Return on Investment

Using NRC Data with Faculty Members, Departments, and Visiting Committees

Lydia Snover, Scott Barge, Jon Daries, and Gregory Harris

Office of the Provost, Institutional Research
Massachusetts Institute of Technology

3/4/2011

Abstract

In this paper we describe several ways in which MIT has deployed information from the NRC data set as a resource for benchmarking and decision making. We discuss data from additional sources that are important to help end users contextualize the NRC data and identify some limitations of the NRC data set. We also identify limits to how MIT has been able to use the NRC data. Overall, faculty members are simultaneously drawn to and skeptical of the rankings data. In many instances, this skepticism has driven ultimately fruitful conversations about the availability and usefulness of program-level data. Specifically, we have advanced our own local understanding of appropriate measures of program quality, improved processes to streamline collection of data, and enhanced the sophistication with which we think about supporting programs with relevant metrics.

Contents

Introduction

In the years leading up to the release of the NRC's 2010 Research Doctorate study, institutions invested significant resources—time and money—in preparing and validating data. While much of the initial buzz surrounding the rankings-related results of the study has passed, the rankings and the data behind them remain as a resource for faculty members, department heads, and visiting committees. In this paper we describe several ways in which MIT has deployed information from the NRC data set as a resource for benchmarking and decision making. We also discuss data from additional sources that are important to help end users contextualize the NRC data and identify some limitations of the NRC data set.

In Part I, we discuss how MIT conveyed the methodology and illustrative rankings to end users: with descriptive statistics, weights, rankings, and combined reports that were a synthesis of all three. For each type of report we frame our discussion around common questions from faculty members and then walk through key data elements and options for presenting and contextualizing program results from the NRC study. For each type of report, we also provide examples of tables and charts that were used to communicate detailed versions of the 20 indicators.

In Part II, we briefly introduce the visiting committee—MIT's longstanding approach to program review—and then identify ways in which the NRC data are incorporated into a broader context for program review. We also discuss this data context, including the use of publications and citations from Academic Analytics.

In both sections, we identify limits to how MIT has been able to use the NRC data. For example, the interdisciplinary indicator has proven difficult to interpret and apply in ways that are meaningful to program faculty. Specific programs—notably computer science and engineering—have identified limitations of the publication and citation data. We are also in communication with NRC researchers regarding our faculty's scrutiny of the methodology for calculating the percentage of doctoral graduates with careers in academe.

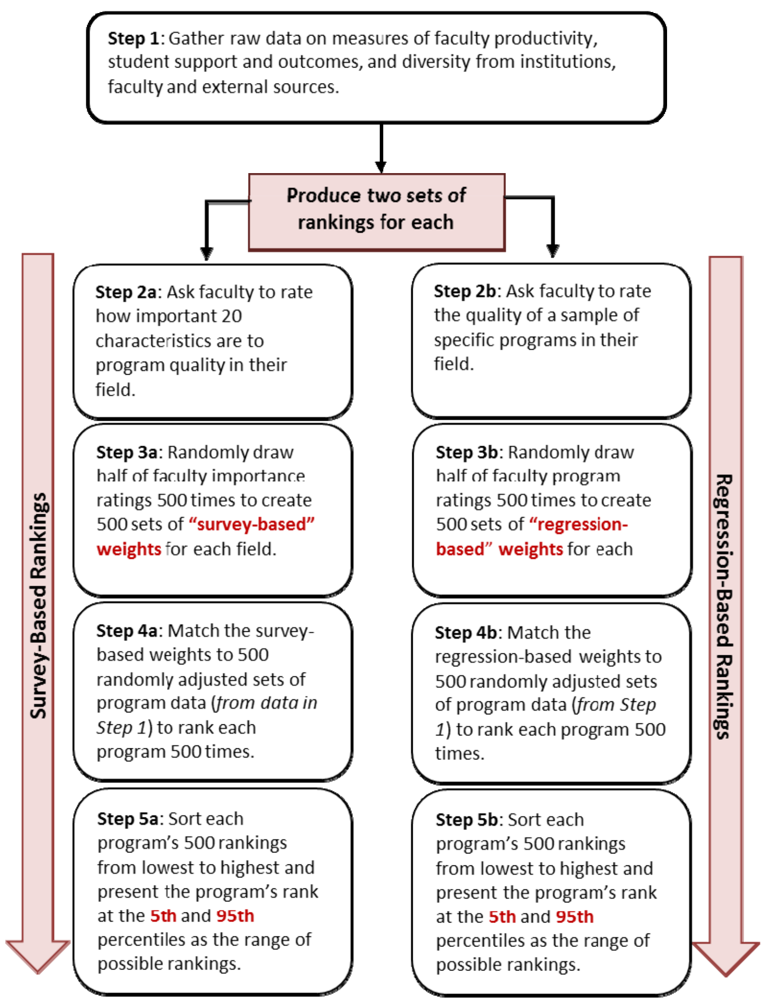
Part I: Dissemination of NRC Results to Campus

