

# NRC Assessment of Doctoral Programs: Ratings and Rankings

## Status Report

# Background

- NRC studies in 1982 and 1995: ratings & rankings
- 1999—Changes Recommended
  - Include Agricultural Sciences
  - Include Bio-medical fields in Medical Schools
  - NRC was asked to review the methodology
- 2002—Committee formed to examine methodology (JPO chair); two year study
  - Virtues of previous studies: authoritative, objective, widely used
  - Defects to address: spurious accuracy, poor dissemination, errors and lack of currency

# Methodology Study: Conclusions

- More rigorous treatment of uncertainties was necessary
- More emphasis needed on quantitative measures; broader coverage of fields; a benchmarking study needed to minimize errors.
- Goals: greater transparency and usefulness to students
- Dissemination of results should be in a web based format, designed for easy utilization by users and analytic uses of data.
- On-going updates of quantitative variables should continue after the study was completed.
- An assessment was worth doing.

# Changes in the Taxonomy

- Added Agricultural Sciences
- Added detail and updated the Biological Sciences
- Addition of PhD programs in medical schools
- Addition of research programs in some professional schools
- Addition of other fields that had grown in size

# Changes in the Quantitative Measures

## More about students

- Completion rates and attrition
- Outcomes tracking for sample of graduates
- Student support
- A trial student questionnaire for 5 fields (Neuroscience, chemical engineering, physics, economics, English)

# Changes to Encourage Use of the Study Data

- Make data easily available via web
- Disseminate through professional societies
- Permit customized comparisons by users
- Provide analytical tools and associated essays

# New Committee Charge (from NRC)

An assessment of the quality and characteristics of research-doctorate programs in the United States will be conducted. The study will consist of 1) the collection of quantitative data through questionnaires administered to institutions, programs, faculty, and admitted to candidacy students (in selected fields), 2) collection of program data on publications, citations, and dissertation keywords, and 3) the design and construction of program ratings using the collected data including quantitatively based estimates of program quality. These data will be released through a web-based, periodically updatable database and accompanied by an analytic summary report. Following this release, further analyses will be conducted by the committee and other researchers and discussed at a conference focusing on doctoral education in the United States. The methodology for the study will be a refinement of that described by the Committee to Examine the Methodology for the Assessment of Research-Doctorate Programs, which recommended that a new assessment be conducted.

# Committee\*

- **Jeremiah P. Ostriker (NAS)**  
*Committee Chair*  
(Astrophysics) Princeton University
- **John I. Brauman (NAS)**  
(Chemistry)  
Stanford University
- **Jonathan Cole**  
(Sociology)  
Columbia University
- **Virginia S. Hinshaw, *Vice-Chair\****  
(Biology)  
University of Hawaii, Mano'a
- **Eric W. Kaler**  
(Engineering)  
University of Delaware
- **Elton D. Aberle**  
(Agriculture)  
University of Wisconsin-Madison
- **Earl Lewis \***  
(history)  
Emory University
- **Norman Bradburn**  
(Statistics)  
University of Chicago

\*Several members of Committee are present or former Deans of graduate schools.

## Committee (2)

- **Joan F. Lorden\***  
(Neuroscience)  
University of North Carolina,  
Charlotte
- **Carol B. Lynch\***  
(Biology)  
University of Colorado,  
Boulder
- **Robert M. Nerem (NAE)**  
(Bioengineering)  
Georgia Institute of Technology
- **Suzanne Ortega\***  
(Sociology)  
University of Washington
- **Robert J. Spinrad (NAE)**  
(Computer Science)  
Xerox Corporation
- **Catharine R. Stimpson\***  
(English)  
New York University
- **Richard P. Wheeler\***  
(English)  
University of Illinois

# Kinds of Data

- Unit of Measurement: the doctoral program
- Data about Institutional Resources and Practices
- Data about Program Faculty, Resources and Practices
- For Selected Fields: Data about Accomplishments and Perceptions of Advanced Students

# Where will the data come from?

- Standardized source providers (eg citations)
- New surveys (eg faculty & students)
- *US institutions of higher education:*
  - *A HUGE TASK*

*Thank You*

# Rating and Ranking

## Background

- Field by field
- Task is rating “overall program quality” not “scholarly quality of faculty” as in the past
- Need informed raters
- Need to convey uncertainty in ratings: range and overlap of ratings will be emphasized.
- Need to construct ratings from quantitative variables

# Ratings: What measures “Quality of PhD Program”

- Usual Approach:

Those who design the study construct measures on an ad hoc basis (eg USN&WR etc).

