Report to the American Sociological Association Council Regarding the 2010 National Research Council Assessment of Doctorate Programs*

Ad-hoc Committee on the NRC Rankings:

John H. Evans, University of California, San Diego
Sally T. Hillsman, American Sociological Association, *ex officio*
Andrew Perrin, University of North Carolina, Chapel Hill
Mario L. Small, University of Chicago (Committee Chair)
Sandra S. Smith, University of California, Berkeley

February 2011

Table of Contents

I. EXECUTIVE SUMMARY

In November 2010, American Sociological Association President Randall Collins asked the authors of the present report to form an ad-hoc committee to assess the 2010 National Research Council rankings of sociology departments. The committee was asked to prepare a set of recommendations to the ASA Council on how, if at all, to respond. We have assessed the rankings, identified both operationalization and implementation problems, and discussed the larger questions that such ranking exercises raise. Based on our work, we recommend that the ASA Council issue a resolution criticizing the 2010 NRC rankings for containing both operationalization and implementation problems; discouraging faculty, students, and university administrators from using the core 2010 NRC rankings to evaluate sociology programs; encouraging faculty, students, and university administrators to be suspicious of the raw data accompanying the 2010 NRC report; and indicating that alternative rankings, such as those based on surveys of departments' reputations, have their own sets of biases.

II. INTRODUCTION

In November 2010, American Sociological Association President Randall Collins asked the authors of the present report to assess the 2010 National Research Council (NRC) rankings of sociology departments. The highly anticipated rankings, titled *A Data-Based Assessment of Research Doctorate Programs in the United States* (Ostriker et al. 2010), had incited controversy, and the committee was asked to prepare a set of recommendations to the ASA Council on how to respond. Critics of the 2010 NRC report argued that it did not evaluate sociology programs effectively, and that strong sociology programs were unfairly ranked poorly. Furthermore, several department chairs feared that the report misrepresented the quality of their programs in ways that university administrators may not appreciate. President Collins requested a report in time for the February 2011 meeting of the ASA Council. Given the short turnaround time and the fact that the NRC is expected to continue releasing documents after February 2011, the committee agreed to assemble a limited evaluation with a set of recommendations to ASA Council considered to be provisional. The committee has worked diligently for several months, and the resulting report follows.

We begin by describing the procedures employed by the NRC to produce its 2010 rankings, including a discussion of how those procedures were designed to improve upon other approaches to rankings. Next, we identify what we consider to be serious problems with the execution of the rankings. We conclude with a set of recommendations for action for ASA Council.

III. BACKGROUND: THE NRC RANKINGS

The 2010 NRC rankings and their potential problems must be understood in light of other existing rankings, whose limitations the NRC attempted to address. Of these, the three most important are the *U.S. News & World Report* rankings; the Gourman Report rankings, now published by Princeton Review; and the 1995 NRC rankings, which preceded those published in 2010. We discuss each in turn.

Existing rankings

U.S. News and World Report began ranking graduate programs in 1987. Its rankings of the sciences, social sciences, and humanities have been based solely on departmental reputations (Morse 2010a, 2010b). For the last report, *U.S. News* collected peer assessment data from deans, directors of graduate studies, and senior faculty or department chairs at institutions that had granted five or more doctorates in each discipline during the five-year period, from 2001 through 2005. These experts were asked to judge the academic quality of programs in their field on a five-point scale: outstanding (5), strong (4), good (3), adequate (2), or marginal (1). If individuals were unfamiliar with a particular school's programs, they were asked to select "don't know." "Scores for each school were determined by computing a trimmed mean (eliminating the two highest and the two lowest responses) of the ratings of all respondents who rated that school; average scores were then sorted in descending order" (Morse 2010b).[1] A department ranked highly to the extent that faculty and deans judged it to be strong or outstanding.

