Measuring Research Data Uncertainty in the 2010 NRC Assessment of Geography Graduate Education

Ashton Shortridge a, Kirk Goldsberry a & Kathleen Weessies a

a Michigan State University, East Lansing, Michigan, USA

Available online: 25 Oct 2011


To link to this article: http://dx.doi.org/10.1080/00221341.2011.607510
ABSTRACT
This article characterizes and measures errors in the 2010 National Research Council (NRC) assessment of research-doctorate programs in geography. This article provides a conceptual model for data-based sources of uncertainty and reports on a quantitative assessment of NRC research data uncertainty for a particular geography doctoral program. Findings indicate that important variables, including faculty totals and allocations and publication counts, are substantially undercounted, with important and negative impacts on program research activity measures. Further, these research measures are highly sensitive to small changes in counts and are particularly problematic for interdisciplinary fields such as geography. We caution against using the 2010 NRC data or metrics for any assessment-oriented study of research productivity.

Key Words: assessment, geography departments, assessment processes

INTRODUCTION
Since its founding in 1916, the National Research Council (NRC) has engaged in projects serving government, the public, and the scientific and engineering communities. The ambitious goal of the latest project, led by the Committee to Assess Research-Doctorate Programs, was to measure, assess, and compare a range of characteristics for over five thousand doctorate programs across 212 universities (NRC 2010b). Twenty variables were collected about each program; while many were compiled during early 2007, they documented program activity from 2000–2006 with one metric casting back to 1981. Intensive data analysis produced program ranking distributions; these distributions, along with the variables on each program, were intended to be employed for assessment of graduate programs by university administrators, active program faculty, and prospective graduate students, among many others. Geography was one of the disciplines surveyed in this endeavor: a total of forty-nine Ph.D. programs submitted data and were evaluated and ranked in the 2010 report; this number is substantially lower than the sixty-seven Ph.D.-granting geography programs in the Association of American Geographers’ (AAG) 2005–2006 Guide (AAG 2006). Rankings in geography were determined via a weighted combination of composite scores encompassing three categories: research activity, student support and outcomes, and diversity. Given the assessment’s focus on doctorate programs, it is notable that no data on teaching activity, curriculum breadth or depth, or graduate student training were collected or used.

Research activity, a composite score derived from variables reflecting publications per faculty, average citations per publication, percent of faculty with grants, and awards per faculty, was highly influential in the overall rankings. On average this composite comprised about half of the weights for the geography rankings (NRC 2010a, Appendix E). It seems evident that any parameter so influential must be as accurate, robust, and informative as possible. To be accurate, the variables comprising research activity need to be correctly and consistently measured for each doctorate program. To be robust, the metrics must not change substantially when subjected to reasonable variation. To be informative, the model characterizing research activity should reflect the overall research priorities of geography as a scholarly discipline. Together, shortfalls in any of these elements lead to an increase in uncertainty in the resulting assessment. This article considers the role, magnitude, and implications of uncertainty in the NRC assessment of research activity in geography, and consists of four parts:

Part 1: A conceptual model of uncertainty for the NRC rankings. We developed a model focusing on data-based uncertainty. It is intended to clarify specific sources of uncertainty that affect metrics and rankings of doctorate programs in geography, and encompasses both inaccuracy and vagueness.

Part 2: Quality assessment of the NRC methodology research activity metrics. We conducted a case study on the research productivity of a single geography program. This study uses the same database and accounting rubrics as the NRC methodology to recalculate key research activity metrics. These metrics are compared to those in the 2010 report (NRC 2010c).

Part 3: Sensitivity assessment of the NRC methodology. Using data collected in Part 2 we evaluated the sensitivity of research
activity metrics to multiple realistic scenarios for measuring faculty and article quantities. If these metrics were robust, they would be relatively insensitive to different scenarios. We then considered conceptual issues of vagueness and their potential for impact.

**Part 4: Implications for doctorate programs in geography.** Given the results from Parts 2 and 3, we identified key issues specific to quantitative assessments of graduate programs in geography. Their implications were considered at multiple scales, from the departmental level to the disciplinary level.

