Since educational statistics, which are relatively easy to obtain, can only attempt to measure "quality," this paper asks how quality in higher education is assessed and how educational rankings, which are defined as benchmarks or attempts to measure, contribute to this process. The paper notes that while attempts to rank institutions of higher education date to the late nineteenth century, two seminal studies published in the 1980s and 1990s generated immense interest in the academic community and prompted much analysis. Since then there has been a proliferation of annual publications by the media, but the paper cautions that many are not published on a timely basis nor are their data comparable. Some discussion is devoted to the four major ranking methodologies—reputational rankings, citation analysis, faculty productivity, and statistical rankings—and five articles, which are considered classic studies, are noted as are efforts at international rankings. The paper then offers eight requirements for perfect educational rankings, including standardized reporting methods, peer review, multidimensional techniques, and focus on interdisciplinary programs and emerging disciplines. The paper concludes that while statistics-based rankings are based on hard facts, consumers need to be aware of the shortcomings of opinion-based rankings. (Contains 34 reference notes.)

(CH)
EDUCATIONAL RANKINGS OF HIGHER EDUCATION: FACT OR FICTION?

A Paper Presented at the
Eighth International Conference on Assessing Quality in Higher Education
July 15, 1996
The Pan Pacific Gold Coast
Queensland, Australia

By
Lynn C. Hattendorf
Associate Professor
of
Library and Information Science
The University of Illinois at Chicago
The University Library (MC/234)
Box 8198
Chicago, IL 60680
USA

FAX: 312-413-0424
Internet: U48417@UICVM.uic.edu
Telephone: 1-312-996-2728

BEST COPY AVAILABLE
Educational Rankings of Higher Education: Fact or Fiction?

Introduction
Assessing quality in higher education—where are we now? Where have we been? Where are we going? These are questions which certainly will accompany us into the next millennium. The objective assessment of any institution or any individual is difficult in the extreme. Judging by the variety of presentations at this conference, it is not an issue for which there are any prosaic panaceas. It is an intellectual area which remains an unmitigated challenge to social science researchers because it invites, indeed, it beckons us with the immense possibilities that exist for inventing assessment strategies that work—that really work. The challenge is in discovering just what these are and how they can accomplish successfully what they are designed to do.

What is an educational ranking? David Webster, the "godfather" of contemporary rankings research, defines an academic quality ranking as one which meets two major criteria. First, it must be arranged according to some criterion or set of criteria which measures or reflects academic quality. The second requirement is that it must be a list of the best colleges, universities, or departments in a field of study, in numerical order according to their supposed quality, with each institution or department having its own individual rank.

Educational rankings provide benchmarks by which prospective students, educational administrators, trustees, legislators and policymakers, managers of public and private funding agencies and others in this arena have something concrete—whether fact or fiction—by which to measure and evaluate the quality of education. Budgetary, funding, recruitment, and personnel decisions, strategic and long range planning and philanthropic efforts are frequently formulated on the basis of published rankings. The immense importance and popularity of rankings cannot be either ignored or denied. They are here to stay as long as society is concerned with the three B's—the biggest, the brightest, and the best.

Educational rankings attempt to measure quality. Quality is both a subjective and a nebulous term. The yardstick of educational quality is primarily based upon the intellectual productivity of an institution's faculty; i.e., published research and citation counts of their research, funded grant proposals, faculty awards, honors, and prizes, institutional patents, etc., all of which lend, at the very least, a veneer of elitism to their respective institutions.

Statistics reflecting academic resources such as educational expenditures per student, faculty-student ratios, library holdings, and retention rates are also used to determine educational quality. These types of statistics are relatively easy to obtain and manipulate. The fact is that every institution measures these factors differently. What these statistics do not tell us and what they cannot tell us is how well the faculty teaches or how much their students learn.

Compiling educational rankings is a difficult, complex, and labor-intensive process because these rankings attempt to measure quality. "Attempt" is the key word.
In terms of their popularity and widespread diffusion, it can be concluded that they have succeeded in their attempt to assign numerical designations to departments, institutions, and individuals. Whether these numerical designations are inherently valid is open to scrutiny. Seeing something in print is oftentimes all that is needed for people to interpret fiction as fact.