The *U.S. News* rankings have been criticized because of the limitations of reputation-based approaches to ranking departments (Brooks 2005). First, reputational approaches can be highly subjective, since evaluators may rank departments highly because of loyalty to their *alma mater*, hearsay, or even gossip. Second, they must rely on evaluators' imperfect information. Since evaluators have unequal and often insufficient access to the current composition and productivity of departments, reputational lags are a threat. Once-stellar departments in decline may be ranked highly long after they stopped being productive; up-and-coming departments with new exceptional faculty may have the opposite problem.[2] Third, reputational rankings can be sensitive to what has been labeled a "halo effect" (Brooks 2005). Weak departments in strong universities may be ranked highly because of the latter's reputation; conversely, strong departments in weak universities may be penalized for similar reasons. The halo effect can also derive from one or two star faculty improving the reputation of otherwise weak departments. Nonetheless, the common perception is that reputational studies capture subtle qualities and distinctions known in the field but difficult to measure accurately with statistical data (Ostriker et al. 2010: 60).

---

[1] The response rate in sociology was about 43%. To rank specialties within disciplines, they identify subjects in which schools most often offer concentrations and ask respondents to nominate schools that have outstanding programs in each specific area. Each respondent is allowed to nominate up to ten schools in any specialty area. *U.S. News* publishes the names of the schools receiving the highest number of nominations in each specialty area, based solely on the number of nominations they receive in each specialty area (Morse 2010a). Whereas *U.S. News* assesses quality of graduate programs in the sciences, social sciences, and humanities based solely on reputational data, assessments of professional schools—business, education, engineering, law, and medicine—are determined based on both reputational and statistical data. In addition to obtaining expert opinion from knowledgeable individuals in academia and practitioners in each profession, *U.S. News* collects data on input and output measures. Input data refers to such factors as the quality of students and faculty, faculty-student ratio, research funding, and other resources brought to the education process. Output measures capture indicators of success, such as passing the bar, job placement rates, starting salaries, and the like (Morse 2010a).

[2] *U.S. News*'s allowance of a "Don't know" category has done little to temper this criticism.

The Gourman Report has been assessing graduate and professional programs since 1980.[3] Since that time, eight editions have been released, the last appearing in 1997. It was published by Princeton Review and distributed by Random House. The rankings' creator, Jack Gourman, a former professor of political science at Cal State Northridge and current Director of Research at National Education Standards, has explained that his rankings are based on four key factors--curriculum, faculty instruction, faculty research, and library resources (Webster 1986).[4] However, Gourman has said little about how he gathers the data (which do not appear to be supplied to Gourman by the institutions), defines his terms, operationalizes or weighs key variables, or performs the analysis (Webster 1984; Selingo 1997; Bedeian 2002).

Because of this missing information, the validity and reliability of the Gourman Report data and rankings are difficult to assess, and the rankings have been discredited (Webster 1984; Selingo 1997; Bedeian 2002). In addition, several of the results lack face validity, such as the presence of identical rankings for all departments in the same institution (Webster 1984). In spite of these critiques, the rankings have been disseminated widely, easily available in websites that publish graduate program rankings, such as http://www.socialpsychology.org/gsociol.htm and through the Princeton Review.

The 1995 NRC rankings, which evaluated 41 fields in 274 institutions, were based on a reputational survey carried out in 1993 (Ostriker and Kuh 2003:10). The survey, called the National Survey of Graduate Faculty, asked about overall scholarly reputation, educational effectiveness, and change in quality since 1982 (http://www.nap.edu/readingroom.php?book=researchdoc&page=summary.html). The survey did not collect objective information about the programs, such as the publication record of faculty. The 1995 NRC rankings were highly correlated with the *U.S. News* report of the same period, because they were both largely based on reputational surveys. In sociology, quality scores and quality rankings overlapped considerably, with Pearson correlations of 0.97 and 0.96, respectively (Rogers and Rogers 1997).