NRC's Methodology

One of the first steps we took at MIT, even prior to the release of the final NRC report, was to familiarize key internal constituents with the methodology used for the rankings. We developed an in-house presentation to describe the ranking methodology for non-statistical users and drew upon summaries and descriptions prepared by others in the broader institutional research community. In addition to distribution of our presentation, we published an article in MIT's faculty newsletter and presented the methodology to key administrators in the run-up to the release of the report.

After the release of the rankings, we distributed program-specific reports to all departments and included a second, brief explanation of the ranking methodology using a flow chart (See Figure 1). The most common questions regarding methodology were about the differences between the regression-based and survey-based ranking approaches. We address those questions below.

Figure 1. Simplified schematic of NRC ranking methodologies



Defining and Describing the Measures

The NRC study is built around 20 quantitative indicators, or measures, that are used to rank doctoral programs.[1] When presenting the descriptive statistics for these indicators to our end users, we took care to introduce and define the variables. Specifically, we aimed to identify the timeframe for the data and the source of the data (internal, NRC, survey, etc.). Our basic reports present MIT's values for each indicator alongside field mean and other descriptive statistics (see Figure 2 below). Each indicator is also linked to the specific definition provided in the NRC's methodology report.

The first of the common questions we received from our faculty concerns transparency. They often wanted to know the source of the data (see second column in Figure 2), the verifiability of the data, and the transparency of the calculations used to generate certain data. In the case of indicators collected and/or submitted by MIT, we owned the verification process and were therefore able to validate estimates when questioned by faculty. However, in the cases of several of the most weighty and important indicators (e.g., faculty productivity), the raw data used were unavailable for verification. In some cases, we were able to use Academic Analytics data to provide some context for the NRC's faculty productivity measures (publications, citations, and honors and awards).

In some cases, the lack of transparency in how indicator values were calculated limited our ability to identify sources of discrepancies between internal figures and NRC data. Publishing the numerator and denominator (or other relevant background data) would have allowed departments to more easily reconcile NRC's estimates with internal estimates. Further, the inclusion of supporting data would both disclose the (sometimes very small) cell counts on which quotients were based and allow recombination of smaller programs for more flexibility in comparisons. For example, an institution could combine data from multiple doctoral programs that are in a single organization unit to provide that unit or department with more applicable evaluations of its programs. Examples of all calculations for all variables used in the study would also have answered many faculty questions regarding the source of the data used.

Among the measures used, three key variables were frequently questioned by MIT faculty: average GRE scores, the number of graduates in academe, and the number of per-capita publications and citations. In some programs at MIT, students are not required to take or submit GRE scores. For those programs, MIT did not report average GRE scores. However, NRC used imputed scores (the average GRE score for all programs in the field), calling into question the applicability of those program's rankings for the faculty members of those programs.

Regarding the post-graduation plans of graduates, MIT's departments expended considerable effort tracking down all of their doctoral recipients over the five-year period covered by the survey and determining their current job status. In spite of this effort, NRC chose instead to use the 2005 NSF Doctoral Records File based on the Survey of Earned Doctorates (SED) to measure percentage of

---

[1] Much of the NRC documentation refers to the 20 quantitative variables as "measures." In our on-campus presentation we often used the term "indicators," as in *indicators of quality*. However, usage has not been entirely consistent and the terms "variable," "indicator," and "measure" should generally be assumed to be interchangeable.

students with academic plans. The NRC may have assumed that all universities require students to complete this survey; however, MIT's Institutional Review Board will not allow MIT to require students to complete *any* surveys, and thus only a fraction of MIT students were included in the data. In addition, MIT's SED response rate differed widely by program, ranging from 40% to 100%. Since the NRC used the ratio of students who responded affirmatively on the SED to the *total* number of students who graduated during those five years to determine this statistic, the assumption was that any student who did not complete the survey did not pursue an academic career, which is very likely incorrect. Even if the SED were to be used, given its low response rate at MIT, it would be more accurate to report the ratio of respondents with academic positions to the total number of respondents. Also, because this survey is administered at the time of graduation, it may not represent the actual percentage of students who eventually take academic positions.

