The Top American Research Universities

2012 Annual Report

The Center for Measuring University Performance

John V. Lombardi
Elizabeth D. Phillips
Craig W. Abbey
Diane D. Craig

Measuring Research Performance:

National and International Perspectives

by Diane D. Craig and John V. Lombardi

University based education and research continue to hold a central place in the international competition for national prosperity. Countries and regions throughout the world struggle with the challenge of finding the right combination of instructional formats and research programs that will produce the most competitive labor force and intellectual and technical expertise. Some advocate for more emphasis on the tool skills of a technically competent work force while others seek greater emphasis on the invention and innovation of basic and applied research programs. All, however, know that it is the educational engine in its instructional and research modes that allows modern societies to compete successfully in the global economy.

In America, as well as elsewhere in the world, research universities hold a place of honor and prestige, setting standards of academic and educational performance. A consequence of this importance has been the rise of an industry devoted to the evaluation, assessment, and ranking of academic research universities on both a national and international basis. Over more than a decade, The Center for Measuring University Performance (the MUP Center) has published an annual report on the Top American Research Universities that categorizes American institutions using reliable indicators of performance. Within the US marketplace, the notion of a high quality research university includes a number of different elements in addition to the production of research results, and the institutional competition for resources and talented people encourages universities to develop profiles that serve as indicators of institutional prestige and significance. Among these we can identify the following for America's premier research universities:

- A physical institutional presence on a campus that projects through the arrangement of buildings and open spaces the image of a charmed and special place of learning and growth.
- A collection of faculty, characterized by the prestige associated with their place of training and by the accumulated weight of their research discoveries and publications.
- A body of students—seen as reflecting in the quality of their pre-collegiate academic preparation and the breadth and depth of their experiences in and of the world—who are a cross section of what many expect will become tomorrow's social, political, and economic elite.
- A set of academic program offerings that closely match those of the most prestigious colleges and universities in America, touching on the humanities, the fine arts, the social sciences, and including strong representations of the mathematical, physical, and biological sciences, and often a significant presence of professional programs related to health, business, engineering, and education. These, within the context of prestigious institutions, will be organized into coherent curricular plans that lead students to graduation, successful transition to employment, or more significantly for the purposes of prestige, to the next stage of graduate education for an advanced degree and a promising professional career.
- And finally, the university places itself among the most significant producers of advanced research in the nation, a presence reflected in publications, research grants, inventions, laboratories, national and international prizes, and the engagement of the faculty, staff, graduate and post graduates whose work claims national and international attention.

When the various enterprises around the nation and the world create rankings and league tables, they seek primarily to quantify the qualities that produce legitimate prestige, separate from institutional advertising and self-promotion or reliance on historical traditions of excellence.

The MUP Center has approached this issue within a framework that speaks to this multi-dimensional definition of institutional achievement. Each of the nine measures collected and verified by the MUP Center and included in the annual Top American Research University reports touches on one or another of these criteria through reasonably robust indicators. Rather than creating a ranked list of institutions, the MUP Center's annual reports present these top institutions within groups, recognizing that small differences among similar institutions are generally not significant. In the materials accompanying these annual reports, the MUP Center has explored the techniques and described the methodology in some detail. On its website, a significant collection of annual institutional data serves to help universities place their accomplishments within the national context.

The International Context for Comparative University Assessment

With the increased international interest in research university development and evaluation, reflected in a variety of ranking systems, the MUP Center has been fortunate to co-sponsor with the United Nations University's International Institute for Software Technology (UNU-IIST) the development of the Global Research Benchmarking System (GRBS). This initiative draws on the large Scopus database of academic journal articles and other peer reviewed materials made available to this project through collaboration with Elsevier. The GRBS develops reliable measures of research productivity and quality applicable to research institutions throughout the world.

Many ranking and league table exercises, some sustaining significant for-profit enterprises, identify measures that address the elements defined above as contributing to research university reputation or prestige. A defect with some of these efforts is their use of the notion of prestige to define prestige, creating a circular and self-reinforcing opinion cycle that offers little basis for understanding the substance of an institution's reputation. The circularity occurs when a league table uses the results of a survey of consumers or experts who are asked to identify the most prestigious institutions as evidence that a university is indeed prestigious. This approach may well identify an element of public opinion, but it does not provide information on the substance of university performance that is the basis of prestige. We might want to know, for example, if public opinion is substantially in error because it identifies an institution as high prestige when, in fact, its performance on a variety of measures that underlie the concept of prestige is less significant when compared to many others.

The data available for national comparisons of institutions are often not easily compared internationally. In the US, for example, a high score from the college entrance examination of the Scholastic Aptitude Test (SAT) can serve as a surrogate measure of an entering undergraduate's academic preparation for college work, and it may well also indicate the individual's socio economic circumstances. Collegiate prestige in America depends in part on the quality of an institution's students making the average SAT of entering students a helpful indicator of prestige. However, the sorting mechanisms that place students within various higher education institutions in other countries are quite different from those in the US, making the SAT a useful national but not international measure of comparative institutional reputation.