Still, the 1995 NRC rankings were criticized for several reasons, a number of them having to do with their reliance on reputational data. For example, they were subject to the "halo effect" problem. However, they were unable to recognize effectively the presence of strong departments in niche fields or smaller programs. Second, they were unable to distinguish scholarly productivity from educational effectiveness, undermining the value of programs that, while not producing a large number of books or articles, were nonetheless particularly effective at training future scholars (Ostriker and Kuh 2003:10). Third, the 1995 rankings provided no measure of the uncertainty regarding the reported rankings. Readers could not evaluate well whether a difference of a few positions in either direction referred to an actual difference or to one due merely to chance.

---

[3] The Gourman Report has been assessing colleges and universities since 1967 (graduate and professional programs since 1980). Since that time, 10 editions have been published, the most recent in 1998.

[4] Rankings for colleges and universities are based on ten key factors---facilities, administration policies, the relationship between professors and administrators, support of faculty members, cooperation among professors, methods of communication between professors and the administration, the openness of the administration, the use of consultants and committees to solve problems, attitudes about scholarly research, and the overall cooperation of the administration (Webster 1984).

The 2010 NRC rankings

The 2010 NRC rankings were expected by many to avoid the problems of reputational approaches (such as those in *U.S. News* and the 1995 NRC) while adding the transparency missing from some existing rankings (such as the Gourman Report). The NRC's solution was to produce rankings based on objective departmental performance data, to provide five different principal rankings of program quality, and to give readers the measures used in each ranking, so that users can assess how departments rank according to, not only each of the five rankings, but also each of the several components of each ranking. Of the five, two core rankings provided overall assessments. The additional three focused on specific dimensions of program quality: "(1) the research activity of program faculty; (2) student support and student outcomes; and (3) diversity of the academic environment" (Ostriker et al. 2009:3). Data were collected on about 20 variables from surveys sent to over 200 universities and faculty members at these institutions between 2006 and 2008. The variables were derived from questions on standard measures such as number of publications, citations, and awards among the faculty, student quantitative GRE scores, and time to degree. These data were used to construct the final rankings, released in 2010 after some delays related to data quality and analysis.

The two overall assessment rankings, which are based on different methods, have produced the most discussion. Both rankings rely directly on program data collected from the institutions. The only difference is how much weight they assign to different components of the program (such as faculty publications, GRE scores, or time to degree) when determining their rank. For the first ranking, a survey of faculty asked them directly to report how important each of the components is to a program's success. Their answers were used to assign weights to the different components, such that programs were ranked highly if their objective data were high on those measures that faculty had separately said were important for a graduate program to have. This procedure produced the "S-rankings" (for "survey-based"). For the second ranking, a smaller sample of faculty in each field was asked to rate the quality of a sample of specific graduate programs from 1 to 6, in a procedure similar to a reputational ranking. The resulting variable was then regressed on the same roughly 20 variables obtained from the institutional surveys in a stepwise fashion, after some modifications. The resulting coefficients were used to assign weights to the different program characteristics, on the idea that they implicitly indicated what faculty believed was important. This procedure produced the "R-rankings" (for "regression-based"). For complete discussion of the procedures employed, see Ostriker et al. (2009, 2010), http://www.nap.edu/rdp/ and http://graduate-school.phds.org/about/ranges.

The NRC decided on its own which program characteristics would be measured. For example, the faculty survey did not offer faculty the opportunity to specify what kinds of scholarly productivity or student quality mattered most, just the extent to which scholarly productivity mattered in comparison to student quality. This decision was consequential. For example, faculty publications in sociology did not include books, and citations to faculty did not include citations to books, because the NRC measure of scholarly productivity in the social sciences was based entirely on articles. Also, GRE scores of incoming students for sociology were measured only using the quantitative GRE score. The decision to impose these restrictions was not based on the faculty survey. Instead, the survey asked respondents about the relative importance of "Number

of publications (books, articles, etc.) per faculty member" and "Median GRE scores of entering students," respectively.