For measures of faculty productivity, the NRC took steps to tailor the counting of publications and citations to particular fields, but we found that faculty members frequently identified discipline-specific limitations of these data. For example, in the engineering fields the fact that conference proceedings were not included was seen as a shortcoming. In many instances, faculty members questioned the accuracy of the productivity indicators and expressed concerns that the estimates were too low in comparison to internally-collected figures.

The last question regarding the indicators themselves was how the data aligned to a given department. Since the programs at specific institutions mapped onto the programs specified in the rankings in different ways, one MIT department could have multiple fields associated with it. Conversely, other institutions could have more than one doctoral program per field (e.g. Harvard submitted three separate economics programs from its different schools). While questions of this nature can be addressed with careful attention to documentation and clear explanation, they highlight the inevitable difficulties that come with making inter-institutional comparisons and integrating external data into local evaluation processes that are often built around internal organizational structures.

Figure 2. Example table of descriptive statistics for MIT's Cell and Developmental Biology program.

Table 1. Program values for each indicator plus overall summary statistics for the field.

Indicator	Source of Data	MIT Value	All Cell and Developmental Biology Programs (n=122) Mean	Min	5th %tile	95th %tile	Max	Std. Dev.
Publications per Allocated Faculty	NRC	**4.725**	1.658	0.286	0.679	2.829	4.936	0.742
Cites per Publication	NRC	**8.785**	4.364	1.122	1.824	8.609	11.256	2.008
Percent of Faculty with Grants	Survey	**98.01%**	85.94%	33.33%	60.37%	100.00%	100.00%	13.82%
Percent Faculty Interdisciplinary	IR	**0.00%**	25.66%	0.00%	0.00%	66.67%	85.71%	21.76%
Percent Non-Asian Minority Faculty	IR	**0.00%**	2.87%	0.00%	0.00%	10.00%	16.67%	3.48%
Percent Female Faculty	IR	**24.44%**	26.88%	0.00%	13.79%	39.39%	63.64%	8.83%
Awards per allocated faculty	NRC	**7.696**	0.635	0.000	0.000	2.155	7.696	1.004
Average GRE-Q	IR	**763**	702	571	632	761	787	40
Percent 1st yr. students w/ full support	IR	**100.00%**	96.08%	0.00%	83.30%	100.00%	100.00%	16.07%
Percent 1st yr students with portable fellowshi	IR	**71.43%**	17.26%	0.00%	0.00%	85.71%	100.00%	28.31%
Percent Non-Asian Minority Students	IR	**10.17%**	11.02%	0.00%	0.00%	28.57%	45.46%	9.78%
Percent Female Students	IR	**58.23%**	53.40%	25.00%	37.04%	71.88%	92.86%	11.16%
Percent International Students	IR	**17.72%**	32.36%	0.00%	5.00%	70.83%	88.24%	20.21%
Average PhDs 2002 to 2006	IR	**14.0**	5.2	1.0	1.4	11.6	49.0	5.1
Percent Completing within 6 years	IR	**55.71%**	50.02%	2.27%	12.67%	83.33%	100.00%	19.54%
Time to Degree Full and Part Time	IR	**6.1**	5.6	3.5	4.7	6.8	8.0	0.7
Percent students in Academic Positions	NRC	**n.a.**	22.64%	0.00%	0.00%	45.46%	62.50%	14.32%
Student Work Space	IR	**1**	1	-1	-1	1	1	1
Health Insurance	IR	**1**	1	-1	-1	1	1	0
Number of student activities offered	IR	**16**	17	10	14	18	18	2

Indicator definitions are available in Appendix 1 of this report (click on indicators for link to definitions)

Green = MIT Value > Mean for Field

Red = MIT Value < Mean for Field

Explaining Weights

The weights used to derive MIT's ranges of rankings for each program were another area of the NRC results that we presented to our internal constituents. For each of our programs we prepared a table presenting the weights for each indicator used to determine MIT's 5th and 95th percentile ranking (See Figure 2). In order to give a sense of the weights used in a given field, we also presented an average of the plus- and minus-one standard deviation of the weights, as provided by NRC, for each field (not shown).

