

POLICY AND GLOBAL AFFAIRS
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

**ASSESSING RESEARCH-DOCTORATE
PROGRAMS: A METHODOLOGY STUDY**



Research-doctorate programs graduate over 40,000 PhDs per year. These individuals staff the nation's laboratories and universities, write books, shape public policy, and help us to understand the human condition. They produce, apply, and disseminate the discoveries that are vital to a vibrant economy.

Surprisingly, doctoral programs are rarely scrutinized, but in the past 20 years, when scrutiny has occurred, it has been carried out by the National Research Council. This report, *Assessing Research-Doctorate Programs: A Methodology Study*, takes a close look at the methodology used in earlier studies and recommends broad changes. These changes will make the next assessment more informative to students, faculty, administrators, and the public concerned with doctoral education in the United States.

The report concludes that it is critical that we, as a society, be able to compare doctoral programs, assess their quality, and provide information about these programs for doctoral students choosing programs, for faculty responsible for developing them, and for administrators charged with making wise program investments. As advanced knowledge and its transmission grow in their importance to our society and economy, the general public needs assurances that our doctoral programs will continue to pursue and deliver effective, high-quality education for our students.

Given the importance of research doctorate programs to the future of our society, economy, and general welfare, it is essential to provide a reliable, impartial, transparent, and responsive means to address questions of quality and stimulate ongoing consideration of improvement by programs, universities, and the state, federal, and philanthropic funders of doctoral study.

WHAT WERE THE STRENGTHS AND WEAKNESSES OF EARLIER STUDIES?

The strengths of the 1995 Study identified by the Committee were:

- *Wide acceptance.* It was widely accepted, quoted, and utilized as an authoritative source of information on the quality of doctoral programs. The ratings were derived from asking faculty involved in doctoral education their views of the scholarly strength of the faculty in peer programs
- *Comprehensiveness.* It covered 41 of the largest fields of doctoral study and over 3000 faculty participated.
- *Transparency.* Its methodology was clearly stated.
- *Temporal continuity.* For most programs, it maintained continuity with the NRC study carried out ten years earlier.

The weaknesses of the 1995 Study were:

- *Data presentation.* The emphasis on exact numerical rankings encouraged study users to draw a spurious inference of precision. The user community often treated the program rankings, which were presented in strict numerical order, as something precise. In fact, the rankings were derived from ratings that varied considerably across raters, and the Committee felt that it was important to convey some sense of that variability
- *Flawed measurement of educational quality.* The reputational measure of program effectiveness in graduate education, derived from a question asked of faculty raters, confounded research reputation and educational quality.
- *Emphasis on the reputational measure of scholarly quality.* This emphasis gave users the impression that a “soft” criterion, subject to “halo” and “size effects,” was being overemphasized for the assessment of programs.
- *Obsolescence of data.* The period of ten years between studies was viewed as too long.
- *Poor dissemination of results.* The presentation of the study data was in a form that was difficult for potential students to access and to use. Data were presented but were neither interpreted nor analyzed.
- *Use of an outdated or inappropriate taxonomy of fields.* Particularly for the biological sciences, the taxonomy did not reflect the organization of graduate programs in many institutions.
- *Inadequate validation of data.* Data were not sent back to providers for a check of accuracy.

PRESERVING THE STRENGTHS, ADDRESSING THE WEAKNESSES

The Committee set up four panels, listed at the end of this summary, to cover key methodological areas. These were:

- Taxonomy and interdisciplinarity
- Quantitative measures
- Student processes and outcomes
- Reputation and data presentation

The Committee reviewed the work of these panels and then, with the help of eight diverse universities, fielded questionnaires that were sent to the institutions, doctoral programs, the faculty of these programs, and to students who had been admitted to candidacy.

The report contains many detailed recommendations. The key ones that address the weaknesses are:

Data Presentation. *The presentation of reputational ratings should be modified so as to minimize the drawing of a spurious inference of precision in program ranking.* Rankings are derived from combining individual ratings. These ratings vary from rater to rater. To convey that variability, the Committee examined rankings derived from different samples of the pool of raters. Repeated sampling results in a range of rankings, which can be graphically represented as a bar. For the most highly ranked programs, bars are shorter, reflecting rater agreement about these programs. Middle rank programs generally show longer bars. These programs are less well known or they may specialize in a particular area and be well regarded by some raters and not others. A number of programs have bars that cover overlapping rankings. You can still say what programs are in the “top 10” or “top 15”, but there are a number of programs whose rank can only be described as “somewhere between 15 and 30”, etc. The Report concludes that this gives a more accurate picture of the distribution of the ratings of programs.