To deal with the fact that any ranking involves uncertainty, the committee followed a procedure that allowed them to report a range of rankings instead of a single rank for each program (Ostriker et al. 2010:13, 53-55). For each ranking type, each program's rank was calculated 500 times, each time using a randomly selected half of the data. NRC originally intended to report the resulting inter-quartile range of estimated rankings for each program (the 25th and 75th percentiles), but eventually decided to report the 5th and 95th percentiles instead. Since the distribution of ranks within that range is not reported—for example, a given program's estimated ranks could be bunched near the top, or somewhere else, or evenly distributed—the NRC discourages only taking the top, or bottom, or average of the estimated rankings as a single measure of a program's rank. Instead, readers should assume that a given program's true rank likely lies somewhere between the reported bottom and top (5th and 95th percentile) figures (Ostriker et al. 2009:18). Nonetheless, many interpreters have disregarded this advice, including the third-party website phds.org. Rankings and raw data are available from the NRC's website, http://www.nap.edu/rdp/.

IV. PROBLEMS WITH THE 2010 NRC RANKINGS

Many sociologists have commented, often through informal memos for department leadership or online posts, on the problems with the NRC rankings (e.g., Freese 2010; Marwell Undated). Two co-authors of the present report, John Evans and Andrew Perrin, have been a part of that discussion as well on the sociology blogs Scatterplot (http://scatter.wordpress.org) and Org Theory (http://orgtheory.net). Our discussion relies on the previous work by these scholars and on our own assessment. The committee believes there are two sets of problems, those dealing with operationalization and those dealing with implementation. While a full discussion is not possible given time constraints, we list those problems we consider to be the most important.

Operationalization problems

- *The rankings exclude books from the measure of faculty publications* (Ostriker et al. 2010:40, 47). This decision was made across the board for science and social science disciplines; humanities disciplines did include book publications at a rate of five articles to one book. Sociology is well known to be both a book and an article discipline, and a large proportion of sociological research is published in book form. More importantly, books serve a distinct role in sociological publishing. They are often the publication outlet for historical, ethnographic, and interview data, are often read across disciplinary boundaries, and are more likely to be aimed at a broader audience (Clemens et al. 1995; Nederhof 2006; Swygart-Hobaugh 2004). The decision to exclude books understates the productivity of departments in which faculty tend to work on qualitative, historical, and theoretical research, and on research aimed at a broad or interdisciplinary audience.

- *The rankings exclude cited books from the measure of citations to faculty* (Ostriker et al. 2010:40, 47). This procedure understates the influence of book-focused departments, and penalizes those subfields of sociology in which books are the primary means of scholarly output. Many of the most important and influential works in the history of sociology and in recent sociology have been books. Examples are Theda Skocpol's *States and Social Revolutions* and William Julius Wilson's *The Truly Disadvantaged*, citations to which would be excluded in the assessments of their departments. [5]

- *The rankings disregard the quality of the outlet in which articles were published* (Ostriker et al. 2010:40). Sociology journals differ in the quality of their output and in their reputation. Nonetheless, the NRC assigned a publication in the *American Journal of Sociology,* the *American Sociological Review*, or *Social Forces* the same weight as a publication in any of the hundreds of minor journals catalogued by the database they employed. By disregarding differences in journal quality, the rankings penalize departments in which faculty tend to publish fewer papers of higher quality in stronger journals. Many scholars would consider the NRC procedure a perverse incentive to produce larger quantities of less important scholarship.

- *The rankings disregard the number of authors in each publication* (Ostriker et al. 2010:40). For example, a paper with six co-authors in the same department is counted six times. As a result, the rankings penalize departments in which faculty tend to work either alone or in small groups, while rewarding departments in which faculty tend to work in many large teams with each member making a small contribution to the whole. This decision, in turn, penalizes departments strong in subfields in which team projects with co-authored publications are not the norm.