Our faculty members often assumed these weights were the same across all institutions. In other words, they thought that the exact same set of weights was used to determine our peers' program rankings at the 5th and 95th percentile. To allay these concerns we most often drew upon our various methodology documents to work through these types of questions. Another common question among end users was how to interpret negative weights. In explaining how the negative weights might be understood, we found it important to highlight the differences between how the weights were derived for the regression-based and survey-based rankings. Further, we felt it was critical to note that a negative weight did not necessarily imply that efforts on the part of the department to increase/decrease the associated indicator would impact future rankings.

Figure 3. Example of a table of weights as presented to internal constituents.

Table 2. Indicators and indicator weights used to calculate MIT's 5th and 95th percentile ranks for regression-based and

	Source of	R Weights (Regression-Based)		S Weights (Survey-Based)	
Indicator	Data	5th	95th	5th	95th
Publications per Allocated Faculty	NRC	0.042	0.000	0.130	0.131
Cites per Publication	NRC	0.109	0.063	0.099	0.105
Percent of Faculty with Grants	Survey	0.055	0.091	0.172	0.168
Percent Faculty Interdisciplinary	IR	0.075	0.024	0.041	0.040
Percent Non-Asian Minority Faculty	IR	0.032	-0.013	0.010	0.011
Percent Female Faculty	IR	-0.060	-0.031	0.018	0.018
Awards per allocated faculty	NRC	0.148	0.055	0.059	0.059
Average GRE-Q	IR	0.044	0.107	0.081	0.077
Percent 1st yr. students w/ full support	IR	-0.021	0.025	0.058	0.060
Percent 1st yr students with portable fellowshi	IR	0.016	0.055	0.039	0.042
Percent Non-Asian Minority Students	IR	-0.044	0.023	0.023	0.023
Percent Female Students	IR	-0.013	0.048	0.020	0.020
Percent International Students	IR	0.070	-0.005	0.008	0.008
Average PhDs 2002 to 2006	IR	0.094	0.160	0.023	0.022
Percent Completing within 6 years	IR	0.066	0.079	0.060	0.060
Time to Degree Full and Part Time	IR	0.012	0.048	-0.033	-0.032
Percent students in Academic Positions	NRC	0.028	0.056	0.078	0.079
Student Work Space	IR	-0.006	0.032	0.005	0.005
Health Insurance	IR	-0.037	-0.071	0.004	0.005
Number of student activities offered	IR	0.028	-0.014	0.037	0.034

Indicator definitions are available in Appendix 1 of this report (click on indicators for definitions in electronic files)

Presentation of Rankings

The rankings themselves drew more interest from senior administrators, deans, department chairs, and individual faculty than any other component of the NRC report. Our most common and direct method of presenting the NRC rankings to our faculty members was through graphical representations of MIT's rankings alone and in comparison to the ranks of peer institutions. We presented these rankings in a variety of ways: all rankings (R, S, and dimensional) for a particular program (see Figure 4 and Figure 6); the range of rankings (R and S combined) as compared to peers for a particular program (see Figure 5); and the range of R and S rankings for all programs across MIT (Figure 7).

For our core presentations of the rankings (Figure 5 and Figure 6), we made the judgment to let the numerically smallest rank (from either R or S) drive the ordering of peers in the chart. Programs were sorted by: (1) highest of R or S rank at the 5th percentile, ascending; (2) lowest of R or S rank at the 95th percentile, ascending; and lastly (3) institution name, ascending.

Figure 4. R, S, and dimensional rankings for a specific program.