- NRC Approach:

Faculty input on a field by field basis determines the measures. Two estimators of faculty values to estimate best measures: explicit and implicit.

# Ratings: Approaches

- Explicit:

Ask a sample of faculty in each field the relative importance of a different factors in determining overall program quality to determine weights.

- Implicit:

Regress quantitative variables on program ratings for a sample of programs. Use resulting weights to construct ratings for all programs in the field. Same principles as in Methodology Study.

# Ratings: Corollary Approaches

- Additional subsidiary indices will also be presented
  - I) "Research Impact" (RA), Possible metrics: citations/faculty member, publications/faculty member, honors and awards etc.
  - II) "Student Support and Outcomes" (RB) Possible metrics: fraction of students with full support, time to degree, attrition rate, fraction with a position in a relevant field on graduation, etc.
  - III) "Diversity of Academic Environment" (RC). Possible metrics: fractions of students and faculty that are female and minority etc.
- Customized indices will be offered

Users will be given capability to construct their own ratings of programs of interest using their own choice of variables and weights.

# Details on Ratings Methodology

- . The methodology for the study will be a refinement of that described by the Committee to Examine the Methodology for the Assessment of Research-Doctorate Programs, which recommended that a new assessment be conducted.

A major issue is to define the technical procedures we shall utilize to conduct task. This necessary task is a difficult one: we must specify a set of weights which, when multiplied by the collected data will produce “quantitatively based estimates of program quality”. There are two primary approaches to this task that have been suggested and they depend on how the rating task is specified; variants of each approach have been suggested that depend on who is asked to perform the ratings.

## Explicit Analysis

1. In this approach, raters are asked to assign weights to the quantitative variables. It is unmanageable to assign weights to every variable, so key variables will need to be identified, either empirically (through factor analysis) or judgmentally. Once the weights are determined, the program rating (and range of ratings) is applied to the program-specific value of each variable. The range of ratings is determined using the variability across the raters of the weight assigned to each variable.

2) Faculty are asked to indicate characteristics of programs that they find important. Then, asked to rate importance of each class of characteristics.

- Program Faculty Quality
- Student Characteristics
- Program Characteristics

Result of weights that can be applied to program data to yield a range of ratings.

## **Implicit Analysis** (see Appendix G of the NRC study “*Assessing Research-Doctorate Programs*” )

For each field

- Draw a sample of faculty who have expressed willingness to be raters
- Provide each rater with data from about 15-20 programs—data sheets will be mailed
- Ask each rater the degree of familiarity with each program
- Ask the rater to rate the program on a 5 point scale (input will be electronic and password protected)

# Sample of programs and faculty in each field

## Criteria

### Program

- Faculty size (quintiles depending on field size)
- Geographic Location (NE, SE, MW, SW, W)
- At least 30 programs sampled in each field

### Faculty in a field

- Geographic location
- Faculty size
- Rank
- At least 100 faculty sampled in each field

## Rater Information Sheet

**Institution:** {name}

**Location:** {place}

**Program:** {name}

**Program URL:** {URL}

<b>Faculty Names (Faculty spreadsheet)</b>	<b>Number of Ph.D.s 2002-2006 (Ques. PC1)</b>	<b>% PhDs in non-academic positions (average 2002-2005) (NSF data)</b>	<b>6 yr. completion rate (average 2001-2006) (Ques. PC16a and PC16b)</b>	<b>Median Time to Degree (average 2004-2006) (Ques. PC2)</b>	<b>Faculty % Female (Ques. PB5)</b>	<b>Faculty % Non-white (Ques. PB7)</b>
--	---	--	--	--	-------------------------------------	--

# What does the sheet do?

- Faculty names convey scholarly quality
- Other data convey important (and generally unknown) characteristics of the program
- Using the URL raters may, if they wish, find out more information about the program

# Final Analysis

1. Explicit and implicit ratings will be combined to provide estimates of range of rating for each program, and from these the ranges of rankings will be derived for overall program quality.
2. Three additional rankings and ratings will be constructed for the following: Research Impact; Student Support and Outcomes; Diversity of Academic Environment.
3. User designed Customized Indices will be offered.
4. Commissioned interpretive essays will be included.

# Further Information

<http://www7.nationalacademies.org/resdoc/index.html>