- *The rankings disregard the non-quantitative portion of the GRE scores in their measure of the competitiveness of incoming students* (Ostriker et al. 2010:67). Science and social science departments were ranked with the quantitative GRE scores; humanities departments, with the verbal scores. This decision privileges sociology departments with quantitative-oriented programs over departments with theoretical, historical, ethnographic, or other non-quantitative orientations.

- *The rankings rely on at least one measure that can change dramatically from year to year.* The measure of student outside support is based on the number of first-year graduate students who had outside support in the single year of 2005 (Ostriker et al. 2010:44). Since the receipt of a prestigious external award has a strong stochastic component, this indicator might have looked rather different in 2004 or 2006, producing substantially different students support rankings.

---

[5] The NRC report indicates that citations to books were excluded because of constraints due to their data source, the Institute for Scientific Information. However, the ISI does permit searches for citations to books. Still, ISI does not include citations in books to either books or articles.

Implementation problems

- *The process by which publication data were gathered is not well documented.* The report states only that publications counted were "obtained by matching faculty lists supplied by the programs to the Thompson Reuters list of publications and cover publications extending back to 1981" (Ostriker et al. 2010, pp. 46, 47, 241). As a result, independent researchers have had difficulty replicating some of the rankings. For example, researchers at the University of North Carolina-Chapel Hill and the University of Virginia have performed recalculations of the faculty productivity measures and produced substantially different results. Whether these differences are real or due to chance is currently unknown because of the report's insufficient documentation. This problem may be resolved if the NRC releases the full raw data, which appears to be currently under consideration.

- *The rankings' approach to dealing with uncertainty is not well justified.* There are two potential issues. One is the technical issue that, in spite of an extensive and otherwise carefully documented methodological report (Ostriker et al. 2009), the NRC did not explain well their rationale either for using 50 percent subsamples ("random halves") or for taking 500 samples rather than a larger number, since statisticians have recommended at least 1,000 samples for estimating percentile ranges for confidence intervals (Mooney and Duval 1993:37). The other is a related but more conceptual issue regarding the method's ability to account for uncertainty. The report is right to acknowledge the presence of uncertainty, but strongly gives the impression that percentile reports (5th to 95th) accurately capture that uncertainty. In sociology, an important source of variability among raters is the kind of sociological work they value. Systematic differences across raters in sociology are not difficult to imagine and are not mere random noise. It was unclear to the committee how the NRC method accounted for this kind of error. Rather than highlighting this possibility, the resulting percentile ranks produced by the report seem to give the impression that any important source of error is captured within the reported 5th and 95th rank percentiles.

Larger issues

In addition to these particular problems, there are larger, conceptual problems that lead us to question the value of the rankings. First, sociology departments cannot be rank-ordered effectively with respect to their excellence in sociological scholarship. Many of the implementation and operationalization problems described above are solvable. For example, the NRC could include verbal GRE measures, include books in the productivity indicator, weigh sole-authored and co-authored articles differently, and develop a reliable method for assessing rank uncertainty. Nevertheless, even if all of theseproblems were addressed, the resulting improved ranking would still suffer from an important problem, that the heterogeneity of sociology is multi-dimensional. For example, while sociology is heterogeneous with respect to both substantive traditions and methodological ones, each substantive tradition is not limited to a single methodological one. The first kind of heterogeneity alone would not constitute an insurmountable problem; *U.S. News* has addressed it by producing subfield rankings. The problem is that different departments can attain excellence in the same subfield while excelling in different methodological traditions, making them non-rankable on any single dimension. For example, a department can excel in urban sociology by having a largely ethnographic or a largely

quantitative program. Single-dimension rankings of scholarly quality in sociology will likely elude us for the foreseeable future.