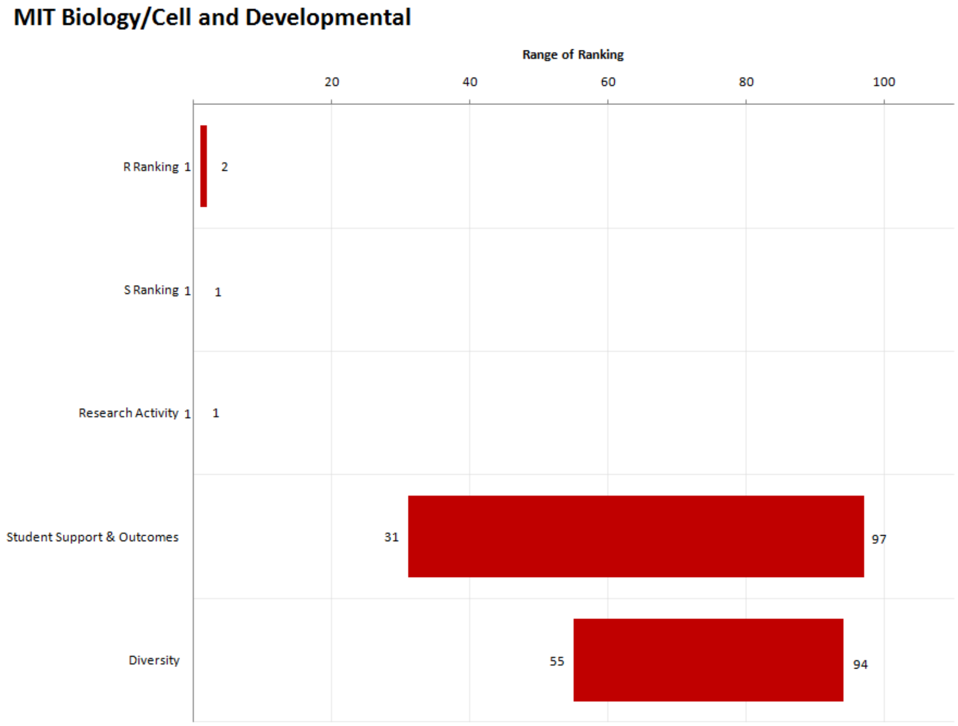
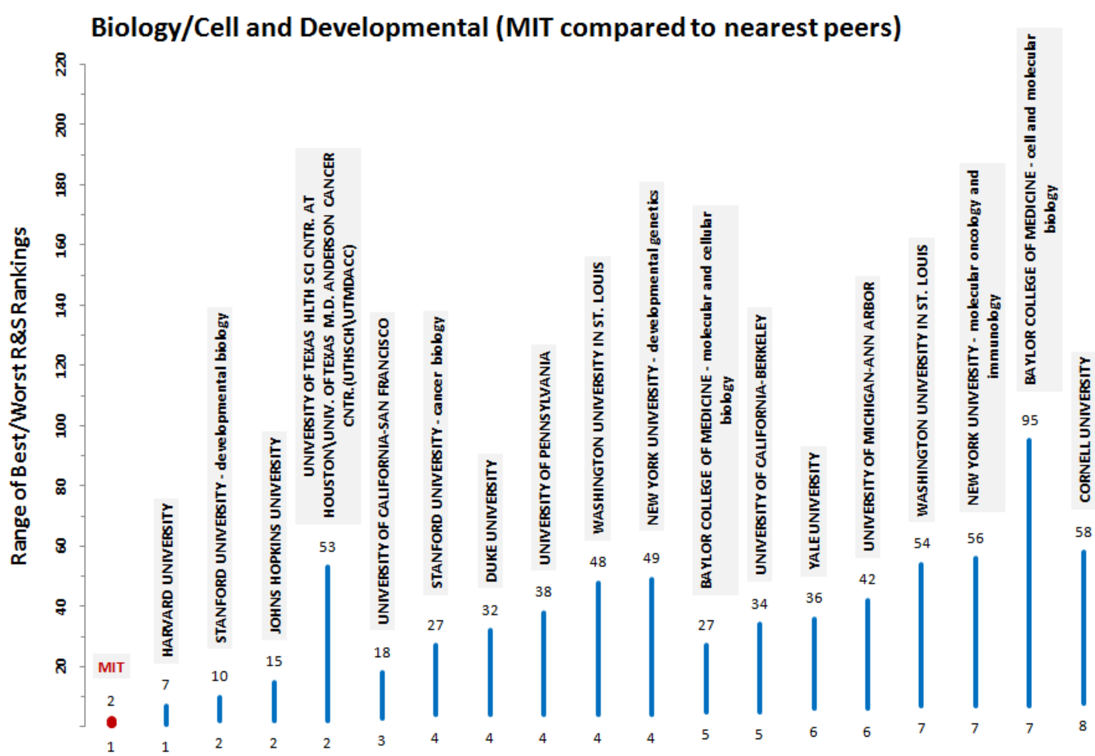


Figure 5. A graphical representation of MIT's ranking range for a program, compared to peers.



MIT and nearest peers based on sorting scheme described in Table 3.