**PART 1. SOURCES OF UNCERTAINTY IN THE NRC ASSESSMENT**

We appreciate the effort involved by the NRC and participating institutions and programs in undertaking this massive study of American doctorate programs; its completion alone is quite an achievement. However, in a study as sweeping as this, uncertainty is inevitable (NRC 2010a). Furthermore, depending on its magnitude, uncertainty could undermine the value of the entire assessment. Certainly many criticisms, both general and specific, have been made (e.g., Cole 2011; Carpenter-Hubin and McClelland 2011; Inside Higher Ed News 2010; Levy, 2010; Stigler 2010). In this section we identify multiple factors that could contribute to uncertainty in the 2010 NRC rankings.

Uncertainty in the final rankings propagates from two sources: (1) the inputs, that is, data-based uncertainty, and (2) the model formulation, or model-based uncertainty. This second source, much discussed in the academic program assessment literature (e.g., Sweitzer and Volkwein 2009; Dill 2006), is concerned with how to translate various measurable characteristics of academic programs into rankings of quality or reputation rankings. Model-based uncertainty has been partially accounted for by the NRC methodology itself, which explicitly calculated ratings intervals by repeatedly sampling the parameter distribution (NRC 2010b). However, data-based uncertainty was not directly addressed, and is the focus of this article. Thoughtful macro-level concerns about data quality have been raised (e.g., Cole 2011; Stigler 2010), but this appears to be among the first reports to quantify the nature and magnitude of this error for a particular program.

To characterize data-based uncertainty we adopt and extend a conceptual model from the geographic information science literature (Duckham et al. 2001). This model, presented in Table 1, identifies two distinct components: inaccuracy, which is the deviation of measured values from the actual values, and vagueness, which concerns imprecision in the definition of concepts. Inaccuracy is a function of measurement; the four research activities measured in the NRC methodology are subject to measurement error. Further, all activity metrics depend directly or indirectly on faculty counts, and upon the weights by which faculty activity are distributed to multiple programs. Inaccuracy in the measurement of these six variables comprises the left side of Table 1. On the right side we consider sources of vagueness for the NRC methodology. In this or any program assessment effort, the subject domain is largely comprised of ambiguous concepts like “program quality,” “publications,” and “faculty.” The definitions and boundaries of such concepts are not inherently crisp or determinate, and reasonable people may not agree on their nature. Imprecision due to vagueness includes the definitions of categories like “Publication,” “Faculty,” and “Award,” which ultimately must be defined in some manner if they are to be measured.

While the conceptual model is useful for understanding sources of uncertainty, evaluating overall impact of data-based uncertainty on the geography rankings is difficult as its magnitude is likely to vary from university to university, program to program, and even scholar to scholar. In this article we largely confine our analysis of this uncertainty to a single geography program, one we know best: Michigan State University (MSU Geography). Further, we focus on error and vagueness in the three most important components of the NRC research activity metric: faculty allocation, publication rates, and citation rates.

### Table 1. A conceptual model of categories and sources of data-based uncertainty.

<table>
<thead>
<tr>
<th>Inaccuracy</th>
<th>Vagueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Counts</td>
<td>Publication Definitions</td>
</tr>
<tr>
<td>Faculty Counts</td>
<td>Faculty Definitions</td>
</tr>
<tr>
<td>Citation Counts</td>
<td>Temporal Relevance</td>
</tr>
<tr>
<td>Faculty Allocations</td>
<td>Grant Impact</td>
</tr>
<tr>
<td>Grant Counts</td>
<td>Award Definition</td>
</tr>
<tr>
<td>Award Counts</td>
<td></td>
</tr>
</tbody>
</table>

**PART 2: INACCURACY**

Here we assess the accuracy of three fundamental research activity metrics in the NRC rankings. A meticulous independent data collection effort was made that enabled us to identify the nature and magnitude of inaccuracy of each metric for MSU Geography. First we briefly define the metrics (NRC 2010b) and then we consider each more fully in turn:

**Allocated faculty:** Sum of faculty contribution to a degree program. Faculty contribution is apportioned to different degree programs at an institution according to the proportion of graduate student committees that member has served on.
Publications per faculty: Average number of articles per allocated faculty from 2000–2006. Publications were obtained by matching faculty with publications in the Thompson-Reuters (formerly ISI) Web of Science/Knowledge database(s).