Flawed measurement of educational quality. *Comparable information on educational processes should be collected directly from advanced-to-candidacy students in selected programs and reported. Whether or not individual programs monitor outcomes for their graduates should be reported.* The Report rejects the reputational measure used in past studies as an indicator of effectiveness of the doctoral program. Rather, it recommends collecting data from programs about important aspects of the education of students such as attrition, time to degree, extent of student support and programs to expose students to employment beyond a university research career. It also recommends a trial of a questionnaire to be answered, for selected fields, by students who have been admitted to candidacy that will ask about their employment plans, the different sorts of professional development that they have been exposed to, their view of the program environment, the infrastructure (having their own workspace, having adequate computing resources, etc.) and whether they have written articles, presented papers at conferences, etc.

Emphasis on the reputational measure of scholarly quality. *The Committee recommends that, in addition to data collected for the 1995 Study, new data be collected from institutions, programs, and faculty.* In addition, *the Committee recommends investigation of the construction of a synthetic measure of reputation for each field, based on statistically derived combinations of quantitative measures.* Program reputation results from a combination of factors, many of which can be measured directly. It is valuable to identify these factors that could contribute to the ability of a program to produce work of high scholarly quality. Examples include: publications and citations, research support, laboratory space, library resources, and interdisciplinary centers. The Report recommends that analyses of the relation between reputational and quantitative variables be conducted.

Obsolescence of data. *Data for quantitative measures should be collected regularly and made accessible in a Web-readable format. These measures should be reported whenever significantly updated data are available.* The Report recommends construction of a database of quantitative measures that universities could update periodically, ideally annually.

Poor dissemination of results. *Extensive use of electronic Web-based means of dissemination should be utilized for both the initial report and periodic updates.* There is no way that an assessment can provide all the interesting analyses of a very rich set of data. However, we recommend that the book for the next study include analytic essays using the data, which will encourage others to conduct their own analyses. These would be done by institutional researchers and scholars who specialize in higher education. The other part of the report--the data itself-- would be available in a database on the Web and on CD-Rom. The Web database would be searchable and users could extract data from it to conduct their own analyses. It would also be updateable, so that as time passed much of the quantitative data could be kept current. The Report recommends making the data more useful to potential students who are searching for the right program for them. Our study cannot give them the answer, but we can make it easy for them to find data that may help them narrow their choices. A number of professional societies also provide helpful information about doctoral programs, and the NRC study should be linked to theirs.

Use of an outdated or inappropriate taxonomy of fields. *The taxonomy of fields should be changed from that used in the 1995 Study to incorporate additional fields with large PhD production. The agricultural sciences should be added to the taxonomy and efforts should be made to include basic biomedical fields in medical schools. A new category, "emerging fields," should be included* The question of what fields should be included is still somewhat open and should be revisited when the full study begins. The Report recommends criteria for the inclusion of fields (500 PhDs in the past 5 years) and programs (at least 5 PhDs in 5 years from at least 25 programs). It recommends coverage of the biomedical science PhD programs that exist in medical schools as well as doctoral programs offered in schools of agriculture. It also adds more life sciences categories than were in the 1995 Study, and two rapidly growing fields: communication and American studies. The Report tries to address interdisciplinarity through the listing of emerging fields, which are separate programs in some universities and included in existing programs in others. Examples are gender studies in the humanities, and biotechnology in the life sciences. The Report also recommends asking faculty who teach in more than one program to name all the programs that they teach or supervise theses in. In the pilot trials, almost 40% of the faculty were involved in more than one program.

Inadequate validation of data. *All data that are collected should be validated by the providers.* In the 1995 Study, lists of faculty were not sent back to the universities to be double-checked. The Report recommends that more care be taken in this area.

THE NRC ASSESSMENT OF RESEARCH-DOCTORATE PROGRAMS AS A RESOURCE FOR HIGHER EDUCATION AND ITS PUBLIC

The methodology described in this report will produce not only an assessment of the reputational quality of doctoral programs, it will produce a valuable data resource for students and for those

who plan, administer, and formulate policy for graduate education in the United States. The conduct of the recommended study will be a major undertaking, requiring time, care, and effort from universities as well as from the NRC. The Committee concluded that doctorate education is important enough to the American future that the collection of data to inform improvement is worthwhile.

GUIDE TO INFORMATION SOURCES

The earlier NRC assessments of research-doctorate programs referred to are:

Jones, L. V., G. Lindzey, and P. E. Coggeshall. 1982. *An Assessment of Research-Doctorate Programs in the United States*. Washington, D.C.: National Academy Press.