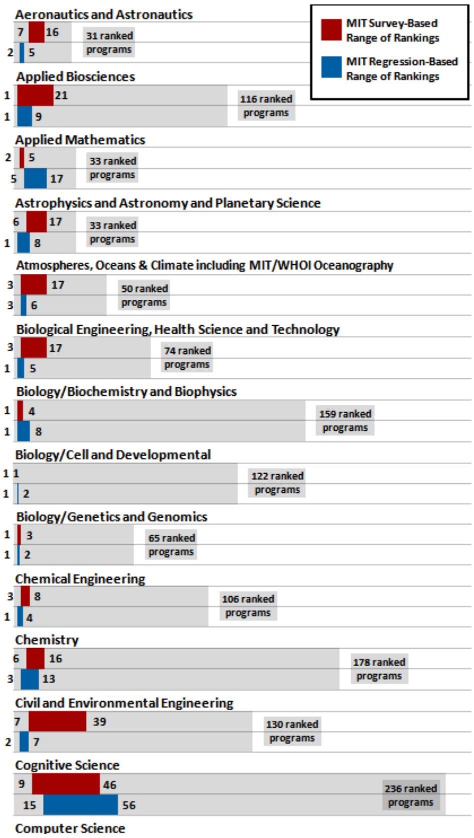
Figure 6. A table of rankings with MIT presented alongside peers.

Table 3. Ranges of rankings for MIT's program compared to other participating programs; R, S, and dimensional rankings (N=122).

	R Ranking (Regression-		S Ranking (Survey-Based)		Best/Worst of R and S Rankings		Research Activity		Student Support & Outcomes		Diversity	
Institution	5th	95th	5th	95th	5th	95th	5th	95th	5th	95th	5th	95th
MIT	1	2	1	1	1	2	1	1	31	97	55	94
HARVARD UNIVERSITY	1	2	2	7	1	7	2	18	22	79	60	99
STANFORD UNIVERSITY - developmental biology	3	7	2	10	2	10	2	7	60	117	69	104
JOHNS HOPKINS UNIVERSITY	3	11	2	15	2	15	3	32	14	74	27	70
UNIVERSITY OF TEXAS HLTH SCI CNTR. AT HOUSTON\UNIV. OF TEXAS M.D. ANDERSON CANCER CNTR.(UTHSCH\UTMDACC)	11	53	2	9	2	53	2	10	5	50	7	25
UNIVERSITY OF CALIFORNIA-SAN FRANCISCO	4	13	3	18	3	18	2	8	87	118	93	115
STANFORD UNIVERSITY - cancer biology	4	16	5	27	4	27	6	55	20	85	29	72
DUKE UNIVERSITY	6	25	4	32	4	32	4	41	5	64	73	107
UNIVERSITY OF PENNSYLVANIA	4	19	5	38	4	38	6	57	9	61	68	103
WASHINGTON UNIVERSITY IN ST. LOUIS	4	16	7	48	4	48	8	71	23	90	82	112
NEW YORK UNIVERSITY - developmental genetics	4	49	4	27	4	49	7	53	3	64	4	16

We published Figure 7 on MIT's website and in the MIT faculty newsletter. This graphic presents regression-based and survey-based rankings separately—by program—listed in alphabetical order. In the public presentation we also represented the total number of programs ranked in each field to give the audience a better perspective on relative rank. We did this in response to faculty questions regarding context. Is a particular program ranked fourth out of ten or fourth out of one hundred? Other than the single public chart, we didn't present rankings within the broader context of the total number of programs ranked.

Figure 7. Table of all MIT programs' R- and S-based rankings displayed graphically.



Faculty members commonly asked about the relative importance of regression-based and survey-based rankings—which one should departments be more concerned with? In general, MIT did better using the regression-based approach because MIT's relatively large programs (in terms of the number of PhD graduates) benefitted from the large weights on the size-related indicator. As departments have begun using the rankings for their own purposes, we have learned that some departments are using regression-based and survey-based combined, only regression-based, or only survey-based to meet their needs.

