

A DATA BASED
ASSESSMENT OF
RESEARCH DOCTORATE
PROGRAMS

A Short History

- Data collection in 2007—using 2005-6 data for the most part
- 2008—reviewing and validating data and the methodology
- 2009—A Guide to the Methodology is published.
- 2010—The final report and accompanying spreadsheets will appear.

This Talk

- A little background
- The data that will become available and how they can be used.
- Thinking about quality of doctoral programs and program characteristics: How do we assess quality using data from programs?
- A variety of measures—overall and in different dimensions
- Study release—and after

CRA and the Study

- Spoke with the CRA Executive Board in February
- Concern that NRC definition of publications did not include all papers from refereed conferences.
- NRC agreed to record data on these pubs from résumés.
- Will not have data on citations.
- Hope to release CS rating ranges at the same time as those for other fields
- That said, let me describe what the NRC is doing.

Methodology Study: Conclusions

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

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, and citations, 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
- **Virginia S. Hinshaw, *Vice-Chair****
(Biology)
University of Hawai'i Mano'a
- **Paul Holland**
(Statistics)
Educational Testing Service
- **Elton D. Aberle**
(Agriculture)
University of Wisconsin-Madison
- **Eric W. Kaler (NAE)**
(Engineering)
SUNY-Stony Brook
- **Norman Bradburn**
(Statistics)
University of Chicago
- **Earl Lewis ***
(history)
Emory University

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

Committee (2)

- **Earl Lewis ***
(history)
Emory University
- **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 New Mexico
- **Catharine R. Stimpson***
(English)
New York University
- **Richard P. Wheeler***
(English)
University of Illinois

What is the Assessment?

- Collection and dissemination of data on important aspects of doctoral programs
 - Programs
 - Students
 - Faculty
- Development of a benchmarking/rating methodologies
 - Compare doctoral programs in a single field across universities
- 220 Universities, 59 fields with ratings

What will be released?

- The Report
 - A “slim volume” discussing what was done in the study, the data, and two *illustrative* methodologies for data based rankings.
- Online spreadsheets
 - Data for 4838 programs for the 20 variables used in the ratings calculation, and for 9 additional variables.
 - Range of rankings for 5 types of illustrative rankings: 2 overall and 3 dimensional
 - Ability to “click through” to get detail of ranking calculations

Release materials (cont'd)

- Demos to show how to query the spreadsheets
- Related effort on [PhDs.com](https://www.phds.com) will permit calculations with user determined weights
- Press release and FAQ's
- Press conference
- Revised Methodology Guide

Later

- Public use database
- Release of all questionnaire data (with individual identities masked) to researchers who request it and sign a confidentiality agreement

Six Months Later

- Conference on analytic uses of the data

Key points

- It is helpful to compare programs that are doing similar things by collecting the same data from all the programs in a field.
- It is possible to compare data values by forming a ranking.
- There are many ways to develop rankings—the NRC did it in two ways—there are many others.
- It is important to know what goes into a ranking.

Some things that will change from the July 2009 Methodology Guide

- The rankings and their ranges
 - 1 overall ranking range \Rightarrow 2 separate ranking range calculations as illustrations of data-based ranking schemes
 - Change in length of range from covering 50% of a program's rankings to covering 90%
- Emphasis
 - Ranges of rankings are *illustrative*. You could get different results with different assumptions.

Audiences

- Prospective graduate students. Give them better information about the various programs to make more informed decisions re where to apply.
- Faculty in the programs to better evaluate their own strengths and weaknesses.
- Those responsible for the health of graduate programs to enable them to better assess the programs under their charge and compare more objectively to those in other institutions.
- Those with more global interests (legislators, boards of trustees, the US government, other nations) to provide more transparency in assessing a vital US national institutional resource.
- During the “Recession of 2008-?” Resource allocation decisions.

Examples of Questions

- A prospective student
 - What do I want to do when I finish and does the program seem to support that aim?
 - Am I likely to get funding?
 - How long will it take to complete?
 - How likely is it that if I start in a program that I will complete in a reasonable amount of time?
 - Will I be the only (woman, minority)?

More Examples of Questions

- A department chair
 - What are the strengths and weaknesses of the program?
 - How does my program compare to peer programs?
- A provost
 - Where can additional resources result in the most improvement?
 - What programs could benefit from being combined with similar programs?

And Yet More Examples of Questions

- A state board of higher education
 - Do we have too many doctoral programs in a given field?
 - Which programs are strong nationally and deserve more support?
 - Which programs do we really need to have for the state's economic vitality?

What is the Assessment?

