal_of_higher_education/v081/81.2.ali.pdf]

Capaldi, Elizabeth D. and Craig W. Abbey. "Performance and Costs in Higher Education: A Proposal for Better Data," Pre-publication version for *Change*, March 2011. [<http://www.performancemethod.com/CapaldiAbbeyChangePrePub2010final.pdf>]

Center for Measuring University Performance, *The Top American Research Universities*, 2000 [<http://mup.asu.edu/research.html>].

Lombardi, John V., Elizabeth D. Capaldi, et al. "The Myth of Number One: Indicators of Research University Performance" (2000). [<http://mup.asu.edu/MythNumberOne.pdf>]

Gater, Denise S. "U.S. News & World Report's Methodology" (2001, Revised). [<http://mup.asu.edu/usnews.html>]

Craig, Diane D. "The Center Top American Research Universities: An Overview" (2002).
[<http://mup.asu.edu/TARUChina.pdf>]

Gater, "A Review of Measures Used in U.S. News & World Report's America's Best Colleges" (2002).
[<http://mup.asu.edu/Gater0702.pdf>]

Gater, "Using National Data in University Rankings and Comparisons" (2003).
[<http://mup.asu.edu/gaternatldata.pdf>]

Lombardi, Capaldi, et al. "Deconstructing University Rankings: Medicine and Engineering, and Single Campus Research Competitiveness" (2005).
[<http://mup.asu.edu/DeconstructingUniversityRankings.pdf>]

Lombardi, Capaldi, et al. "Rankings, Competition, and the Evolving American University" (2007).
[<http://mup.asu.edu/research2007.pdf>]

Chang, Gordon C. and J.R. Osborn. "Spectacular Colleges and Spectacular Rankings: The 'U.S. News' rankings of American 'best' colleges," *Journal of Consumer Culture* (5:3, 2005).
[<http://joc.sagepub.com/content/5/3/338.full.pdf+html>]

Bastedo, Michael N. and Nicholas A. Bowman. "U.S. News and World Report College Rankings: Modeling Institutional Effects on Organizational Reputation," *American Journal of Education* (116, 2010).
[<http://www.jstor.org/stable/10.1086/649437>]

Bialik, Carl. "Small Change by U.S. News Leads To New Controversy in Rankings," *Wall Street Journal: The Numbers Guy*, April 7, 2005.
[<http://online.wsj.com/article/0,,SB111279937006999640,00.html>]

Dillon, Erin. "America's Best Master's Universities and Baccalaureate Colleges," *Washington Monthly* (2010).
[http://www.washingtonmonthly.com/college_guide/feature/americas_best_masters_universi_1.php]

Emens, Stephanie C. "The Methodology and Manipulation of the U.S. News Law School Rankings," *The Journal of the Legal Profession* (34, 2009).
[<http://web.ebscohost.com/ehost/viewarticle?data=dGJyMP Pp44rp2%2fdV0%2bnjisfk5Ie46bZMt6e3T7Kk63nn5Kx95uXxjL6prUm0pbBIr6eeSrimtlKzq55oy5zyit%2fk8Xnh6ueH7N%2fiVa%2bmtEiuprFPrq2khN%2fk5VXj5KR84LP-fUeac8nnls79mpNfsVbOnsk62qq5IpNztiuvX8lXk6%2bqE8tv2jAAA&hid=10>]

"Failing the Grade: The Craze for Ranking Humanities Journals," The Editors. *Kritika: Explorations in Russian and Eurasian History* (10:1, 2009).
[<http://muse.jhu.edu/journals/kritika/v010/10.1.article.pdf>]

Fehrman, Craig T. "Preprofessionalism: Rankings, Rewards, and the Graduate Admissions Process," *College Literature* (36:3, 2009).
[http://muse.jhu.edu/journals/college_literature/v036/36.3.fehrman.pdf]

Sehgal, Ashwini R. "The Role of Reputation in U.S. News and World Report's Rankings of the Top 50 American Hospitals," *Annals of Internal Medicine* (152:8, 2010).
[<http://web.ebscohost.com/ehost/viewarticle?data=dGJyMP Pp44rp2%2fdV0%2bnjisfk5Ie46bZMt6e3T7Kk63nn5Kx95uXxjL6prUm0pbBIr6eeTbimsIKwp55oy5zyit%2fk8Xnh6ueH7N%2fiVbWstEq0rbBNrpzqeezdu33snOJ6u9e3gKTq33%2b7t8w%2b3%2bS7TLepr0yyr7Y%2b5OXwhd%2fq4ji3MSN6uLSfffbq&hid=10>]

U.S. News and World Report: "College Rankings & Lists"
[<http://colleges.usnews.rankingsandreviews.com/best-colleges>]