

OVERVIEW OF THE NRC ASSESSMENT OF RESEARCH-DOCTORATE PROGRAMS

The National Research Council, which conducted two prior assessments of research doctorate programs, has conducted a new study with a very different methodology. The NRC report, in the form of Excel files for each field, is scheduled to be released **by the end of 2009**. The actual ratings for each program will not be known in advance of the report's public release.

Forty-seven of Stanford's doctoral programs were rated. The ratings are the result of a lengthy and complex data collection and statistical analysis process. Each program will receive two scores that demarcate a range of rankings (e.g., 14 - 27). This reflects the inherent variability among raters.

The data will be reported in Excel spreadsheets that alphabetically list the rated programs in a given field. Each program will receive four ratings: an Overall measure, Research Activity, Student Support and Outcomes, and Diversity of the Academic Environment. For each scale two scores, the "Third Quartile" (high end) and "First Quartile" (low end), will be reported. Each department will thus have a "range of rankings" for each of the four scales (the range scores are more fully explained on pp. 2-3).

The NRC indicates that the report for each field will look something like this, where university programs are listed alphabetically in the far left column:

University	Overall Measure		Research Activity		Student Support and Outcomes		Diversity of the Academic Environment	
	Third Quartile	First Quartile	Third Quartile	First Quartile	Third Quartile	First Quartile	Third Quartile	First Quartile
A	17	20	14	17	78	89	93	100
B	75	83	78	88	55	75	39	59
C	12	14	22	27	10	19	72	89
D	48	57	47	60	65	84	3	6
E	3	5	2	3	5	13	64	81

The remainder of this document describes in greater detail the data on which the ratings and rankings are based, and the procedures used to derive the ratings. The full methodology report (191 pages) can be downloaded from <http://sites.nationalacademies.org/pga/Resdoc/index.htm>

DATA COLLECTION

This is the third time the NRC has assessed the quality of research doctoral program in U.S. universities. The methodology of the current study is a marked departure from the earlier studies published in 1982 and 1995. The data collection was carried out in late fall 2006 and spring 2007 via surveys to universities, programs, faculty, and, in a few fields, students. Most of the data pertain to the academic year 2005-2006. The surveys collected data on a number of variables related to scholarly productivity of program faculty, effectiveness of doctoral education, research resources, demographic characteristics of students and faculty, resources available to doctoral students, and characteristics of the doctoral program.

In the current assessment NRC has defined 63 fields of study, whereas the prior assessment included 41 fields. Stanford will be fully rated for 47 programs in 40 fields (in some cases, like anthropology, Stanford had more than one program when data were collected); the full list is on pp. 5-6.

The NRC used four campus-based data collection instruments to derive the ratings. An *institutional questionnaire* collected institutional data, and a *program questionnaire* asked about programs and faculty participating in the programs. At Stanford these two surveys were completed by our Institutional Coordinator, the department of Institutional Research and Decision Support, and school coordinators (see list on p. 8) in close cooperation with department chairs. Faculty members were surveyed with a *faculty questionnaire*, and a subset were also surveyed in the *rating of program quality questionnaire* (what the NRC refers to as the “anchoring study”).

CALCULATING THE OVERALL PROGRAM RANKINGS

In the previous NRC assessments of doctoral programs, rankings were provided for doctoral programs based on reputational ratings, resulting in a rank-ordered list of programs within a field. The current assessment attempts to mitigate the methodological shortcomings of a reputational ranking in two ways. (The three “dimensional ratings” are described on p. 3.)

1) Quantitative Indicators with Field-Specific Weights. The new program ratings rely on a wider range of indicators of program quality: faculty research activity, student support and outcomes, and faculty and student demographics. The indicators come from the extensive data provided by the institutions themselves as well as some data collected directly by the NRC (e.g., faculty awards, publications, and citations). The 20 variables (see pp. 7-8) that make up the overall rating are weighted to produce quantitative estimates of program quality. The weights differ by field, recognizing that faculty members in different disciplines value different aspects of doctoral research programs. The values of the indicators multiplied by the weights provide the final rating.

2) Range of Rankings. Ratings from many raters were aggregated and arranged in order to yield a program ranking. However, the rankings will be reported in *ranges* instead of as a single number. For example, “Alpha University’s department of Beta is ranked between 27th and 35th.”