Inaccuracy in Counting and Allocating Faculty

The NRC methodology established a population of allocated faculty for each geography program in 2006. First, each faculty member was classified as core in their home program if they had served as a member of a doctoral dissertation committee or as new if they had been hired in the past three years. Also, core faculty members were deemed to be associated with any other degree program at their institution if they had served on dissertation committees in that program (NRC 2010b). Faculty members’ contributions were apportioned across doctorate programs according to the proportion of graduate student committees that they had served on; a faculty member who served only on committees in his or her core program would have an allocation rate of 100 percent to that program, while one who served on committees in other programs would be identified as an associate in those other programs, and a proportion of their research contributions would be allocated accordingly. In turn, this apportioning influenced the NRC metric, “Publications per faculty,” which is a weighted sum of articles by allocated faculty: for example, two articles by a professor with 80 percent allocation to a program netted that program 1.6 allocated articles, while the remaining 0.4 articles were allocated to other programs. Consequently, the influence of interdisciplinary doctoral committee participation on “research activity” was enormous:

- Only publications by allocated faculty counted towards the rankings

To assess the effect of allocation, the authors obtained the official NRC list of 2006 MSU Geography faculty with allocation percentages from the university’s graduate school office. There were thirty total faculty allocated to the program: twenty-three of these were core faculty, four were new faculty and only three were associated faculty. Allocation rates ranged from 100 percent to 10.5 percent. Total weighted allocated geography faculty published in the 2010 NRC program rankings was 24.89 (NRC 2010a). Table 2 documents the adjustments to total core/new faculty in the left column, and to associated faculty in the central column. The program effectively lost the contributions of 3.2 core faculty through allocation and gained the contributions of 1.6 associated faculty.

The reported NRC associated faculty numbers for MSU Geography are very low in comparison to internal records. During 2000–2006, eighteen MSU Geography Ph.D. students graduated; between them, these students had twenty-five external members, twenty-four of whom were MSU faculty. Only one of these external members was assigned to MSU Geography as an associated faculty; the twenty-three other faculty were not formally associated, and no part of their research activity was apportioned to MSU Geography. Furthermore, an additional seven geography doctoral students graduated in 2007; none of their external committee members appeared in the 2006 geography allocation lists either. Consequently the total NRC count of associated faculty (three) is inaccurate and very low; we conservatively estimate the actual value of associated faculty at twenty-five. The magnitude of error in the allocations is unknown but presumably quite large.

Inaccuracy in Article Counting

Along with faculty allocations, the number of research articles in the assessment period was another key metric in determining a department’s publication rate and overall research productivity. Identifying the entire set of research publications for dozens of faculty members over a seven-year period is obviously a formidable task with considerable potential for inaccuracy. The NRC did this by specifying a methodology and contracting the work of extracting and counting research articles from the Thompson-Reuters/ISI databases to a private firm; they...
Table 3. Metrics of four scenarios for MSU Geography allocation and activity. Changes in how articles are counted and faculty allocations result in strikingly different publication activity statistics. Rankings are relative to published 2010 NRC Geography publications per faculty rates and are intended solely to demonstrate the impact of alternative counting scenarios on research activity metrics. See text for scenario details. (Column 1 data drawn from (NRC 2010c), remaining columns compiled by the authors.)