Some countries find that research funding can serve as a proxy for research significance. In the US, for example, the federal research funding competition serves as a major determinant of university research prestige because it is primarily peer reviewed and the competition is conducted on an open nationwide basis. In other countries, national governments award research funding to institutions based on an assessment of the institution's research performance or promise, but in the US the federal funds come from an individual competition among research scholars or groups proposing projects for funding. The institutional indicator of federal research support in the US, then, is the sum of the successful grants earned by individual researchers who work at the institution and is difficult to compare with the funding grants given by national governments to institutions in support of research programs.

The International Focus on Publication

The GRBS attempts to overcome these nationally specific research measures by focusing on research publications, the most universally recognized token of research performance. Whatever else differs among countries in the development and promotion of research universities, everyone who seeks research prestige must do so by publishing results available to the world. Thanks to the advent of sophisticated bibliographic databases such as the Elsevier's Scopus, it is now possible to approximate a universal view of the world's research publications.

As the GRBS team has developed its data, measurements, and systems and acquired strong collaboration and support throughout the world, it has also been careful to outline the limitations of the publication-based assessments. The source database includes primarily the academic articles published in peer-reviewed journals, and also includes reviews and conference papers although the coverage of these is not as complete as it is for articles. As a result the data exclude some important research results that appear in books, some electronic publications, or other formats that may not yet be included within the database. The GRBS focuses on science, technology, engineering, and other related fields and currently does not include the social sciences or the humanities, fine arts, and some of the professions. An additional characteristic is that a large majority of the research captured by the database appears in English, regardless of the language or nationality of authors or institutions. Scientific results appearing in other languages are less likely to be included within this system.

Using Publication Database Information

Within this context, the database allows a reasonable approximation of the quantity and quality of publications by scholars associated with a large and growing number of international universities within a wide range of specific fields and subfields. From these data, we can then measure the research staff's contributions and produce a more consistent indication of the relative accomplishments of research universities that serve as the basis for their reputations and international prestige.

Unfortunately, the issue of research prestige is not easily resolved. While the GRBS can count publications and associate them with individuals and the individuals with their institutions, it is difficult to distinguish precisely those publications of high significance from those of relatively modest importance. A university whose faculty consistently produce large numbers of highly significant research publications will surely be deserving of higher prestige than those whose staff publish less important work.

The GRBS uses a number of the available bibliometric tools scholars have developed to measure the significance of published research work. These methods recognize that a particular journal article or other published item is a major contribution to the field when others who are expert cite this article within their own published work. The more citations that a journal article receives from others who publish, the more significant we believe the journal article to be.

As not all journal articles are equally important, neither are all academic journals of equal significance. If one journal publishes articles that rarely earn citations in the articles of other scholars, then we can conclude that this journal is not of major importance. However, if another journal publishes articles that usually earn a significant number of citations in other articles, then we can assume that this other journal is a prestigious place to publish. This notion of prestige functions because journals can only publish a few articles out of the many they receive. The more prestigious a journal the more submissions it will get, and thus the journal will choose only the most significant articles to publish. If it uses its scarcity well, and selects important items to publish, those items will be widely cited, increasing the journal's reputation, earning it more submissions, and allowing it to continue to be a preeminent journal.

In developing its system for assessing the research strength of universities, the GRBS has engaged these topics, and for the purposes of this discussion we can review some of the methods used to calculate the significance of research publications in the wide range of subject fields, the extensive number of journals, and the widely varying patterns of publication and citation characteristic of the different scholarly fields.

Specific Internationally Applicable Measures of Publications and Citations

A variety of different methods help identify the significance of research publications, each developed to address one or both of the twin issues of quantity and quality. For the purposes of our discussion here, we can recognize some of the most significant that affect the comparison of research university performance within the context of the GRBS.

A frequently used measure of research significance is what is known as the *h*-index (named after its inventor Seymour Hirsch). Hirsch defined this relatively simple measure as follows: "A scientist has index h if h of his/her Np papers have at least h citations each, and the other (Np – h) papers have no more than h citations each." This measure grows larger over the length of a scholarly career. When using the *h*-index to measure current institutional performance, the GRBS uses the most recent 4-year window for calculating this indicator. The principal difficulty with this measure is that it directly reflects the characteristics of publications and citations within each field. If one field has the pattern of many publications with many citations, scholars in that field will have higher *h*-indexes than scholars of equivalent merit in fields with lower publication rates and citation numbers. GRBS calculates the *h*-index for each university within a particular subfield and four-year window.