Developing the Weights

The field-specific weights are based on faculty opinions. In a spring 2007 survey, faculty were asked in two ways what they value most in PhD programs. The results of the two approaches were combined to produce overall weights, and then an overall rating.

1) Explicit weights for indicator variables. The survey asked faculty to rate the importance of 21 quantitative variables (including faculty quality, student characteristics and program characteristics) that determine overall program quality. This questionnaire was administered across all fields to all faculty members surveyed.

2) Implicit weights using perception survey. In the second step, the “implicit” or “anchoring study,” faculty were asked to rate a sample of programs in their field. This is essentially the process used in previous studies. Regression analysis was then used to determine which quantitative variables most closely predicted the program ratings in each field.

Ranking the Programs

Of course, different raters value different things; that is, a different set of respondent faculty would have differently weighted the variables. The NRC study attempts to compensate for this uncertainty by using a “random halves” procedure in which weights are calculated based on the responses of a randomly selected subset of faculty respondents. This is done 500 times, calculating 500 ratings that are rank ordered each time. The 500 resulting rankings are ordered from best to worst, and the bottom

quarter (the lowest 125 rankings, aka the First Quartile) and top quarter (the highest 125 rankings, aka the Fourth Quartile) are dropped. This results in two scores for each program – the so-called “inter-quartile rating for the program.” One end of the range is the highest remaining score (the top of the third quartile) and the other is the lowest remaining score (the top of the first or lowest quartile). In some cases this range will be large (third quartile=45, first quartile=56), in others, it will be quite small (third quartile=3, first quartile=5).

This method takes into account variability in rater assessment of what contributes to program quality within a field, variability in values of the measures for a particular program, and the range of error in the statistical estimation. These statistical techniques give a range of rankings, not an exact ranking for any program.

As a practical matter, it appears that the largest coefficients, and thus the most highly weighted variables, in most fields, are those relating to faculty research activity, although in any given field, other factors will come into play.

Three Additional Assessment Scales

In addition to the overall rating described above, three additional rating scales, designed to highlight specific dimensions of program quality, will be provided. Each scale is comprised of a few of the variables used in the overall rating, as detailed in the table on pp. 7-8.

With the exception of the research activity measure, these “dimensional ratings” produce program rankings that are quite different from the overall ratings. This reflects the premise that excellence in doctoral programs is not uni-dimensional. Just as with the overall rating, faculty survey data were used to determine how the variables included in each dimension were weighted in each field.

1) Research activity. This dimensional measure relates to various ways to gauge the contribution of research by counting faculty publications, citations, research grants, and honors and awards. NRC reports that the weight placed on publications per faculty member is remarkably consistent (about 30%) across fields. Research activity is the dimensional measure that most closely tracks NRC’s overall measure of program quality.

2) Student support and outcomes. This measure combines data on the percent of students fully funded in the first year, completion rates, time to degree, and some employment-related information. NRC found that faculty typically placed a larger weight on student support and completion rates than on the other components, and that there is “surprising uniformity” across broad fields on the weights.

3) Diversity of the academic environment. The measures that are included in this dimensional measure are: the percentage of faculty and students who are female and who are from underrepresented domestic minority groups, and the percentage of students who are international. Most fields place the highest weight on the percentage of students from underrepresented minority groups. In the health sciences, social sciences, and humanities, relatively high weights are also placed on the percentage of faculty who are underrepresented minorities. The percentage of international students was not highly weighted, except for in the physical sciences.

TIMELINE FOR THE RELEASE AND RESPONDING TO THE MEDIA

It is not yet known when the data and the report will be released; although we anticipate it will be in the Fall of 2009. Stanford will receive at least one day's advance notice of the release. We will immediately notify campus leaders, department chairs, and program directors.

Each department chair will receive a one-page summary of the Stanford-provided data that are used in the ratings (listed on pp. 7-8) in advance of the release.

You may receive calls from the media. The resources below can provide assistance in developing a response. The University's public position on the NRC Assessment is:

"Stanford believes that the NRC's assessment of research doctoral programs offers useful information on programs in these fields, and the University collected and contributed data about its programs.

Stanford does not comment on the specific rating or ranking of the individual departments and programs included in the National Research Council's report. Questions about specific departments and programs should be directed to the department chairs and school deans.