Second, any ranking produced by a respected scientific association will be read as a best-to-worst ordering of sociology departments by prospective students, deans, and faculty, regardless of any disclaimers offered by the producers of the ranking. The NRC, cognizant of the many limitations of its own and other rankings, offered a number of caveats in its 2010 report, including that the rankings were sensitive to errors in data input and that some departments may be ranked worse than they are because of the complexity of the ranking process (see, e.g., Ostriker et al. 2010:47). It also provided an Excel workbook allowing users to sort departments along not only the five core rankings but also the multiple dimensions. Nonetheless, with respect to overall rankings, it insisted that, "*In general, the range of rankings captures well the relative standing of most Ph.D. programs*," essentially providing readers a reason to read the rankings as a best-to-worst ordering of departments (Ostriker et al 2010:63; italics in original). This attitude about rankings, inherent in the ranking process itself, has practical consequences in the short and long term.

In the short term, the NRC ranking system can have detrimental effects on departments whose rankings are lower than their actual quality. University administrators seem to have taken the results seriously. After the release of the 2010 rankings, many universities claimed that the rankings provided evidence of their excellence. For example, all of the nine University of California campuses with Ph.D. programs and several Ivy League universities released statements indicating that the rankings revealed the excellence of either the university as a whole or its various departments.[6] This attitude may result in short-term punitive actions by administrators against departments whose rankings, unfairly or not, are low. In fact, many of the initial critiques of the 2010 NRC rankings that circulated informally among ASA members were offered by faculty aiming to explain to administrators why their ranking differed so markedly from previous rankings. Departments may also have reason to worry that, in their next review, administrators may direct them to change their programs to improve their status in future NRC rankings, for example, by hiring faculty who write articles, rather than books.

---

[6] See the following: http://ucsdnews.ucsd.edu/newsrel/awards/09-28DoctoralPrograms.asp,

http://newsroom.ucla.edu/portal/ucla/ucla-places-strong-in-comprehensive-172131.aspx,

http://newscenter.berkeley.edu/2010/09/28/nrc-rankings/, http://www.ia.ucsb.edu/pa/display.aspx?pkey=2330,

http://today.uci.edu/news/2010/09/nr_nrc_100928.php,

http://www.news.ucdavis.edu/search/news_detail.lasso?id=9625,

http://newsroom.ucr.edu/news_item.html?action=page&id=2437, http://news.ucsc.edu/2010/09/nrc-report.html,

http://news.ucsf.edu/releases/ucsf-graduate-programs-receive-top-rankings-in-national-survey/,

http://news.harvard.edu/gazette/story/2010/09/high-marks-for-doctoral-programs/,

http://www.princeton.edu/main/news/archive/S28/56/94S46/index.xml?section=newsreleases,

http://www.upenn.edu/pennnews/news/national-research-council-ranks-penn-s-graduate-programs-among-nations-best.

In the long term, the rankings may have an impact if departments permanently alter the work they support in response to the NRC evaluations. Any long-term impact would depend on the widespread acceptance of the rankings, particularly among prospective students and university administrators, as the sole objective, data-based evaluation of graduate programs. This outcome is not difficult to imagine. In the 2010 rankings, the most obvious problems are data quality and the four-year lag. Rapid advances in communication, data access, and data management make it possible that, in the near future, the NRC begins to release rankings much more frequently. One can picture a system wherein deans require departments to upload data to an online portal on a yearly basis, after which the calculation of rankings (at least of the S-rankings, which do not require asking faculty to rank departments) is easy to perform quickly. If so, it is possible that university administrators and prospective students will come to perceive the NRC rankings as the most legitimate, up-to-date, and objective measure of program quality. In fact, the 2010 NRC report explicitly sought this recognition, touting itself—even its title—as the scientific alternative to reputational rankings. Improvements in data quality and the correction of implementation errors may be all that is needed for NRC to become the *de facto* doctoral program ranking system.