Another method the GRBS uses to help identify high quality producers of scholarship comes from a measure called the Source-Normalized Impact per Paper or SNIP. This calculation as described by its inventor Henk F. Moed "measures a source's contextual citation impact. It takes into account characteristics of the source's subject field, especially the frequency at which authors cite other papers in their reference lists, the speed at which citation impact matures, and the extent to which the database used in the assessment covers the field's literature. SNIP is the ratio of a source's average citation count per paper, and the 'citation potential' of its subject field. It aims to allow direct comparison of sources in different subject fields." GRBS obtains the most recent SNIP value from Elsevier's Scopus database for each source title (journals, conference proceedings, and book series).

The MUP Center and the GRBS Approaches to University Research Performance

While the MUP Center's annual reports continue to provide a stable and useful guide to overall American university performance within the context of the nine measures used in that study, the success of the GRBS in developing international bibliometric indicators for assessing research publication gives an important alternative perspective on university performance. The MUP Center focuses on a

variety of measures indicative of institutional research capacity and success that address indirect evidence of research performance. Federal and total research expenditures for example reflect the success of institutional faculty and staff in competing for research funding but they do not measure the academic results of the research. Similarly, the indicators of faculty awards, post-docs, doctorates, and academy membership all touch on elements that are characteristic of high performing research universities and their faculty but again do not directly measure academic productivity. The measure of SAT scores, as mentioned above, speaks to a special circumstance of US research universities where the presence of high quality undergraduate students appears to indicate a campus context that first rank researchers find congenial. While this may be a significant asset in recruiting and retaining highly productive faculty and staff, it does not reflect the actual production of academic research.

For the purposes of the MUP Center, this approach has several advantages. Research success in America relies heavily on the organization and support of research universities, institutions with many functions only some of which are directly related to the production and publication of research results. Nonetheless, research universities in the US appear to have a set of characteristics that encourage and sustain research productivity, and the measures identified for the MUP Center's annual reports speak not only to those most indicative of research competition (federal research expenditures) but also those most indicative of a supportive institutional context.

The GRBS measures published research results directly without regard to institutional context, national research organization, or differential patterns of funding. From the MUP Center's perspective, a comparison of the American institutions that fall within the top performing categories of the GBRS data with those in the top categories of the MUP Center's annual report on American universities offers an opportunity to understand better the elements of institutional research success. The additional detail and sophistication of GBRS bibliometric data and indexes might appear to offer an easier approach to the question of aggregate academic performance by scholars associated with universities, but the added information also increases the challenge of defining high academic performance.

The GRBS currently provides measures that address the volume and significance of publications while continuing to develop other views of research performance, such as levels of international collaboration. The GRBS reports publications separately by academic area of specialty, allowing measurement of research success within 15 subject fields defined broadly at a top level and another 253 defined beneath these within subfields. This depth and detail recognizes that while many research institutions pursue a broad agenda competing in almost all areas of research significance, others take a more focused approach, seeking distinction in only a few fields where their resources will be sufficient to support outstanding performance. In this essay, however, we focus on the institutional level of performance across the 15 top-level subject areas, as this most closely matches the perspective of the MUP Center's Top American Research Universities annual report.

In reporting institutional score, the GRBS uses percentile groupings. A composite score is created for each of the fields or subfields by choosing optional weighting patterns for the seven indicators (described below). The composite score is normalized within each field and ranges from 0 to 100. Results are reported in "bands," with each band corresponding to a decile range. For example, band 1 is a composite score of 91-100, band 2 is 90-81, etc. In other words, a university that is in band 1 in a particular field is in the top 10% worldwide based on the measures and weights of the fields selected for review. A university within bands 1 through 3 will be in the top 30%.

We can anticipate that the aggregate measures of the GRBS should produce general clusters of American research universities that generally match those developed by the MUP Center and published in the annual MUP Center's reports.

Constructing GRBS University Assessments

Before comparing the results of the MUP Center and GRBS assessments, however, the sophistication of GRBS data provides us with additional complexity. In producing results for entire universities, the GRBS offers choices about how to weight the various indicators of publication and citation quantity and quality aggregated for an institutional score. Similar to the MUP Center's clustering methodology, GRBS' overall institutional score is based on how many times the institution falls into the top 30% (or top 3 bands) worldwide within each of the 15 top-level subject fields.

GRBS provides seven indicators that when combined identify bands of top performing institutions:

1. Total number of publications
2. Total number of citations
3. Percent of publications published in the top 10% journals, based on SNIP value
4. Percent of publications published in the top 25% journals, based on SNIP value

5. Percent of citations received from publications in the top 10% journals, based on SNIP value

6. Percent of citations received from publications in the top 25% journals, based on SNIP value

7. Four year *h*-index

The first two are indicators of total productivity and total recognition of that productivity within the academic publication marketplace through citations. Items 3 through 6 are four indicators of quality. The percent of publications in the top 10% or 25% of journals, based on their SNIP values for the last year of the time frame window (or 2011 in the case of the 2012 version of GRBS), indicate the proportion of total publications that appeared in the most prestigious journals in a specific field. The percent of citations in the top 10% or 25% of journals indicate the proportion of total citations to those publications that appeared in items published in the most prestigious journals. Item 7 offers a somewhat different balance of quality and quantity by calculating an institutional four-year *h*-index that includes the entire staff at an institution that published in a specific field.