**Historical Perspective**

Rankings of institutions of higher education and of individual departments and disciplines have been in existence since the late nineteenth century. Their origins can be traced as far back as 1865 to several European studies which promulgated the nature versus nurture debate. That is to say, whether heredity or environment was the determining factor in producing men of genius. At this point in history, a few women were included in these studies although they were always referred to under the rubric of men. An analysis of this phenomena is best left to scholars in other disciplines. These first attempts to measure the overall quality of institutions and the scholars affiliated with them were concentrated in the areas of science and medicine. They received widespread discussion throughout the academic sector as evidenced by the existing literature. There is no question that these early rankings influenced greatly the thinking of educators concerning quality assessment.

North America is the most prolific producer of educational rankings. There are many reasons for this but chiefly it is one of quantity. In the United States there are 3,600 institutions of higher learning from which students must choose. It is obvious that numbers play a crucial role in the process of ranking institutional quality. There is an inverse correlation at work here: more choices increase the difficulty of identifying accurately two of the three B's--the best and the brightest.

Six multidisciplinary reputational rankings of graduate departments in the United States were published between 1925-1979. Only one of these, the 1959 Keniston study, totalled its departmental ratings to derive institution-wide ones. These studies were instrumental in serving as springboards for two massive attempts to measure educational quality during the 1980's and 1990's.

The fascination and preoccupation with where an institution or an institution's departments rank in comparison to others escalated with the publication in 1982 of the five-volume *Assessment of Research-Doctorate Programs in the United States*. This now landmark study generated immense brouhaha from the academic community and prompted social science researchers to devote years of intense analysis to dissecting its findings. Webster dubbed the *Assessment* "the Rolls-Royce of academic quality rankings," because it was "by far the biggest, best, most expensive, most thoughtfully conceived and carefully carried-out academic quality ranking yet done."

The *Assessment* examined sixteen measures of quality at 228 universities and covered 2,699 programs in thirty-two disciplines. A major undertaking to be sure but it was far from perfect. Institutions in each discipline were listed not in descending order from the highest to the lowest-ranked but rather in alphabetical order. By deciding not to rank the institutions by quality, the compilers made it exceedingly difficult to determine
an institution's standing in any discipline or to compare institution's programs with each other.\textsuperscript{10}

Over a decade later, in September of 1995, the (NRC) National Research Council, published what is already considered the definitive stand-alone analysis of doctoral-granting institutions. This 740-page report, \textit{Research-Doctorate Programs in the United States: Continuity and Change}, took more than four years of planning, research, and preparation.\textsuperscript{11} It has surpassed the NRC's 1982 \textit{Assessment} in its scope, methodology, and analysis. It elicited an immediate response from all sectors of American education and government. This long-awaited second study increased its coverage to 3,634 doctoral programs at 274 American universities in forty-one disciplines,\textsuperscript{12} the largest single attempt ever of this type.

Charlotte Kuh, executive director of the NRC's Office of Scientific and Engineering Personnel, stated:

"This book is not casual, light reading, but a reference document for at least the next decade."\textsuperscript{13}

The major deficiency of the 1995 study is that it does not indicate how individual Ph.D. programs' ratings changed from 1981 to 1993, when the second survey was conducted. Daniel Zawlewski summed succinctly the deficiencies of the 1995 Report. He stated,

"Its first 146 pages form a sort of how-not-to manual, in which the authors detail the hundreds of things the data don't reveal."\textsuperscript{14}

It was not until the 1980's, however, that educational rankings became highly visible to mainstream society. Why was this? More than any other factor, it has been the media in all its many manifestations which has been responsible for bringing to the attention of everyone, not just academics, the power that rankings exert upon society as a whole. Publicity by the media of published rankings has at once embellished, enhanced, distorted, and confused their original purposes.

The 1980's witnessed a proliferation of annual publications of educational rankings studies in the United States and Canada. The best known and most popular of these are the \textit{U.S. News & World Report}'s fall rankings of America's best undergraduate colleges and its spring rankings of graduate and professional programs. \textit{Money} magazine, \textit{Business Week}, and \textit{Redbook} picked up on this trend by devising and publishing their own rankings of the best colleges and universities for tuition and expenses, those having the best business schools, and the states with the best elementary schools respectively. Not to be outdone, the Canadian magazine, Maclean's, began rating and ranking their institutions of higher education based on a methodology similar to that utilized by \textit{U.S. News & World Report}. Rankings published in popular magazines have the potential to reach the greatest number of users even though better rankings data may be available in scholarly journals which are published for a limited audience.\textsuperscript{15}

Since their inception, these annual published rankings by the popular press have constituted each magazine's top-selling issues. The media have capitalized on these rankings to the point that each new issue makes the front page of the national newspapers and receives top billing in the nightly newscasts. The widespread availability of the
Internet has promoted spirited ongoing discussions among academics from around the world concerning the validity, reliability, and methodology of these rankings.