<table>
<thead>
<tr>
<th>Baseline 2010 NRC</th>
<th>“Recount”</th>
<th>“Program Fortress”</th>
<th>“Hide The Children”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Publications</td>
<td>N/A</td>
<td>188</td>
<td>165</td>
</tr>
<tr>
<td>Allocated Publications</td>
<td>87.1</td>
<td>143.5</td>
<td>165.0</td>
</tr>
<tr>
<td>Allocated Faculty</td>
<td>24.89</td>
<td>24.89</td>
<td>25.00</td>
</tr>
<tr>
<td>Publications per Faculty</td>
<td>0.50</td>
<td>0.82</td>
<td>0.94</td>
</tr>
<tr>
<td>Pubs/Faculty Rank (of 49)</td>
<td>33RD</td>
<td>11IH</td>
<td>9IH</td>
</tr>
</tbody>
</table>

We conservatively chose to only count the total of 188 research articles rather than include other WOS publication categories. Weighted by allocation, this corresponded to 145.49 articles, or about 0.82 articles per year per allocated faculty (Table 3); according to this count the 2010 NRC publication rate for MSU Geography was 40 percent too low. This substantial discrepancy in measured research activity due to data inaccuracy in article counts underscores a critical point from the NRC methodology documentation itself:

Small differences in the variables can result in major differences in the range of rankings, especially when a program is very similar on other measures to other programs in its field. But individual instances of programs that should have been ranked considerably lower or significantly higher than the tables indicate may emerge, and so it is strongly recommended that the rankings of individual programs be treated with circumspection and caution and analyzed carefully. (NRC 2010a, 63)

In addition, since we have article counts by individual faculty, we are able to assess the net impact of faculty allocation to determine the effect of the allocation process on research activity metrics. Table 2 shows the changes in MSU Geography article counts due to allocation. Essentially the left portion of the table represents net contributions by geography program faculty to its own activity numbers, while the central portion represents net contributions by other faculty. It depicts the degree to which allocation changed the contribution of program and associated faculty. This impact is profound. MSU Geography faculty produced 165 articles during the reporting period. Nearly twenty-six (15.6%) of these articles were lost to the program due to allocation. In turn, the program gained only 4.2 articles from associated faculty, not nearly enough to offset the outflow from core/new faculty to other units. The net change of −13 percent had a substantial effect on critical research activity metrics and ultimately on the program’s rankings within the discipline. While we cannot measure the impact of lost productivity from uncounted allocated faculty, this may also be considerable. The decision to effectively penalize a program’s measured research activity for faculty service on external graduate committees is one of the more baffling aspects of the NRC assessment methodology.

Inaccuracy in Citation Rates

In the 2010 assessment release, the NRC identified 0.76 average annual citations per publication for MSU Geography (NRC 2010c). The authors conducted an independent citation count by searching the WOS for all publications by 2006 allocated faculty for the period 1981–2006. Article totals by faculty, by year, were assembled, following the official methodology (NRC 2010b). Only journal article publications were employed; book reviews, commentary, including special issue introductions, and conference abstracts published in journals were not included. Citations of these articles during each year from 2000 to 2006 were
toted for each faculty member, along with the number of articles published by that faculty member up to and including that year.

Articles and citation counts were multiplied by that faculty member’s allocation weight and these allocated totals were summed for each year. The allocated citations per publication was then calculated as their ratio:

\[
\text{Year} \times \text{cites} / \text{publication} = \text{total allocated cites} / \text{total allocated publications} \tag{2}
\]

Average allocated citation rates per publication were then calculated for each year from 2000–2006, along with an overall average. The average annual citation rate was 1.13, nearly a 50 percent increase over the official 2010 NRC assessment. As with the publication rate, allocation played a substantial role in the final citation rate calculation. Articles by MSU Geography faculty generated 1,935 citations, while allocated citations by core and new faculty were 1,640.4, a loss of nearly 295 citations, or about 15 percent. MSU Geography gained 48.4 citations from associated faculty outside of the program; Table 2 reports on the net outflow of citations from the program; the net change is strongly negative with a relative magnitude similar to that of publications.