The flexibility of the GRBS permits a selection of all or some of the top level 15 academic fields to include in a calculation of university research performance, and then to assign weights to the seven indicators within the selected fields to achieve a desired balance. Through this mechanism, the GRBS can calculate the relative performance of institutions in accord with a wide range of assumptions about the significance of quality vs. quantity and the importance of particular research fields. The virtue of this system is the flexibility that permits an assessment of university research performance within different perspectives. It does not, however, provide a simple rank ordering of universities from the best to the worst, a popular but often misleading result of many league table ranking systems.

Comparing the MUP Center Reports and GRBS

The MUP Center's annual reports compare universities within a specific US competitive context, and it is of considerable interest to observe how the US assessment compares with the publication-driven international university assessments generated by the GRBS. Ideally, the US system that categorizes research university performance on a variety of institutional characteristics related to university support for and success in achieving research capability should reflect the publication performance evaluation of the GRBS. The purpose of research universities, whatever the national context for their operations, is to generate significant scholarly work reflected most generally through publication.

However, the MUP Center's reports capture a much wider range of institutional characteristics in identifying top research universities, many of which may not contribute to publication results. For example, federal research expenditures (perhaps the most directly research specific element among the MUP indicators) includes some federal research awarded for activities that do not lead directly to publication, for example construction of such expensive physical objects as telescopes or the provision of experimental physics equipment. Other federal research may be awarded for projects in the national defense, for which no publication is expected or even allowed. The federal competition awards funds for projects that may provide practical information that is not normally published in prestigious peer reviewed scholarly journals. Nonetheless, in general terms, the individual institutions identified as top American research universities should, however identified, also appear as stellar performers when measured on the international scale of publication quality and quantity.

For this exercise then, we have identified two sets of top US research universities using two different rating models from the GRBS and matched these sets against the 2011 MUP Center's group of institutions and against the top US performers in terms of federal research expenditures. The MUP Center identifies its top performers by their ranking within the top 25 on each of the nine measures. Using this method, illustrated and described in prior annual reports, some 49 universities fall into this top category of high performing institutions by having at least one indicator in the top 25 nationally. The MUP Center reports exclude specialized institutes and medical institutions, since its focus is on full service universities, and we have also excluded those institutions from the ratings developed from the GRBS data.

The first set of institutions, Model 1, from the GRBS data uses the following criteria and weights applied against the 15 top-level field categories. To avoid double counting, we set the weights of elements 3 and 5 to zero:

GRBS Model 1

1. 10% weight on the Total number of publications

2. 10% weight on the Total number of citations

3. 0% weight on the Percent of publications published in the top 10% journals, based on SNIP value

4. 35% weight on the Percent of publications published in the top 25% journals, based on SNIP value

5. 0% weight on the Percent of citations received from publications in the top 10% journals, based on SNIP value

6. 35% weight on the Percent of citations received from publications in the top 25% journals, based on SNIP value

7. 10% weight on the Four-year *h*-index

While the range of weights possible is of course wide, this set offers a modestly rigorous selection by focusing on total publications and citations, the h-factor, and those publications and citations in the top 25% of SNIP journals. This model gives significant weight to the volume of publications and citations with the use of the h-index, in addition to an emphasis on the quality as reflected in the SNIP rating of the journals. This set appears in Table 1.

The second set of criteria, Model 2, offers a more rigorous application of the publication and citation metrics. For this set we used the GRBS data applying the following criteria and weights applied against the 15 top-level field categories. Again, to avoid double counting, in this model we set the weights of elements 4 and 6 to zero. This model applies more rigorous criteria for the rated institutions by focusing on only the top 10% journals (based on SNIP value) for publication and citation measures. This second model uses only metrics that reflect the quality of publications and citations, by setting weight for the h-index to zero and the weights for the total numbers of publications and citations to zero. This set appears in Table 2.

GRBS Model 2

1. 0% weight on the Total number of publications

2. 0% weight on the Total number of citations

3. 50% weight on the Percent of publications published in the top 10% journals, based on SNIP value

4. 0% weight on the Percent of publications published in the top 25% journals, based on SNIP value

5. 50% weight on the Percent of citations received from publications in the top 10% journals, based on SNIP value

6. 0% weight on the Percent of citations received from publications in the top 25% journals, based on SNIP value

7. 0% weight on the Four year h-index

Results of GRBS and MUP Report Ratings

As Table 1 shows, the criteria for Model 1 are met for all 15 of the top-level fields by 17 institutions out of the top 54 identified in this group. Even the last institutions in this group qualified within the top three bands for 9 of the 15 fields according to these criteria. We have indicated the 43 of these 54 institutions that also appear among the 49 Top American Research Universities identified in the last annual report. Clearly, this view shows a rather close convergence between the universities identified by the MUP Center methodology and the GRBS ratings.