Many end users questioned the methodology behind the regression-based rankings and, specifically, to what extent they were reputational. It was difficult for end users to understand the way in which the regression-based rankings were not reputational, but yet included a reputational component in the methodology. In general, the survey-based approach was more straightforward to understand. There was, however, some sentiment among faculty members and department heads that a truly reputational component to a ranking project was justifiable. In meeting with faculty, it was clear that they have a strong sense of who their peers are, and where the pockets of excellence in their respective fields exist. Peer review is an integral part of the academic culture at research universities, and a reputational ranking that is based on faculty assessment of programs in their field would have credibility with faculty. Furthermore, in the eyes of some faculty, the absence of well-regarded, high-quality programs in the higher ranks diminished the credibility of the NRC results.

The way in which interdisciplinarity was evaluated in the rankings was of common concern to our faculty members. NRC originally planned to measure the percentage of faculty with interdisciplinary research through a survey of faculty. Later, NRC changed the methodology by instead counting the ratio of associated faculty to total faculty members. This second methodology generated some confusion because it was not widely understood that this was a mechanism for measuring interdisciplinarity, and at MIT departments tend to designate *all* faculty as core because they are integral to the mission of that program regardless of their departmental affiliation(s). Consequently, most MIT departments did not include any associated faculty members. Moreover, many MIT departments contain faculty members who have dual appointments with another unit (50% appointments in both units), yet these faculty members were not counted as interdisciplinary because they were not listed as associated.

More generally, our faculty members often voiced concern over how program faculty lists were determined, even though they had responsibility for the determination of their lists. One problem was the lengthy delay between submitting faculty lists and the release of the rankings. The individuals who compiled the list and the people associated with the program had changed in many cases. Ambiguity in the criteria for determining which faculty members were eligible to be included resulted in inconsistencies in how departments reported their faculty members. Since the number of faculty in a program was used to produce many of the per-faculty indicators, the impact of some of these fundamental concerns reverberated into doubts about the accuracy of other measures. Although NRC provided departments with the opportunity to revise faculty lists, faculty members felt that without an understanding of how the lists would impact related variables they could not make informed decisions about how to report their faculty members according to NRC's intended criteria.

Faculty were puzzled by the inclusion of at least three indcators: Student Work Space, Health Insurance, and Number of Student Activities offered. In the case of Health Insurance, this indicator was compiled at the institutional level, beyond the control of individual programs. Number of Student Activities is largely institutionally driven at MIT. Another weight that faculty found problematic was Percent International Students. In addition to the use of positive and negative weights across disciplines, faculty wanted more nuance in terms of the *quality* of international students, given that two programs with the same makeup of students could differ widely. While some institutions recruit international students in order to fill seats and increase tuition income, other institutions admit a smaller number of high-quality international students.

One of the most important faculty productivity metrics in science and engineering relates to the amount of external research funding that faculty and programs are able to attract. External research funding is evidence of the quality of research conducted by the faculty in a program and the resources available to support graduate students. The NRC metric only captured the percentage of faculty with grants and therefore was of limited use. In addition, at MIT some department chairs reported that the NRC metric did not correspond to actual funding patterns for the faculty in their program.

Combined Presentation

For all of MIT's ranked programs, we generated reports that combined the above measures into a single analysis of the program. These included the MIT values of indicators, the field mean, the field standard deviation, and MIT's relative standing on each indicator, along with the weights used for that program.

The primary question these reports were designed to answer was which indicators most heavily influenced the ranking of the program. In the example report (Figure 8 below), end users could see the value of each indicator and its respective weight using the color-coded *z*-scores and histogram weights. In this type of presentation, it is easier to see the differential impact of program size on the rankings (higher impact for regression-based than for survey-based results). This layout allowed faculty to see how MIT's relative performance on the indicator combined with the weight to drive the ranking. For some constituents, we augmented this chart with data for selected peers to reveal how specific schools fared on specific measures.

Figure 8. An example of a combined report of NRC results for a particular program.