**PART 3: SENSITIVITY AND VAGUENESS**

This part of the article has two subsections: in the first we examine the inherent sensitivity of the NRC research activity metrics to variation in input, while in the second we consider sources and impact of vagueness on the assessment’s definition of research activity elements. Ideally the metrics should be insensitive to reasonable variation in either measurement or definition. Both sections raise distinct and important questions about the validity and appropriateness of the metrics for the NRC assessment of doctorate programs in geography.

**SENSITIVITY**

This section considers four scenarios for faculty allocation and their impacts on publication rate and rankings, enabling us to explore the sensitivity of publication rates and rankings due to variability in counts and allocations.

- **Baseline:** employs the current NRC faculty allocations and publication rates.
- **Recount:** summarized in the previous section, this uses the NRC allocation rates but an adjusted WOS publication total, based on our manual recount.
- **Program Fortress:** also uses the recount data but adjusts faculty allocations by considering only core and new faculty, and allocates their publication contributions 100 percent to the program. This scenario ignores contributions from external associated faculty members. Program Fortress effectively models program research activity as a function of total core and new faculty output.
- **Hide the Children (and Administrators):** removes new faculty and faculty with administrative, nonresearch/nonteaching appointments, along with their articles, from the assessment. The idea is to account only for faculty members with active research and doctoral student advising roles over the 2000–2006 time period.

Using data collected for MSU Geography, publications per faculty were calculated for each of these four scenarios to evaluate the sensitivity of the NRC assessment; results are provided in Table 3. For each scenario the table indicates total article publications, if known, allocated articles and faculty, publication rate, and the rank-order of that publication rate against the published 2010 geography program publication rates.

The baseline scenario is clearly the lowest performer. The second scenario, which was summarized above, introduces the largest jump in rank by more effectively recounting MSU Geography WOS publications. The third scenario ranks two places higher, in spite of the lower article count. This is because the program receives 100 percent of the 165 articles by core and new faculty. Finally, the fourth scenario results in the highest ranking, with a publication rate 2.5 times higher than the baseline. Even though the article count is lower than two of the other scenarios, this is offset by the reduction in the allocated faculty count, and the publication rate is high enough to rank third of forty-nine geography programs.

What justification might a program have for the third and fourth scenarios? The third scenario captures program productivity very directly: for prospective graduate students and their undergraduate mentors considering possible advisors, university administrators concerned with unit output, or researchers interested in comparing program output across an academic discipline, this perspective may well be preferable. The fourth scenario removes faculty with limited opportunities to either advise doctoral students or to have published output over 2000–2006 (or indeed for new faculty, 1981–2006). It restricts analysis to the most relevant faculty. This should have been accounted for in the NRC methodology: according to the NRC definitions faculty should not be assigned as core or associated to a program if they have not served on a Ph.D., admissions, or curriculum committee during the time period (NRC 2010a, 14–15). However, inspection of the MSU Geography list indicates that this procedure was not followed correctly, as at least two core faculty did not meet these criteria. Confusion about calculating faculty allocation was evidently widespread in many universities and doctoral programs during the data collection phase (Carpenter-Hubin and McClelland 2011; Stigler 2010), and this factor alone may explain substantial variation in research activity metrics between geography programs.

We do not claim that MSU Geography’s research article productivity should be ranked 3rd, 9th, 11th, or 33rd; rather, our goal is to determine the sensitivity of the
metrics. Drastic swings in research activity metrics are achievable using modest adjustments to the manner in which underlying metrics are counted. While the analysis here is for a single program, we suspect that many, if not all, geography programs are subject to similar variability in their publication rates under these or other plausible scenarios, and that most if not all geography programs’ research activity metrics would improve with a careful recount of ISI articles and faculty.

**VAGUENESS**

Substantial definitional questions plague efforts to quantify complex phenomena. In this section we consider the impact of these problems for the NRC ratings. A full range of challenges related to defining concepts for an assessment is beyond the scope of this article, but we briefly treat three definitional questions related to conceptualizing geographic research productivity in this section: Who is being counted? What is being counted? When is it being counted?