However in Table 2, using the second, more rigorous set of criteria in Model 2, no institution rates inclusion within the top three bands of the GRBS for more than 10 of the top level fields. Within the 63 institutions with performance at this level, the last group of institutions has only 3 of the possible 15 fields qualifying within the top three bands. This group of 63 institutions includes 41 from the 49 MUP Center institutions, again illustrating the convergence between the results of GRBS and MUP methodologies in identifying high performing research universities.

These examples, but two out of the many that could be constructed using different weights for the publication and citation measures, demonstrate the importance of fields in this type of institutional evaluation. No university can be the best at everything, although some perform at high levels in many fields. The GRBS serves as a tool to help institutions identify where their faculty and staff's publications have the most impact. The institution can then direct its development strategies to enhance work in fields where the university is already successful and to expand into other fields that might offer the greatest opportunity for a high return from additional investments.

An inspection of the two tables also highlights the different grouping of institutions achieved with the different weights. Three institutions remain in the top group in both tables (MIT, Stanford, and Berkeley). Others in the top group in Table 1 fall into one of the bottom two groups in Table 2 (Duke, Minnesota, Penn State, and UC-Davis). These results demonstrate that many institutions will have good performance across many fields, but stellar performance in fewer fields. A careful review of the data on the GRBS website will identify the different ratings by field. Finally, of the eleven universities in the top three groups of institutions within the MUP Center categories all are also within the top group of GRBS institutions in Table 1, but only seven fall within the top three groups of GRBS institutions among US institutions in Table 2. This, again, indicates that with more rigorous criteria for inclusion, fewer institutions remain competitive.

Table 1

Model 1: Top GRBS Universities vs. Top American Research Universities

Institutions	No. of Times in Top 3 Bands (2012 GRBS)	In Top 25 Nationally (2011 MUP)
Columbia U	15	x
Cornell U	15	x
Duke U	15	x
Harvard U	15	x
Massachusetts Inst of Technology	15	x
Pennsylvania State U - University Park	15	x
Stanford U	15	x
U of California - Berkeley	15	x
U of California - Davis	15	x
U of California - Los Angeles	15	x
U of California - San Diego	15	x
U of Chicago	15	x
U of Michigan - Ann Arbor	15	x
U of Minnesota - Twin Cities	15	x
U of Washington - Seattle	15	x
Washington U - St. Louis	15	x
Yale U	15	x
Boston U	14	
Johns Hopkins U	14	x
Ohio State U - Columbus	14	x
Princeton U	14	x
U of Pennsylvania	14	x
California Institute of Technology	13	x
Georgia Institute of Technology	13	x
U of California - Irvine	13	x
U of Illinois - Urbana-Champaign	13	x
U of Maryland - College Park	13	x
U of North Carolina at Chapel Hill	13	x
U of Texas - Austin	13	x
U of Wisconsin - Madison	13	x
New York U	12	x
Northwestern U	12	x
U of Southern California	12	x
Carnegie Mellon U	11	
Rice U	11	x
Texas A&M U	11	x
U of California - Santa Barbara	11	x
U of Florida	11	x
U of Iowa	11	
U of Pittsburgh	11	x
U of Utah	11	
Vanderbilt U	11	x
Arizona State U	10	x
Brown U	10	x
Purdue U - West Lafayette	10	x
Rutgers U - New Brunswick	10	
U of California - Santa Cruz	10	
Emory U	9	x
Florida State U	9	
U of California - Riverside	9	
U of Colorado - Boulder	9	
U of Massachusetts - Amherst	9	
U of Rochester	9	
U of Virginia	9	x

Model 1 weights: 10% total publications; 10% total citations; 35% top 25% SNIP publications; 35% top 25% SNIP citations, 10% h-index.