- Collection and dissemination of data on important aspects of doctoral programs
 - Programs
 - Students
 - Faculty
- Development of a benchmarking/rating methodology
 - Compare **doctoral programs** in a **single field** across universities
- 212 Universities, 59 fields with ratings

Where do the data come from?

- Standardized source providers (e.g. citations from ISI, NSF for post-graduate student plans)
- New surveys (e.g. faculty & students)
- For a *few* fields, publications compiled from résumés
- ***US institutions of higher education:
– A HUGE TASK***

Summary Descriptive Information for Each Program

Program Characteristics

- **Size**
- **Publications per faculty member**
- **Citations per publication (except for humanities fields and CS)**
- **Percent of faculty with grants**
- **Honors and awards per faculty member**

Student Support and Outcomes

- **Number of PhDs**
- **Percent receiving financial support in first year**
- **Median time to degree**
- **Percent of entering cohort(s) completing within six years (eight for the humanities)**
- **Percent of graduates with definite employment or postdoc plans (from NSF)**

Summary Descriptive Information for Each Program

Program Diversity

- **Faculty:**
 - Gender diversity
 - Racial/ethnic diversity
- **Students**
 - Gender diversity
 - Racial/ethnic diversity
 - International diversity

Program Interdisciplinarity

- **Percent of faculty associated with other programs**
- **Identification of “umbrella” programs**

How will the ratings/rankings work?

Two Approaches

- Asked faculty what they thought was important to the quality of a doctoral program and developed weights (S-weights).
- Asked a sample of faculty in each field how they would rate a sample of programs. Related those ratings to 20 program characteristics through a regression (R-weights)
- Calculated ratings using each approach for all programs in a field, based on program values for the 20 characteristics.
- The rankings will be *illustrative*.

Overall Rating AND Dimensional Measures

- Student Treatment and Outcomes
- Diversity of the Academic Environment
- Research Activity of Program Faculty

The Twenty Key Variables used in the Rankings

- Publications per allocated faculty
- Citations (exc. Humanities and CS) per publication
- Percent faculty with grants
- Awards per faculty
- Percent 1st Yr. Full Support
- Percent Completing in 6 yrs. or less (8 yrs. for humanities)
- Median Time to degree
- Students with Academic Plans
- Collects Outcomes data
- Percent Faculty Minority
- Percent Faculty Female
- Percent Students Minority
- Percent Students Female
- Percent Students International
- *Percent Interdisciplinary*
- *Average GRE-Q*
- *Number of PhDs 2002-2006*
- *Student Workspace*
- *Student Health Insurance*
- *Student Activities*

Ratings: What measures “Quality of PhD Program”?

- Usual Approaches:

Those who design the study construct measures on an ad hoc basis.

- Based on reputation
- Based on refinements of scholarly productivity measures

- NRC Approach:

Faculty input on a field by field basis determines the measures. Two estimators of faculty values to estimate best measures: direct (S) and regression-derived (R).

Sources of uncertainty for any rating

- Differences among raters
- Year-to-year variation in the data
- Range of error in any statistical estimation
- ➔ Every rating has a range, and so do the rankings derived from the ratings
- We settled on a broad range—one that covers 90% of the estimated rankings for a program
- Unincluded and unquantifiable factors may also matter—but the committee focussed on what could be quantified.

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 (later)

What will the spreadsheets provide?

- An online spreadsheet for each of the 62 fields with programs arranged alphabetically within a field
 - Rankings will be shown as a range.
 - Variables that go into the rankings will be shown first, then additional data
 - There will be the three supplemental measures along different dimensions in addition to the overall measure.
 - Can click through to rating calculation details
 - There will be data for individual variables.
 - Data on variables for unrated and emerging fields will be shown, where available.

Some Selected Preliminary Statistics for Programs in Computer Science

| | | | |
|---|-----|--------------------------|-----|
| Number of programs | 126 | Faculty | |
| Public | 89 | Average Total Faculty | 37 |
| Private | 37 | Average Percent Tenured | 63% |
| Average Enrollment | 89 | Faculty Diversity | |
| Programs with enrollment under 25 | 12 | Average percent female | 13% |
| Programs with enrollment over 50 | 83 | Non-response | 1 |
| Funding | | Average percent URM | 2% |
| Programs with full funding for >50% first year students | 110 | Non-response | 4 |
| Programs with full funding for 100% first year students | 55 | Student Diversity | |
| Completion and Time to Degree | | Average percent female | 21% |
| Programs with over half of their students completing in 6 years or less | 9 | Non-response | 0 |
| Programs with a median time to degree under 6 years | 83 | Average percent URM | 8% |
| | | Non-response | 3 |

Some Selected Preliminary Statistics for Computer Science

Research Activity

| | |
|-----------------------------|-----|
| Average Publications | TBC |
| Percent faculty with grants | 74% |

Problem with CS fields: Don't have papers from refereed conferences unless they appeared in conference proceedings for ISI-indexed journals. The Report will include these as publications.