Those institutions which are fortunate enough to win a place in these journalistic rankings studies have used their rankings as a basis for recruitment and as a vehicle for public relations. In 1990, an article in The College Board Review cautioned that those in the education profession were allowing journalists to measure the quality of institutions and education, to act as authorities in a field that is much more complex than they could possibly understand. Furthermore, it was stated that this is largely the fault of educators who do not possess consistent standards with which to measure quality. In effect, the education profession has largely abdicated its rightful position by failing to counteract these journalistic endeavors through devising alternative and better methods of assessing quality and compiling educational rankings.

Most educational rankings do not appear on an annual or even a timely basis. Because of the complexity of devising research methodologies and the prohibitive length of time it requires to conduct surveys, compile results, submit them for publication, and ultimately get the results published, the existence of this time lag presents a very real concern. For example, the publications by Franklin and Marshall College, Baccalaureate Origins of Doctorate Recipients, ranked over 900 private undergraduate colleges and universities which were responsible for producing the most graduates who went on to earn doctoral degrees over time. The seventh and last edition was published in March, 1993, and analyzed data from 1920-1990. The results of these studies are particularly significant because they identify indubitably the best and the brightest although not necessarily the biggest institutions, which have consistently produced highly academically successful graduates.

In 1967 the National Academy Press began issuing an annual publication called Summary Report: Doctorate Recipients from United States Universities. This report extends the findings of the Franklin and Marshall studies by calculating such data as the countries with the most students earning Ph.D.'s in the United States, the percentage of Ph.D.'s awarded to minorities, and the number of Ph.D.'s awarded each year by discipline, citizenship, race, and ethnicity. The Institute of International Education publishes two annual statistical compilations, OPEN DOORS and Profiles: Detailed Analyses of the Foreign Student Population. Both of these receive high visibility within the professional literature. All of these findings create a set of interesting educational rankings based on graduate numbers although they do not, in essence, measure institutional or departmental quality.

The Annual Survey of Colleges, published by The College Board, Rankings of the States published annually by the National Education Association, and the annual Digest of Education Statistics published by the National Center for Education Statistics provide the statistical fodder for countless publications which manipulate, massage, and repackage this data into what are often misconstruable renditions of the original numbers.
Ranking Methodologies

There are four major ranking methodologies which are currently employed in the quest to measure institutional, individual, and departmental quality. Some educational rankings are based solely on one of these four, while others employ a combination of them. Briefly summarized, they are as follows:

1. Reputational Rankings

Rankings predicated from this methodology are derived from the opinions of individuals who are considered to be in positions to know who are the most influential and prolific scholars in a field or which are the highest-quality institutions. These studies are based on the subjective opinions of select groups of people. The technique of peer assessment requires faculty to rank their peers at other institutions. In effect, they are being asked to rate their former professors and/or students. It is obviously in their best interest to rank their alma maters highly. The large number of graduates from highly ranked institutions at all levels enables them to play a major role in shaping opinion within a discipline. This technique perpetuates academic inbreeding ad infinitum. Herein lies the danger that rankings based principally on reputation are subjective in the extreme.

2. Citation Analysis

This is a very popular method of assessing the influence and intellectual significance of research over time. This technique counts the number of times a published paper is cited in the professional literature. In essence, individual departments and institutions are ranked according to the impact their faculties exert within their individual research areas. The more citations to one's research, the higher the rank order of a department or an institution when compared with other similar departments and institutions. Its use is predicated on the assumption that the more a publication is cited, the greater its quality and/or impact. Citation analysis has been called the "Nielsen rating" of science because of its widespread use to judge quality and because of its relative ease of use. While there are many drawbacks to this technique of assessing educational quality, the primary flaw is that citation analyses do not distinguish between good, neutral, or unfavorable citations nor do they make any adjustments for self-citations.