Table 2

Model 2: Top GRBS Universities vs. Top American Research Universities

Institutions	No. of Times in Top 3 Bands (2012 GRBS)	In Top 25 Nationally (2011 MUP)
Harvard U	10	x
Massachusetts Inst of Technology	10	x
Stanford U	10	x
U of California - Berkeley	10	x
Rice U	9	x
Yale U	9	x
Columbia U	8	x
U of California - Los Angeles	8	x
California Institute of Technology	7	x
Georgia Institute of Technology	7	x
Princeton U	7	x
U of California - San Diego	7	x
U of California - Santa Barbara	7	x
U of Chicago	7	x
Carnegie Mellon U	6	
Cornell U	6	x
U of Colorado - Boulder	6	
U of Pennsylvania	6	x
Washington U - St. Louis	6	x
Boston U	5	
Northwestern U	5	x
Tufts U	5	x
U of Michigan - Ann Arbor	5	x
U of Massachusetts - Amherst	5	
U of Utah	5	
U of Washington - Seattle	5	x
Arizona State U	4	x
Brown U	4	x
Dartmouth College	4	x
Duke U	4	x
New Jersey Institute of Technology	4	
New York U	4	x
Pennsylvania State U - University Park	4	x
Stony Brook U	4	
U of California - Riverside	4	
U of California - Santa Cruz	4	
U of California - Irvine	4	x
U of Cincinnati	4	
U of Illinois - Urbana-Champaign	4	x
U of Maryland - Baltimore County	4	
U of Maryland - College Park	4	x
U of Minnesota - Twin Cities	4	x
U of North Carolina at Chapel Hill	4	x
U of Southern California	4	x
U of Texas - Austin	4	x
Boston College	3	
Clarkson U	3	
Colorado School of Mines	3	
Emory U	3	x
Johns Hopkins U	3	x
North Carolina State U	3	
Ohio State U - Columbus	3	x
Oregon State U	3	
Purdue U - West Lafayette	3	x
U of California - Davis	3	x
U of Delaware	3	
U of Idaho	3	
U of New Mexico	3	
U of Pittsburgh	3	x
U of Texas - Dallas	3	
U of Wisconsin - Madison	3	x
Wake Forest U	3	
Yeshiva U	3	

Model 2 weights: 50% top 10% SNIP publications; 50% top 10% SNIP citations.

Clearly, then, the seven GRBS metrics are capable of producing significantly different ratings of institutions as a result of the weighting and the choice of fields. These choices, while in some sense arbitrary, nonetheless offer an opportunity to focus on different aspects of research university performance. The emphasis on the most prestigious publication opportunities can provide a refined result, but for many national and institutional purposes, it may be just as significant to have as many universities as possible publishing often-cited articles in good academic journals as it is to have the most prestigious results possible.

Alternative ratings are also possible by focusing closely on particular fields of interest either to universities themselves or to their sponsors. The ratings would be substantially different below the top institutions with only some of the 15 top-level fields included. While these results clearly illustrate the power of the GRBS to provide a fine grained perspective on research publication performance, they also demonstrate the complexity of research evaluation and illustrate the crude and unsatisfactory approximation of single number league tables that gain widespread public notoriety.

Research Funding and Publication Metrics

A final perspective on these data demonstrates the importance of understanding the purpose of measuring university performance before presenting rating information. As indicated above, in the United States, the competition for federal research funding produces one of the key indicators of institutional research preeminence. These funds, from a variety of federal agencies, are awarded for projects in a wide range of fields using a primarily peer reviewed processes. Although there are many issues with the criteria, the selection of general fields to support, and political concerns that may influence the process, the results of this competition are nonetheless seen as touchstones of American research performance. The MUP Center reports use annual federal research expenditures as a key measure among its nine indicators and reports this data element online for all institutions that receive federal support.

However, as also indicated above, federal research expenditures measure resources applied to research, not the actual publication produced by that research. To test the relationship between the publication data available through the GRBS and the federal research expenditures reported by the MUP Center in its annual reports, we compared the 63 institutions in Model 2 on page 9 (the more selective weighting of GRBS data) with the top 63 institutions reporting federal research expenditures (Table 3). Note that both the GRBS list and the federal research expenditures list exclude special purpose and medical only institution. As is the usual pattern in these comparisons, a group of universities performs at the top level from both perspectives. The top 20 US institutions in federal research expenditures also appear among the top 63 within the GRBS data generated through the Model 2 process. An additional 22 institutions within the top 63 in federal research expenditures appear within the top 63 GRBS Model 2 list. However, some 20 universities that rank within the top 63 US federal research expenditures list do not appear within the GRBS publication based list. Looked at from the publication perspective, of the 63 top institutions in the GRBS Model 2 list, 21 institutions fall outside the top 63 in US federal research expenditures (Table 4). Indeed, some of these institutions are well down the list based on federal research expenditures with the lowest ranking at 293.

These data clearly indicate that while federal research expenditures in the US are a significant indicator of research performance leading to publication, it is not a perfect reflection of a worldwide publication view of research achievement. Moreover, the two domains, federal research expenditures and publication information collected by Elsevier's Scopus, capture research activity in significantly different ways as we have discussed above. Scopus captures field specific information from a defined set of worldwide, primarily journal, publications and weights this data based on criteria related to a calculation of impact and significance both of articles and the journals in which they appear. The GRBS data also reflect field and discipline characteristics of publications.

The US measure of federal research expenditures captures a key variable in the development and support of research productivity in the US, which is also an imperfect but significant indicator of an institution's total financial support for research of all types. An institution may have significant federal research funding in the US, but work in fields of an applied nature with minimal publication opportunities in the prestige journals included within the GRBS data. In other cases, universities in the US may have a relatively low level of federal research funding but include faculty who receive research support from other sources (state, local, private, medical) and publish in fields with a high presence in prestige journals.

Understanding University Research Performance

From these preliminary, broad brush views of GRBS results compared to the MUP Center annual reports, we can perhaps draw some conclusions.