Some Selected Preliminary Statistics for Programs in EECE

| | | | |
|---|------------|--------------------------|-----|
| Number of programs | 136 | Faculty | |
| Public | 98 | Average Total Faculty | 45 |
| Private | 38 | Average Percent Tenured | 66% |
| Average Enrollment | 98 | Faculty Diversity | |
| Programs with enrollment under 25 | 21 | Average percent female | 9% |
| Programs with enrollment over 50 | 77 | Non-response | 21 |
| Funding | | Average percent URM | 5% |
| Programs with full funding for >50% first year students | 119 | Non-response | 23 |
| Programs with full funding for 100% first year students | 45 | Student Diversity | |
| Completion and Time to Degree | | Average percent female | 18% |
| Programs with over half of their students completing in 6 years or less | 49 | Non-response | 20 |
| Programs with a median time to degree under 6 years | 114 | Average percent URM | 13% |
| | | Non-response | 22 |

Some Selected Preliminary Statistics for EECE

| Research Activity | |
|-----------------------------|-----|
| Average Publications | 1.3 |
| Average citations/pub | 1.4 |
| Percent faculty with grants | 78% |

Problem with CE fields: Don't have papers from refereed conferences unless they appeared in conference proceedings for ISI-indexed journals.

Some Selected Preliminary Statistics for Programs in Computer Engineering

| | | | |
|---|----|--------------------------|------|
| Number of programs | 20 | Faculty | |
| Public | 14 | Average Total Faculty | 30.3 |
| Private | 6 | Average Percent Tenured | 69% |
| Average Enrollment | 40 | Faculty Diversity | |
| Programs with enrollment under 25 | 9 | Average percent female | 10% |
| Programs with enrollment over 50 | 6 | Non-response | 1 |
| Funding | | Average percent URM | 8% |
| Programs with full funding for >50% first year students | 16 | Non-response | 2 |
| Programs with full funding for 100% first year students | 9 | Student Diversity | |
| Completion and Time to Degree | | Average percent female | 15% |
| Programs with over half of their students completing in 6 years or less | 8 | Non-response | 1 |
| Programs with a median time to degree under 6 years | 17 | Average percent URM | 8% |
| | | Non-response | 3 |

Some Selected Preliminary Statistics for Programs in Computer Engineering

| Research Activity | |
|-----------------------------|-----|
| Average Publications | 1 |
| Average citations/pub | 1.2 |
| Percent faculty with grants | 79% |

Problem with CE fields: Don't have papers from refereed conferences unless they appeared in conference proceedings for ISI-indexed journals.

- Tried to separate CE from EECE—but only 20 fields—couldn't do rankings
- Definition of a “publication”—we didn't have data on all refereed conference papers, although we did have data on papers in published proceedings

Some Examples of Comparisons

| Characteristic | Program A | Program B | Program C | Program D | Program E |
|---|-----------|-----------|-----------|-----------|-----------|
| Publications per Allocated Faculty | 4.993 | 4.328 | 4.448 | 2.937 | 2.379 |
| Cites per Publication | 3.573 | 3.401 | 2.782 | 2.819 | 2.386 |
| Percent Faculty with Grants | 88.6% | 100.0% | 95.5% | 90.5% | 73.4% |
| Percent Faculty Interdisciplinary | 71.4% | 0.0% | 38.1% | 18.8% | 0.0% |
| Percent Non-Asian Minority Faculty | 0.0% | 5.0% | 0.0% | 2.9% | 3.1% |
| Percent Female Faculty | 16.2% | 13.6% | 8.0% | 17.9% | 8.8% |
| Awards per Allocated Faculty | 1.929 | 7.291 | 1.896 | 0.640 | 0.424 |
| Average GRE-Q | 712 | 772 | 767 | 703 | 673 |
| Percent 1st yr. Students w/Full Support | 100.0% | 100.0% | 100.0% | 100.0% | 86.0% |
| Percent 1st yr Students with External Funding | 0.0% | 0.0% | 22.2% | 0.0% | 0.0% |
| Percent Non-Asian Minority Students | 2.8% | 1.9% | 3.2% | 8.0% | 13.6% |
| Percent Female Students | 39.3% | 39.1% | 39.8% | 42.2% | 37.3% |
| Percent International Students | 23.0% | 42.7% | 37.2% | 45.1% | 31.3% |
| Average PhDs 2002 to 2006 | 31.6 | 17.4 | 20.2 | 11.400 | 19.800 |
| Percent Completing within 6 years | 49.3% | 77.8% | 67.6% | 41.6% | 54.0% |
| Time to Degree Full and Part Time | 5.7 | 5 | 4.9 | 4.3 | 5.000 |
| Percent students in Academic Positions | 17.2% | 32.1% | 25.6% | 20.0% | 12.2% |
| Student Work Space | 1 | 1 | 1 | 1 | 1 |
| Health Insurance | 1 | 1 | 1 | 1 | 1 |