The previous section identified multiple scenarios that used somewhat different definitions for allocated faculty to address issues of sensitivity associated with who is being counted. These definitional variations resulted in substantially different measures of research activity. A profound challenge to efforts to develop standardized rankings for comparative purposes is that definitions as employed may vary from faculty member to faculty member, from program to program, from institution to institution, and from discipline to discipline. We provide evidence concerning this variability for faculty allocation at two of these levels:

- Seventeen of the twenty-three core MSU Geography faculty were allocated to at least one program other than geography according to MSU documentation for the NRC assessment effort. Conversely only one of approximately thirty external Ph.D. committee members during the time period had any reported allocation to geography. Differences in reporting standards between programs at Michigan State University appear to have led to this gap, which in turn influenced measured research activity.

- Of the nine doctorate geography programs in the Committee on Institutional Cooperation (CIC) consortium (all Big Ten schools), five had three or fewer associated faculty, while the others had twelve, twenty-two, twenty-four, and forty-two associated faculty (all data from NRC 2010c). While this large gap might reflect real differences in graduate advising policies or cultures, we believe this demonstrates vagueness in the NRC data collection methodology and widespread variation in faculty counting practices at different institutions.

Equally important is another source of vagueness: what is being counted? Citations and articles, of course, but these are indirect measures of other things. Citation counts serve as a proxy for research quality and impact. Their use in this context has been controversial for many years (Bornmann and Daniel 2008; Lindsey 1989) We know that the NRC uses journal articles as one proxy for research activity, but we could not determine from NRC’s documentation (NRC 2010b) what categories of articles within the WOS were actually counted; in this study, we took a conservative approach and included only research articles; clearly a more inclusive approach that counted reviews or commentaries would increase the program’s publication rate. But perhaps more critical for geographers is simply the methodology’s reliance on the WOS databases. We note that a number of journals that regularly publish work by geographers do not appear in these databases, including, as of spring 2011, the following highly incomplete list: Cartographica, Cartography and Geographic Information Science, Environment and Planning C, Journal of Cultural Geography, and Transactions in GIS.

Geography is of course a broad field with diverse approaches to research production; many of these approaches do not involve journal articles or grants. For some geographers, and some geography programs, books are valuable or even essential scholarly outputs. For others, peer-reviewed papers in published conference proceedings volumes are important measures of scholarship. However, neither these nor other published works were used in the assessment of geography programs. Geography was evaluated within the social science domain; the publication metric for all programs in the social sciences was WOS article counts. In an independently conducted self-study for the period 2002–2008, MSU Geography faculty reported a total of 282 journal articles, sixty-eight book chapters, and twenty-nine books; clearly the program’s internal definition of publications is more inclusive than the NRC methodology’s, with consequent impact on measured research activity. All of these publications: research articles in journals not included in the WOS, books, book chapters, and other publications, were essentially invisible within the 2010 NRC assessment framework. Indeed, they depressed program activity metrics, as time spent on them reduced the amount of content a scholar could generate that was visible to the NRC methodology.

A third definitional challenge is the problem of when; over what time period should research activity be assessed to develop metrics and rankings, and what is the shelf life of those rankings? We maintain that it is not long. Of MSU Geography’s twenty-seven core and new faculty included in the assessment, fourteen were not associated with the program for the entire 2000–2006 duration. Further, between early 2007 and fall 2010 when the rankings were published, five faculty on the 2006 list had left the program and six new faculty had joined. Given this amount of change, it is difficult to see how, even with perfectly measured data, a prospective graduate student or administrator might employ the historic NRC assessment data to make present-day judgments.
PART 4. IMPLICATIONS AND CONCLUSION

In this final section we consider several implications of our findings: the impact of the decision to use interdisciplinary faculty associations and the use of WOS databases on the NRC rankings. We then evaluate the overall utility of these rankings, and conclude with some open questions concerning future assessment efforts of geography graduate programs.