- First, assessing institutional research performance is a complex process that does not yield simple answers. A list of top universities without careful specification of the criteria and frame of reference for determining the evaluation is likely to be of little use.

Table 3

Top 63 Federal Research Universities* vs. GRBS Model 2 Top 63 Universities

Institutions	2009 Federal Research ($000s)	2009 Federal Research Natl Rank	Top 63 in 2012 GRBS
Johns Hopkins U	1,587,547	1	x
U of Michigan - Ann Arbor	636,216	2	x
U of Washington - Seattle	619,353	3	x
Massachusetts Inst of Technology	532,618	4	x
U of California - San Diego	511,428	5	x
U of Wisconsin - Madison	507,898	6	x
U of Pennsylvania	499,498	7	x
Columbia U	483,111	9	x
Stanford U	477,507	10	x
U of California - Los Angeles	467,505	11	x
U of Pittsburgh - Pittsburgh	463,192	12	x
Duke U	438,767	13	x
U of North Carolina - Chapel Hill	431,837	14	x
Washington U - St. Louis	414,045	15	x
U of Minnesota - Twin Cities	390,602	16	x
Pennsylvania State U - U Park	386,635	17	x
Harvard U	385,704	18	x
Yale U	378,914	19	x
U of Southern California	375,024	20	x
Ohio State U - Columbus	339,820	21	x
Vanderbilt U	336,405	22	
Georgia Institute of Technology	322,452	23	x
Case Western Reserve U	313,044	24	
U of Texas - Austin	309,125	25	x
California Institute of Technology	305,682	26	x
U of Chicago	301,159	27	x
Northwestern U	300,619	28	x
U of Alabama - Birmingham	300,130	29	
U of Rochester	295,963	30	
U of California - Davis	295,924	31	x
Emory U	295,831	32	x
U of Illinois - Urbana-Champaign	288,013	34	x
U of Arizona	287,889	35	
U of California - Berkeley	262,069	38	x
Texas A&M U	261,491	39	
U of Colorado - Denver	256,007	40	
Boston U	255,178	41	x
U of Iowa	252,336	42	
U of Maryland - College Park	246,985	43	x
U of Colorado - Boulder	239,687	44	x
Cornell U	238,022	45	x
U of Florida	232,737	47	
U of Cincinnati - Cincinnati	229,324	48	x
U of Virginia	218,499	49	
Colorado State U	211,890	50	
U of Hawaii - Manoa	203,453	52	
New York U	202,535	53	x
U of Illinois - Chicago	196,702	54	
U of Utah	192,354	56	x
U of South Florida - Tampa	190,949	57	
U of California - Irvine	177,098	58	x
Purdue U - West Lafayette	175,302	59	x
U of Miami	172,000	60	
Carnegie Mellon U	170,260	61	x
U of Maryland - Baltimore	169,081	62	
Michigan State U	164,198	63	
U at Buffalo	152,146	64	
Rutgers U - New Brunswick	151,122	65	
Virginia Polytechnic Inst and State U	148,411	66	
U of Kentucky	145,483	68	
Wake Forest U	144,454	69	x
Yeshiva U	137,108	70	x
North Carolina State U	135,318	71	x

** Excluding standalone medical and other specialized institutions.*

Table 4

GRBS Model 2: Top 63 Universities vs. Top 63 Federal Research Universities*

Institutions	No. of Times in Top 3 Bands (2012 GRBS)	In Top 63 Federal Research	2009 Federal Research ($000s)	2009 Federal Research Natl Rank
Harvard U	10	x	385,704	18
Massachusetts Inst of Technology	10	x	532,618	4
Stanford U	10	x	477,507	10
U of California - Berkeley	10	x	262,069	38
Rice U	9		56,270	137
Yale U	9	x	378,914	19
Columbia U	8	x	483,111	9
U of California - Los Angeles	8	x	467,505	11
California Institute of Technology	7	x	305,682	26
Georgia Institute of Technology	7	x	322,452	23
Princeton U	7		128,876	77
U of California - San Diego	7	x	511,428	5
U of California - Santa Barbara	7		113,837	90
U of Chicago	7	x	301,159	27
Carnegie Mellon U	6	x	170,260	61
Cornell U	6	x	238,022	45
U of Colorado - Boulder	6	x	239,687	44
U of Pennsylvania	6	x	499,498	7
Washington U - St. Louis	6	x	414,045	15
Boston U	5	x	255,178	41
Northwestern U	5	x	300,619	28
Tufts U	5		102,330	98
U of Michigan - Ann Arbor	5	x	636,216	2
U of Massachusetts - Amherst	5		80,163	112
U of Utah	5	x	192,354	56
U of Washington - Seattle	5	x	619,353	3
Arizona State U	4		134,598	73
Brown U	4		93,753	103
Dartmouth College	4		134,113	74
Duke U	4	x	438,767	13
New Jersey Institute of Technology	4		42,656	157
New York U	4	x	202,535	53
Pennsylvania State U - University Park	4	x	386,635	17
Stony Brook U	4		107,396	94
U of California - Riverside	4		53,971	141
U of California - Santa Cruz	4		76,085	115
U of California - Irvine	4	x	177,098	58
U of Cincinnati	4	x	229,324	48
U of Illinois - Urbana-Champaign	4	x	288,013	34
U of Maryland - Baltimore County	4		53,867	144
U of Maryland - College Park	4	x	246,985	43
U of Minnesota - Twin Cities	4	x	390,602	16
U of North Carolina at Chapel Hill	4	x	431,837	14
U of Southern California	4	x	375,024	20
U of Texas - Austin	4	x	309,125	25
Boston College	3		22,672	204
Clarkson U	3		7,105	293
Colorado School of Mines	3		25,109	186
Emory U	3	x	295,831	32
Johns Hopkins U	3	x	1,587,547	1
North Carolina State U	3	x	135,318	71
Ohio State U - Columbus	3	x	339,820	21
Oregon State U	3		118,252	85
Purdue U - West Lafayette	3	x	175,302	59
U of California - Davis	3	x	295,924	31
U of Delaware	3		87,090	108
U of Idaho	3		42,207	159
U of New Mexico	3		133,334	75
U of Pittsburgh	3	x	463,192	12
U of Texas - Dallas	3		25,651	184
U of Wisconsin - Madison	3	x	507,898	6
Wake Forest U	3	x	144,454	69
Yeshiva U	3	x	137,108	70