If the NRC approach becomes institutionalized, long-term consequences may well ensue. The legitimation of formal rankings certainly appears to have affected the decisions of departments in other fields, such as law. The *U.S. News* law rankings have been released annually since 1990. Espeland and Sauder (2007:24) found that law schools reacted in three ways: changing their allocation of resources, altering organizational procedures such as job requirements and admission policies to conform with ranking criteria, and outright manipulation through gaming strategies. For example, since the method measures the percent of graduates who are employed, some schools hire their own graduates in temporary jobs to inflate their numbers. Other schools re-define the student status of the weaker students to remove them from calculations. If the NRC rankings become regularly adopted in their current form, it is possible that sociology departments will begin to value article-based, citation-heavy, copious output, team-based and quantitative research over other models in order to compete in the rankings. This shift would have serious consequences for the development of sociological knowledge, particularly because it would discourage many kinds of work, such as ethnographic or historical projects requiring long periods of gestation, ambitious quantitative projects requiring original data collection, or high-risk "big idea" theoretical books requiring years for their significance to be recognized.

V. RECOMMENDATIONS

The committee recommends that the ASA Council issue a resolution criticizing the 2010 NRC rankings for containing both operationalization and implementation problems; discouraging faculty, students, and university administrators from using any of the five 2010 NRC rankings to evaluate sociology programs; encouraging faculty, students, and university administrators to be suspicious of the spreadsheet data produced by the report; and indicating that reputational surveys have their own sets of biases.[7]

---

[7] While reputational surveys have the advantage that they measure reputation better than the 2010 NRC measured objective performance, they are still flawed indicators of department quality for the reasons discussed above.

The committee also recommends making the present report (or selections thereof) public. Since the NRC is currently considering the release of further documents and continues to evaluate the future of its ranking program, any additional actions by ASA Council vis-à-vis the NRC depend on conditions uncertain at the time of this writing.

Beyond these short-term, targeted actions, the committee discussed what future actions the ASA might undertake. Rankings, and debates about their meaning, are unlikely to disappear. Debates about rankings during the 1960s, for example, centered on the same core questions of the relative merits of rankings based on reputation vs. those based on observed productivity (Lewis 1968). Yet administrators will continue to have to evaluate whether their programs are doing well and students will continue to need guidance on where to apply. The committee discussed several possibilities. One possibility was for the ASA to produce its own reputational rankings, to retain control over what is perceived as quality in sociological research. This possibility was dismissed by the committee for its logistical difficulties and the limitations of single-dimension rankings of any kind as discussed above. Another possibility was for ASA to produce a database and associated multi-dimensional measures for those interested to be able to evaluate different programs across different criteria. This model would eschew overarching rankings of any kind, forcing readers to compare programs only along dimensions they find important. The committee also declines to make this recommendation because of the logistical difficulties involved and because future developments in the availability of data on sociology programs may make the endeavor unnecessary. This endeavor also might lead to unintended consequences with regard to subsets of the sociological profession whose contributions might be difficult to measure or quantify.

At this time, the committee believes that ASA should encourage prospective students, faculty, university administrators or others evaluating a given program to avoid blind reliance on rankings that claim explicitly or implicitly to list departments from best to worst. The heterogeneity of the discipline suggests that evaluators should first determine what characteristics they value in a program and then employ available sources of information to assess the program's performance. In addition, the ASA should help facilitate, within available means, the dissemination of such information.

REFERENCES AND FURTHER READING

Alderman, Geoffrey. 1996. "Audit, Assessment and Academic Autonomy." *Higher Education Quarterly* 50(3):178-92.

Bedeian, Arthur G. 2002. "Caveat Emptor: The Gourman Report," http://www.siop.org/tip/backissues/tipjan02/07bedeian.aspx. (2002).

Brooks, Rachelle L. 2005. "Measuring University Quality," *The Review of Higher Education* 29(1):1-21.