A Sample Comparison

R and S-based Rankings for 5 Programs in a Field

| Institution Name | R5 | R95 | S5 | S95 |
|-------------------------|-----------|------------|-----------|------------|
| Institution A | 4 | 17 | 10 | 29 |
| Institution B | 4 | 27 | 3 | 10 |
| Institution C | 13 | 37 | 8 | 23 |
| Institution D | 31 | 79 | 31 | 86 |
| Institution E | 52 | 102 | 91 | 150 |

Dimensional Rankings for the Same

| Institution Name | RA5 | RA95 | SS5 | SS95 | D5 | D95 |
|------------------|-----|------|-----|------|-----|-----|
| Institution A | 7 | 29 | 9 | 66 | 81 | 131 |
| Institution B | 3 | 12 | 31 | 110 | 97 | 147 |
| Institution C | 9 | 39 | 6 | 42 | 101 | 151 |
| Institution D | 21 | 85 | 21 | 93 | 42 | 97 |
| Institution E | 53 | 124 | 53 | 133 | 77 | 128 |

Coefficients for Chemistry Programs

| Characteristic | R5 | R95 | S5 | S95 |
|--|--------|--------|--------|--------|
| Publications per Allocated Faculty | -0.011 | 0.144 | 0.146 | 0.151 |
| Cites per Publication | 0.037 | 0.086 | 0.125 | 0.130 |
| Percent Faculty with Grants | 0.066 | 0.118 | 0.163 | 0.167 |
| Percent Faculty Interdisciplinary | -0.002 | 0.083 | 0.033 | 0.036 |
| Percent Non-Asian Minority | -0.027 | 0.049 | 0.007 | 0.009 |
| Percent Female Faculty | -0.061 | 0.011 | 0.011 | 0.013 |
| Awards per Allocated Faculty | 0.015 | 0.088 | 0.081 | 0.086 |
| Average GRE-Q | -0.011 | 0.062 | 0.066 | 0.070 |
| Percent 1st yr. Students w/ Full Support | 0.045 | 0.101 | 0.053 | 0.057 |
| Percent 1st yr Students with External Funding | -0.049 | 0.005 | 0.043 | 0.047 |
| Percent Non-Asian Minority Students | -0.062 | -0.007 | 0.015 | 0.017 |
| Percent Female Students | -0.023 | 0.037 | 0.016 | 0.018 |
| Percent International Students | -0.068 | -0.022 | 0.007 | 0.009 |
| Average PhDs 2002 to 2006 | 0.101 | 0.181 | 0.038 | 0.041 |
| Percent Completing within 6 Time to Degree Full and Part | -0.025 | 0.026 | 0.045 | 0.048 |
| Percent students in Academic Positions | -0.019 | 0.028 | -0.025 | -0.023 |
| Student Work Space | -0.026 | 0.055 | 0.067 | 0.069 |
| Health Insurance | 0.006 | 0.076 | 0.005 | 0.006 |
| Number of Student Activities Offered | 0.022 | 0.082 | 0.003 | 0.004 |
| | 0.062 | 0.117 | 0.022 | 0.024 |

Big Points

- Data-based ranking is not a simple task
- Rankings depend on the values of measures used and the weight that is put on them.
- The NRC is not endorsing any method as “best”
- The NRC study will be complex. We will try to make it useful—but that is also up to you.

More things to remember

- You will be able to access the values that went into the calculation of the 5th and 95th percentile values.
 - Note: the calculation uses standardized values. We will also show the actual values for the program and the standardized value in the rating calculation
- The dimensional rankings spotlight program characteristics not prominent in the overall rankings

Does it matter that it's “late?”

- There is a trade-off between speed and accuracy
 - We spent a lot of time trying to get universities to provide comparable data and developing our model.
- In most fields, and especially now, doctoral faculty change relatively slowly, as do patterns of publication
- We would like to update the study in the next two years
 - Now that we have developed the statistical machinery, it is a data updating task, which could be carried out on-line.
 - But, we will need to obtain funding

To Learn More About the Study

<http://sites.nationalacademies.org/pga/Resdoc/index.htm>

Or contact

ckuh@nas.edu

Or

jvoytuk@nas.edu