** Excluding standalone medical and other specialized institutions.*

- Second, tools such as the GRBS provide institutions and their constituents with the ability to define their competitive framework in a detailed, consistent, and reliable fashion, and then mobilize their resources to compete within that framework. Few universities can perform at high levels in all fields, and with very high standards for performance criteria, no university can demonstrate high levels of achievement in all fields.

- Third, the GRBS project, by demonstrating the challenge of accurately measuring research significance across 15 major fields and 253 subfields, encourages the careful benchmarking of specific academic specialties to help guide institutional investments into the most productive uses.

- Fourth, in most instances, when evaluating top institutional performance, the same group of ten to fifteen institutions will almost always appear at the top of the list. This indicates that a few institutions have the resources, infrastructure, and long term commitment that translate into sustained high levels of performance in most academic research fields. Other institutions, however, will perform exceptionally well in some fields and not in others, making the choice of criteria for any evaluation of particular importance in developing strategies for improving competitive institutional performance.

This conversation about rating institutions may frustrate because it does not provide simple answers to the complicated question of defining individual and institutional academic performance. Mostly this is the result of a lack of clarity in defining the question we hope to answer. Do we want to know how well research institutions perform on a range of functions including research, graduate instruction, doctoral degree production, financial resource acquisition, and the like? If so, the MUP Center's annual reports provide, at least for the United States, a useful answer.

Do we want to know the relative quality of the research publications generated by the staff of an institution? If so, the GRBS data provide an opportunity to construct an answer. The question here, however, is too broad. If we want to know how well the publications appearing within a particular field or set of fields rank on a carefully constructed set of prestige criteria, then the GRBS offers an opportunity to answer such a carefully circumscribed question. If, however, we simply want to create a league table of great universities, the GRBS data will not give a good answer precisely because we have asked the wrong question. GRBS can tell us what the top 50 to 70 best performing research groups in the United States might be, but it will not rank order them in any useful fashion because such a ranking would offer only the illusion of accuracy.

Research institutions are critical engines of quality for every nation, and the GRBS provides an international perspective on the elements that drive institutional quality in not only the 15 top-level fields but also the 253 subfields included within its data. By using the GRBS data carefully, and exploring the rich resources available on the website at http://researchbenchmarking.org, institutions can benchmark their performance to an appropriately chosen group of institutions to help determine how best to maintain current exceptional performance in various fields and where new investment might provide additional high levels of competitive quality.

Note: The literature on bibliometrics is extensive, detailed, and often highly technical. For a review of the metrics used by the GRBS see the website section on methodology at http://www.researchbenchmarking.org/web/guest/methodology. In addition, for the *h*-index see Jorge E. Hirsch, "An Index to Quantify an Individual's Scientific Research Output". Proceedings of the National Academy of Sciences (102:46, November 15, 2005); for SNIP see Henk F. Moed, "Measuring Contextual Citation Impact of Scientific Journals," Journal of Informetrics (4:3, 2010). A useful survey of world rankings is in Andrejs Rauhvargers, "Global University Rankings and Their Impact," European University Association (EUA) Report on Rankings, 2011 and a defense of citation metrics is in Henk F. Moed, Lisa Colledge, Jan Reedijk, Felix Moya-Anegon, Vicente Guerrero-Bote, Andrew Plume, Mayur Amin, "Citation-based Metrics Are Appropriate Tools in Journal Assessment Provided that They Are Accurate and Used in an Informed Way," Scientometrics (92:2, 2012).