Geography is among the most highly interdisciplinary of the sixty-two fields included in the assessment. The decision by NRC to assign portions of faculty to programs (and thereby portions of their scholarly output) based on their dissertation committee service was devised specifically to benefit departments that rely on contributions from faculty in allied fields. The consequent confusion and evident variability in faculty allocation reporting between faculty, programs, and universities may obscure any meaningful measured variation in research productivity. Indeed, this was identified by Stigler (2010) as one of the most significant problems with the 2010 rankings; this article is the first to quantify its effects for a particular program.

The NRC report (2010a) expresses regret at the imperfection of the WOS data for capturing certain kinds of publications. Indeed, as discussed above, many kinds of scholarly output do not appear in the WOS. But to this we also must consider the difficulty in attributing articles properly to their authors. Each author’s affiliations over the multidecade assessment period, as submitted to their publishers, subsequently to WOS, and as independently reported to the NRC in 2007, are subject to large amounts of error and variation. In WOS queries conducted for this research, we noted a misspelling of an author’s last name and variably reported ZIP codes. In other cases, author affiliations may be made to a funding organization or employ an alternate department name. Our meticulous literature search revealed more than double the number of peer-reviewed articles than those found by the NRC contractor. These articles were contained within WOS, but third-party searchers failed to find them due to variations in author names and affiliations. We maintain that this approach to publication counting favors faculty with unique names and who lack multiple affiliations.

Given the import and potential influence of the NRC assessment’s results it is imperative that it be accurate, robust, and relevant to the disciplines being evaluated; the results of our meta-assessment reveal significant concerns in each of these three areas.

- **Accuracy:** Substantial numbers (twenty or more) of associated faculty were not counted, only 65 percent of allocated publications appearing in the WOS database were actually identified by the NRC contractor, and citations averages were off by 50 percent.
- **Robustness:** Variations in various faculty- and article-counting scenarios produced large changes in the research productivity metrics. Furthermore, variation in how different universities and programs interpreted and counted faculty appeared to be significant.
- **Relevance:** Much scholarship was ignored by the rankings. There has been substantial faculty turnover since 2000. It is not clear that the variables directly, indirectly, or effectively measured the program characteristics they meant to capture.

Geographers are well aware that data error and model uncertainty readily propagate into model outputs; the famous phrase “garbage-in, garbage-out” is a common mantra in lectures around the country. In the present case, lack of data reliability hinders the interpretation and utility of the NRC assessment results. Our attempt, in good faith, to reconstruct the research activity metrics for a single geography program exposed troubling discrepancies. We doubt our findings are unique; indeed, we expect that many geography programs would identify serious discrepancies with the NRC data. We encourage self-studies by other geography programs to help identify the extent and variability of data inaccuracy problems in the 2010 NRC assessment.

How appropriate then is the NRC’s recent assessment model for geography? Does the design and implementation of this assessment enable it to accurately differentiate our doctorate programs? The results of our data-based evaluation of this assessment suggest that the answer is a resounding “no,” which is in line with the findings of others (e.g., Carpenter-Hubin and McClelland 2011; Cole 2011). Indeed, for considering research activity metrics at least, we disagree with those who suggest that the 2010 assessment data have any clear utility (e.g., Inside Higher Ed News 2011; Colglazier and Ostriker 2010), as any observed variation in metrics or ratings may simply reflect systematic collection error, definitional confusion, or demographic differences in doctorate programs. The revised rankings, released in April 2011, did not address the data quality issues raised in this article.

In light of this, we look forward and pose a few generic questions on the nature of assessment generally for geography doctorate programs:

- What are the key concepts that distinguish our successful doctorate programs?
- Can those concepts be crisply defined and accurately measured across all programs?
- How imperfect can an assessment be before it is worse than no assessment at all?

We hope that this article contributes to an active and informed discussion within geography on the worth, nature, utility, and design of future assessments of our discipline.

REFERENCES

NRC Research Data Uncertainty


———. 2010c. A Data-Based Assessment of Research-Doctorate Programs in the United States; Data Table in Excel. Washington, D.C.: National Academies Press Committee to Assess Research-Doctorate Programs.