Stanford takes great pride in the quality of our many graduate programs, which have long track records of innovation and excellence across a broad range of fields. Stanford is particularly known for promoting interdisciplinary research and education that has yielded generations of outstanding leaders in academia, industry and government.

Many aspects of a high quality education cannot be reduced to numerical measures. The University does not use such ratings and rankings to shape its curriculum or academic policies. Every student's assessment of the best place to pursue graduate studies should be based on his or her own analysis of what the program has to offer. The decision of where to enroll should not be based on a rating or ranking from any organization."

RESOURCES

Every department will receive a data summary sheet listing the Stanford-supplied data for their department. Once the NRC report is released, the Office of the Vice Provost for Graduate Education (VPGE) and Office of Institutional Research will offer a workshop for faculty members interested in understanding the NRC data and how they may use it to understand aspects of their program, including information on how to compare some features of programs across universities.

Responding to media inquiries:	Lisa Lapin, Assistant Vice President for University Communications. lapin@stanford.edu (650) 725-8396
Details on Stanford-supplied data and NRC methodology:	Rana Glasgal, Associate Vice Provost for Institutional Research and Decision Support. rana@stanford.edu (650) 725-1327
Using the data in Stanford programs:	Patricia J. Gumport, Vice Provost for Graduate Education. gumport@stanford.edu (650) 736-0775

PROGRAMS RATED IN THE CURRENT ASSESSMENT

Stanford Program to be Rated ¹	NRC Field	# Programs Rated
Aeronautics and Astronautics	Aerospace Engineering	31
Anthropological Sciences	Anthropology	82
Art and Art History	History of Art, Architecture and Archaeology	58
Biochemistry	Biochemistry, Biophysics, and Structural Biology	159
Biological Sciences	Integrated Biology	120
Biomedical Informatics ²	Emerging Fields: Bioinformatics ²	17
Biophysics	Biochemistry, Biophysics, and Structural Biology	159
Cancer Biology	Cell and Developmental Biology	122
Chemical and Systems Biology	Pharmacology, Toxicology and Environmental Health	116
Chemical Engineering	Chemical Engineering	106
Chemistry	Chemistry	178
Civil and Environmental Engineering	Civil and Environmental Engineering	130
Classics	Classics	31
Communication	Communications	83
Comparative Literature	Comparative Literature	46
Computational Mathematics and Engineering ²	Emerging Fields: Computational Engineering ²	4
Computer Science	Computer Sciences	126
Cultural and Social Anthropology	Anthropology	82
Developmental Biology	Cell and Developmental Biology	122
Drama	Theatre and Performance Studies	27
Economic Analysis and Policy	Economics	117
Economics	Economics	117
Electrical Engineering	Electrical and Computer Engineering	136
Energy Resources Engineering	Chemical Engineering	106
English	English Language and Literature	119
French	French and Francophone Language and Literature	43
Genetics	Genetics and Genomics	65
Geological and Environmental Sciences	Earth Sciences	140
Geophysics	Earth Sciences	140
German Studies	German Language and Literature	29

¹ More than one Stanford program may be rated in the same NRC field.

² Programs in “Emerging Fields” and the “Languages, Societies and Cultures” categories will have data but will not be rated or ranked.

Stanford Program to be Rated ¹	NRC Field	# Programs Rated
History	History	137
Immunology	Immunology and Infectious Disease	78
Japanese ²	Language, Societies, and Culture ²	94
Linguistics	Linguistics	52
Management Science and Engineering	Operations Research, Systems Engineering and Industrial Engineering	72
Materials Science and Engineering	Materials Science and Engineering	83
Mathematics	Mathematics	127
Mechanical Engineering	Mechanical Engineering	127
Microbiology and Immunology	Microbiology	74
Modern Thought and Literature ²	Language, Societies, and Culture ²	94
Molecular and Cellular Physiology	Biochemistry, Biophysics, and Structural Biology	159
Music	Music	63
Neurosciences	Neuroscience and Neurobiology	94
Philosophy	Philosophy	90
Physics	Physics	160
Political Science	Political Science	105
Psychology	Psychology	236
Religious Studies	Religion	40
Slavic Languages and Literatures ²	Language, Societies, and Culture ²	94
Sociology	Sociology	118
Spanish	Spanish and Portuguese Language and Literature	60
Statistics	Statistics and Probability	61

¹More than one Stanford program may be rated in the same NRC field.