The Philadelphia, Pennsylvania-based Institute for Scientific Information (ISI) founded by the doyen of citation analysis, Eugene Garfield, publishes the Science Citation Index, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. Begun in 1961, this method of assessing the research productivity of individuals and institutions has escalated into a number of additional publications which are avidly consulted by scholars from all disciplines. These are Current Contents, The Scientist, and Science Watch. The ISI rankings are also the subject of intense and caustic criticism by those who are not convinced that high citation counts are necessarily indicative of quality or that high citation counts, in themselves, mean what they appear to mean.

3. Faculty Productivity

The research and publication productivity of faculty is measured by counting the number of publications scholars have published during a particular time period. Faculty
productivity is measured solely by published research output, regardless of whether this research has ever been cited. The more publications faculty have, the higher the rankings of themselves, their departments, and their respective institutions. While many arguments have been made against these types of studies because they measure quantity, not quality or teaching, the counter argument is that faculty who perform research on a continuing basis will stay current with the latest developments in their fields. This, in turn, should make them better teachers, but that is not always the case. To state the obvious, quantity and quality do not necessarily go hand-in-hand. Nevertheless, the facts speak for themselves. In almost all educational rankings, the highest-ranked institutions consistently have tenured faculty who are prolific scholars in the areas of research and publication.23

4. Statistical Rankings

These are numerical rankings derived from such arbitrary information as institutions having the most selective admissions policies, the most National Merit Scholars, the highest tuition and fees, the lowest campus crime rates, etc. The statistics are often significant in themselves but what do they actually reveal about educational quality? The major deficiency with this type of assessment technique is that too much importance or the wrong significance will be attached to them in measuring the quality of education.24 Enormous possibilities exist for the distortion, manipulation, and endless repackaging of this type of data.

Arthur Rothkopf, President of Lafayette College, stated that although many college administrators have reservations about the value of rankings as a measure of the quality of academic programs, the college-going public takes them very seriously. Some institutions are "fudging" or "cooking" the statistics they provide for the rankings. If the published numbers cannot be relied upon, tens of thousands of students are making important decisions based on false data—or fiction. Many faculty members, administrators, and trustees also are making decisions and gauging their own institution's performance based on what may be inaccurate information. The resulting potential for damage is daunting.25

**Classic Rankings**

There is a body of educational rankings which can be considered "classics." The term classics as applied to rankings means those studies which stand alone because of their focus on an individual facet of education over time. These one-of-a-kind rankings studies often span decades and spawn further research. This genre of rankings research stands alone and its value increases over time because these surveys are unique, retrospective, and comprehensive.26 The following five articles are representative of some of the best and most interesting of these classic studies. Their titles reflect the scope of these rankings:

1. "Top-Published Authors in Communication Studies 1915-1985."
5. "Top 50 Institutions in Molecular Biology Ranked by Citation Impact, 1981-91."[27]

**International Rankings**

Interest in educational rankings is by no means limited to North America, even though it is the largest single producer of this genre. For example, the Federal Republic of Germany which had 244 institutions of higher learning as of 1988, conducted several major studies on research performance evaluation from 1975-1988. Although this first generation of German evaluation studies exemplifies interesting methodological approaches, it should come as no surprise that they have not yet developed a standardized evaluation methodology.[28]

Penelope Murphy, of the Centre for Research Policy at the University of Wollongong, presented benchmarks for publications productivity for Australian academic researchers. These were based on an empirical study of all reported 1991 publications output from the 36 university-level institutions. Murphy found that 58 percent of all journal articles published by Australian academics were published in journals which were included in the citation indices of the ISI. At the institutional level, the tendency to publish in journals indexed by the ISI was seen, at the aggregate level, as a good indicator of the research quality of the institution.[29] While the number of citations does not measure actual quality, they do provide a measure of the use which the research community makes of an article. If good articles are to be used they must also be visible and a widely circulated journal is the best vehicle for achieving this.[30]

Another Australian researcher has conducted an unprecedented type of survey which attempts to assess quality based on whether a university is strongest in the areas of knowledge generation, knowledge transmission, or knowledge application. Wolfgang Grichting of Australian Catholic University devised a ranking of the 35 Australian public universities according to each of these factors.[31] The results exhibit great disparities among these institutions, and Grichting's survey displays a refreshing new way of approaching quality assessment.