Goldberger, M. L., B. A. Maher, and P. E. Flattau. 1995. *Research-Doctorate Programs in the United States: Continuity and Change*. Washington, D.C.: National Academy Press.

A Website containing questionnaires, the taxonomy, and other information about the study can be found at: www.nationalacademies.org/resdoc



For More Information...

Copies of *Assessing Research-Doctorate Programs: A Methodology Study* are available for sale from the National Academy Press; call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area), or visit the NAP home page at www.nap.edu. The full text of this report is available at <http://www.nap.edu/catalog/10859.html>.

This study was supported by the National Institutes of Health Award# N01-OD-4-2139, TO #107; the National Science Foundation Award# DGE-0125255; the Alfred P. Sloan Foundation Grant No. 2001-6-10, and the United States Department of Agriculture Award# 43-3AEM-1-80054 (USDA-4454). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the National Research Council's Committee to Examine the Methodology for the Assessment of Research-Doctorate Programs and are not necessarily those of the funding agencies.

The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Copyright ©2003 by the National Academy of Sciences. All rights reserved.

Permission is granted to reproduce this document in its entirety, with no additions or alterations.



COMMITTEE TO EXAMINE THE METHODOLOGY FOR THE ASSESSMENT OF RESEARCH-DOCTORATE PROGRAMS

Jeremiah P. Ostriker (*chair*), Professor of Astrophysical Sciences and Provost *emeritus*, Princeton University, and Plumian Professor of Astronomy and Experimental Philosophy, University of Cambridge; **Elton D. Aberle**, Dean of the College of Agriculture and Life Sciences, University of Wisconsin-Madison; **John Brauman**, J.G. Jackson-C.J. Jackson Professor of Chemistry and Cognizant Dean for the Natural Sciences, Stanford University; **George Bugliarello**, President *emeritus* and University Professor, Polytechnic University; **Walter Cohen**, Vice-Provost and Professor of Comparative Literature, Cornell University; **Jonathan Cole**, John Mitchell Mason Professor of the University and Provost, Columbia University; **Ronald Graham**, Irwin and Joan Jacobs Professor, Department of Computer Science and Engineering, University of California-San Diego; **Paul Holland**, Frederic M. Lord Chair in Measurement and Statistics and Acting Director, Center for Statistical Theory and Practice, Educational Testing Service; **Earl Lewis**, Dean of the Graduate School and Vice-Provost for Academic Affairs-Graduate Studies, and Elsa Barkley Brown and Robin D. G. Kelley Collegiate Professor of History and African-American Studies, University of Michigan; **Joan F. Lorden**, Provost and Vice Chancellor for Academic Affairs, University of North Carolina at Charlotte; **Louis Maheu**, Dean and Vice President of Graduate Studies, Université de Montréal; **Lawrence Martin**, Dean of the Graduate School and Associate Provost for Analysis and Planning, Study Brook University; **Maresi Nerad**, Director of the National Center for Innovation and Research in Graduate Education, Associate Dean of the Graduate School, and Research Associate Professor for Educational Leadership and Policy Studies in the College of Education, University of Washington; **Frank Solomon**, Professor of Biology, Massachusetts Institute of Technology; **Catherine R. Stimpson**, Dean of the Graduate School of Arts and Science and University Professor, New York University.

Board on Higher Education and Work Force Liaison:

John D. Wiley, Chancellor, University of Wisconsin-Madison

Panel Members (in addition to the Committee)

Richard Attiyeh, Vice Chancellor for Research and Dean of Graduate Studies and Professor of Economics, University of California-San Diego; **Adam Fagen**, Postdoctoral Research Associate, Division of Earth and Life Sciences, The National Academies; **Robert F. Jones**, Vice President for Institutional and Faculty Studies, Association of American Medical Colleges; **Marsha Kelman**, Associate Vice President and Director of the Office of Institutional Studies, University of Texas at Austin; **George Kuh**, Chancellor's Professor of Higher Education, Indiana University Bloomington; **Leonard K. Peters**, Vice Provost for Research, Virginia Polytechnic Institute and State University; **Charles E. Phelps**, Provost and Professor of Political Science and Economics, University of Rochester; **Brenda Russell**, Professor of Physiology and Biophysics, Bioengineering and Medicine and Associate Vice Chancellor for Research, University of Illinois at Chicago; **Susanna Ryan**, Preceptor, Expository Writing Program, Harvard University; **Donald Rubin**, Chair and Professor of Statistics, Harvard University; **David Schmidly**, President, Texas Tech University; **Peter Syverson**, Vice President for Research and Information Services, Council of Graduate Schools.