²Programs in “Emerging Fields” and the “Languages, Societies and Cultures” categories will have data but will not be rated or ranked.

VARIABLES INCLUDED IN THE OVERALL RATING FOR EACH PROGRAM

The Overall Rating for each program is based on a regression equation combining the 20 variables listed here. (How the variables are weighted and the ranges of rankings derived are described above.) In addition (as described on p. 3), there will be quartile scores on three dimensional scales: Research Activity (based on 4 variables), Student Support & Outcomes (based on 5 variables ¹), and Diversity of Academic Environment (based on 5 variables).

Several faculty measures are based on the number of “Allocated Faculty” for that program. Every program has “Core” faculty members (generally Academic Council members with a primary, secondary, or joint appointment in the department), “New” faculty members (like Core faculty members, but with an appointment beginning between 2003-2006), and “Associated” faculty members who are affiliated with the program through a Courtesy, Acting or other similar appointment, or dissertation advising. Assignment of Core, New and Associated faculty was done by Stanford’s institutional coordinator, in consultation with school coordinators, department chairs, and program directors. NRC then used data about dissertation committee supervision and membership to allocate faculty members on a proportional basis to all departments with which they were affiliated.

A one-page summary of all of the Stanford-provided data for each rated department or program will be provided to all department chairs and program directors before the NRC report release.

Variable	Dimensional scale in which variable is also included	Source of data
Publications per <i>allocated</i> faculty member (Average # of articles from 2000-06; Humanities: Average # articles plus # books x 5 from 1986-2006.)	Research Activity	Institute for Scientific Information (ISI). Humanities: Submitted faculty resumes.
Citations per publication (not included for humanities fields)	Research Activity	ISI
Percent of faculty with grants	Research Activity	Faculty survey
Percent faculty interdisciplinary (% who are Associated)		Provided by Stanford
Percent non-Asian minority Core and New faculty	Diversity of Academic Environment	Provided by Stanford
Percent female Core and New faculty	Diversity of Academic Environment	Provided by Stanford
Awards per <i>allocated</i> faculty	Research Activity	From scholarly societies
Average GRE-Q or GRE-V (humanities), 2004-2006		Provided by Stanford
Percent 1st yr. students w/ full support	Student Support & Outcomes ¹	Provided by Stanford
Percent 1st yr students with external funding		Provided by Stanford
Percent non-Asian minority students	Diversity of Academic Environment	Provided by Stanford
Percent female students	Diversity of Academic Environment	Provided by Stanford

Percent international students	Diversity of Academic Environment	Provided by Stanford
Average # of PhD's granted, 2002 -2006		Provided by Stanford
Percent completing within 6 years (8 yrs in humanities)	Student Support & Outcomes	Provided by Stanford
Median time to degree for full- and part-time students	Student Support & Outcomes	Provided by Stanford
Percent students with definite plans for academic positions (including postdocs) upon graduation	Student Support & Outcomes	NSF's Doctoral Records File
Student work space provided (yes or no)		Provided by Stanford
Health insurance provided (yes or no)		Provided by Stanford
Number of student activities (of a list of 18) offered by the institution or program ²		Provided by Stanford

¹ The Student Support & Outcomes dimension also includes a variable not included in the Overall Rating: Whether the department collects student outcome/placement data.

² The 18 activities include many that are offered by Stanford to doctoral students in all fields, such as orientation for new graduate students and language screening/support prior to teaching. Others varied by program, such as an annual review of all enrolled doctoral students, financial support for the graduate student association, or travel support to attend professional meetings.

STANFORD SCHOOL COORDINATORS

The NRC data collection effort at Stanford involved many people who devoted hours to ensure the accuracy of the data. The University offers them thanks.

Business:	Glenn Carroll, Ken Singleton, Rob Urstein
Earth Sciences:	Amy Balsom, Gordon Brown, Jerry Harris
Education:	Rania Sanford
Engineering:	Sally Gressens, Brad Osgood
Humanities & Sciences:	Ellie Fischbacher Maldonato, Iain Johnstone, Susan Stephens
Law:	Lawrence Friedman
Medicine:	Jim Rollins, Julia Tussing, John Pringle, William Talbot
SLAC:	Helen Quinn, Amy Rutherford
Dean of Research:	Ann George
Institutional Coordinator:	Gail Mahood (2006 - June 2007); Chris Golde (June 2007 - present)