**Requirements for the Perfect Educational Ranking**

1. The adoption of across-the-board standardized reporting methods by colleges and universities.
2. Peer review within specializations but not across specializations.
3. Multidimensional techniques which utilize several of the current major methodological strategies and which are based on multiple variables.
4. Measuring the quality of education that students receive in terms of what they learn.
5. More educational rankings are needed which focus on interdisciplinary programs and on important fields of study within departments, what David Reisman has labeled
"microclimates." For example, a graduate student in English who aspires to become a Milton scholar, will be more interested in those institutions which are the strongest in this area rather than in an institution's overall ranking in the discipline of English.

6. New and emerging disciplines need to receive attention, especially as students prepare themselves to enter a radically changing workforce.

7. The contributions of ancillary departments to the department being ranked should be acknowledged. Scholars from other departments often publish in core disciplinary journals which are outside the realm of their major expertise. For example, David Webster, one of the foremost academic quality rankings experts, has published an article about educational rankings in RQ, a major journal in the field of library and information science.

8. Assessment of research performance should cover a period of at least three to five years in order to reflect the stability of sustained research, changes in research performance, and to allow a reasonable length of time for these publications to appear in the citation indices.

9. In the assessment of individual departments, those possessing collective strength should be distinguished from those with individual strength. There are many departments which have only one or two prolific scholars. A ranking formulated on faculty productivity of an overall basically unproductive department will present a distorted view in the extreme.

**Conclusion**

The question remains--are educational rankings fact or fiction? Statistics-based rankings are based on hard facts. As such, they are inherently factual. Opinion-based rankings are just that, regardless of the reputations of those who are asked to make these judgments. The majority of the public who consult rankings accept them as fact even though they may well be fiction. Most consumers do not analyze the methodology or the results of educational rankings with jaundiced eyes as do the majority of social scientists. Consumers of rankings and rankings researchers need to be aware of the shortcomings that exist in these rankings, what impact these shortcomings have on the rankings, and what methodological changes might improve them.32

Because the organizational patterns of departments in all disciplines and fields of study vary substantially, there are no national or international rankings that have been developed which can be applied systematically across disciplines or even within one discipline.33 All educational rankings must be examined and compared with other rankings. No one ranking should be evaluated or judged in a vacuum. There is as yet no individual ranking in any area of education that can be considered the definitive ranking.34 As we approach a new millenium, we must strive to create at least one methodology which can serve as a model for the educational rankings of the twenty-first century so that fiction can be transformed into fact.
References


3. Ibid., p. 19-23, 158.


7. Ibid., p. 19-23, 158.


9. Webster, p. 121.

10. Ibid., p. 15.


12. Ibid., p. 2.


23. Ibid.

24. Ibid.


33. West, p. 151-52.
I. DOCUMENT IDENTIFICATION:

Title: Educational Rankings of Higher Education: Fact or Fiction?

Author(s): Lynn C. Hattendorf

Corporate Source: The University of Illinois at Chicago

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

XX Sample sticker to be affixed to document

Check here

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

_______ Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample or here

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

_______ Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Lynn C. Hattendorf

Printed Name: Lynn C. Hattendorf

Position: Associate Professor

Organization: The University of Illinois at Chicago

Address: The University of Illinois at Chicago, The University Library (MC/234), Box 8198

Telephone Number: 312 (996-2728

Date: September 30, 1996

City: Chicago, IL 60680-8198
III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of this document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents which cannot be made available through EDRS).

<table>
<thead>
<tr>
<th>Publisher/Distributor:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Price Per Copy:</th>
<th>Quantity Price:</th>
</tr>
</thead>
</table>

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

<table>
<thead>
<tr>
<th>Name and address of current copyright/reproduction rights holder:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
<td></td>
</tr>
<tr>
<td>Address:</td>
<td></td>
</tr>
</tbody>
</table>

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ASSESSMENT AND EVALUATION (TM)
Catholic University of America
210 O'Boyle Hall
Washington, D.C. 20064-4035

If you are making an unsolicited contribution to ERIC, you may return this form (and the document being contributed) to:

ERIC Facility
1301 Piccard Drive, Suite 300
Rockville, Maryland 20850-4305
Telephone: (301) 258-5500

(Rev. 12/91)