Clemens, Elisabeth S., Walter W. Powell, Kris McIlwaine, Dina Okamoto. 1995. "Careers in Print: Books, Journals, and Scholarly Reputations." *American Journal of Sociology* 101:2 (Sep., 1995), pp. 433-494.

Hoover, Eric. 2010. "The Rankings, Rejiggered," *The Chronicle of Higher Education*, August 17.

Espeland, Wendy Nelson and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113(1):1-40.

Freese, Jeremy. 2010. "NRC Rankings for Sociology." Unpublished memorandum.

Lewis, Lionel S. 1968. "On Subjective and Objective Rankings of Sociology Departments." *The American Sociologist*. 3(2): 129-131.

Marwell, Jerry. Undated. "Latest NRC Departmental Ratings." Unpublished memorandum.

Mooney, Christopher Z. and Robert D. Duval. 1993. *Bootstrapping: A Nonparametric Approach to Statistical Inference*. Newbury Park, CA: Sage.

Morse, Robert. 2010a. "Frequently Asked Questions: The Grad Rankings," http://www.usnews.com/education/best-graduate-schools/articles/2010/04/15/frequently-asked-questions-the-grad-rankings.

Morse, Robert. 2010b. "The Social Sciences and Humanities Rankings Methodology," http://www.usnewsuniversitydirectory.com/graduate-schools/methodology-sshumanities.aspx.

Nederhof, Anton J. 2006. "Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A review." *Scientometrics* 66:1 (2006): 81-100.

Ostriker, Jeremiah P., Paul W. Holland, Charlotte V. Kuh, and James A. Voytuk, eds. 2009. *A Guide to the Methodology of the National Research Council Assessment of Doctoral Programs*. Washington, D.C: National Academies Press.

Ostriker, Jeremiah P., Paul W. Holland, Charlotte V. Kuh, and James A. Voytuk, eds. 2010. *A Data-Based Assessment of Research-Doctorate Programs in the United States*. Washington, D.C: National Academies Press.

Ostriker, Jeremiah P. and Charlotte V. Kuh, eds. 2003. *Assessing Research-Doctorate Programs: A Methodology Study*. Assisted by James A. Voytuk. Washington, D.C: National Academies Press.

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.

Rogers, Evan and Sharon J. Rogers. 1997. "'High Science' vs. 'Just Selling Magazines'? How the NRC and U.S. News Graduate Rankings Compare," *AAHE Bulletin* 7-10, May.

Sauder, Michael and Wendy Nelson Espeland. 2009. "The Discipline of Rankings: Tight Coupling and Organizational Change." *American Sociological Review* 74:63-82.

Sauder, Michael and Ryon Lancaster. 2006. "Do Rankings Matter? The Effects of U.S. News & World Report Rankings on the Admissions Process of Law Schools." *Law and Society Review* 40:105-34.

Selinger, Jeffrey. "A Self –Published College Guide Goes Big-Time, and Educators Cry Foul; They say The Gourman Report is based on shoddy research that will confuse students." *Chronicle of Higher Education*, 44(11): A45-A46.

Shore, Cris and Susan Wright. 1999. "Audit Culture and Anthropology: Neo-Liberalism in British Higher Education." *Journal of the Royal Anthropology Institute* 5:557-75.

Strauss, Valerie. 2007. "U.S. News's College Rankings Face Competition and Criticism," *The Washington Post*, August 17.

Swygart-Hobaugh, Amanda J. . 2004. "A Citation Analysis of the Quantitative/Qualitative Methods Debate.s Reflection in Sociology Research: Implications for Library Collection Development." *Library Collections, Acquisitions, & Technical Services* 28 (2004): 180-195.

Webster, David. 1984. "Who Is Jack Gourman: And Why Is He Saying All Those Things about My College?" *Change* 16(8):14-19, 45-56.

Webster, David. 1986. "Jack Gourman's Rankings of Colleges and Universities: A Guide for the Perplexed." *Reference Quarterly* 25(3):323-31.