MIT Biology/Cell and Developmental Indicator Values and Weights								
Field: Cell and Developmental Biology								
				Program Rank:	1	2	1	1
	Indicators				Weights			
Indicator	**MIT Value**	Field Mean	Field St. Dev.	**MIT Above (green) or Below (red)(Z-Score)**	R @ 5th	R @ 95th	S @ 5th	S @ 95th
Publications per Allocated Faculty	**4.7245**	1.65802	0.741828	**4.13**	0.04	0.00	0.13	0.13
Cites per Publication	**8.785**	4.36445	2.007733	**2.20**	0.11	0.06	0.10	0.11
Percent of Faculty with Grants	**98.0%**	0.85942	0.138204	**0.87**	0.06	0.09	0.17	0.17
Percent Faculty Interdisciplinary	**0.0%**	0.25656	0.217621	**-1.18**	0.08	0.02	0.04	0.04
Percent Non-Asian Minority Faculty	**0.0%**	0.02866	0.03482	**-0.82**	0.03	-0.01	0.01	0.01
Percent Female Faculty	**24.4%**	0.26879	0.088262	**-0.28**	-0.06	-0.03	0.02	0.02
Awards per allocated faculty	**7.6957**	0.63495	1.004412	**7.03**	0.15	0.06	0.06	0.06
Average GRE-Q	**763.226**	702.299	39.63468	**1.54**	0.04	0.11	0.08	0.08
Percent 1st yr. students w/ full support	**100.0%**	0.96084	0.160686	**0.24**	-0.02	0.03	0.06	0.06
Percent 1st yr students with portable fellowshi	**71.4%**	0.17259	0.283073	**1.91**	0.02	0.06	0.04	0.04
Percent Non-Asian Minority Students	**10.2%**	0.11018	0.097775	**-0.09**	-0.04	0.02	0.02	0.02
Percent Female Students	**58.2%**	0.53401	0.111574	**0.43**	-0.01	0.05	0.02	0.02
Percent International Students	**17.7%**	0.32365	0.202144	**-0.72**	0.07	-0.01	0.01	0.01
Average PhDs 2002 to 2006	**14**	5.24426	5.145777	**1.70**	0.09	0.16	0.02	0.02
Percent Completing within 6 years	**55.7%**	0.50016	0.195373	**0.29**	0.07	0.08	0.06	0.06
Time to Degree Full and Part Time	**6.1**	5.6483	0.724867	**0.62**	0.01	0.05	-0.03	-0.03
Percent students in Academic Positions	**n.a.**	0.22639	0.143167		0.03	0.06	0.08	0.08
Student Work Space	**1**	0.68852	0.728204	**0.43**	-0.01	0.03	0.01	0.01
Health Insurance	**1**	0.86885	0.497113	**0.26**	-0.04	-0.07	0.00	0.01
Number of student activities offered	**16**	16.6311	1.65245	**-0.38**	0.03	-0.01	0.04	0.03

Part II: NRC Data & Program Review at MIT

Background

MIT has a well-established system for external program review through its visiting committee structure, which has been in place since 1875. Visiting committees at MIT operate as advisory groups to the Corporation and the administration, providing valuable insight on current activities and future directions for each academic program and other major areas of the Institute (31 in total). Each visiting committee convenes every two years.

As part of the visiting committee process, in consultation with the provost, dean of the school, and department head, Institutional Research prepares a series of reports summarizing departmental strategic indicators and outcomes. These reports showcase ten years of departmental data, often in the context of internal peers and more recently with external peers. The reports cover a broad spectrum of topics: faculty demographics, faculty productivity measures, student majors and degrees, student outcomes, postdoctoral and graduate student appointments, instructional indicators, research indicators, financial indicators, graduate admissions, doctoral time to degree, doctoral cohort analysis, and rankings.

Use of NRC Rankings

Institutional Research works with department chairs to determine which, if any, NRC data are included in their Strategic Indicators report. To date, most department heads have opted to include the NRC overall rankings data only, usually selecting the combined approach to presentation shown in Figure 5 above. The NRC rankings are often shown alongside other graduate-level rankings data (e.g. U.S. News & World Report). We have not included individual NRC indicator data in the strategic indicator reports, although our current reports have incorporated similar metrics (publications, citations, honors and awards, research expenditures)—often informed by lessons learned from the NRC project. Currently, data from sources other than NRC provide us better flexibility, transparency, and timeliness.

Use of Additional Data from Other Sources

Three reasons drive our decision not to use the individual indicator data from the NRC study for our departmental reports. First, as the NRC data collection occurred in 2006, the indicators are quite dated. We update our strategic indicators every two years and strive to use the most current data available. Second, we are somewhat constrained by the NRC taxonomy for program disciplines, which does not align perfectly with our departmental structure. While this is to be expected with any externally-generated taxonomy, we do not have access to record-level (at the faculty level) NRC data in order to configure the NRC metrics to match other data used in our departmental reports. This is a constraint we face with faculty productivity measures pulled from other sources as well. Third, in most cases we cannot verify the accuracy of the NRC indicators collected from external sources. In Table 1 we identify the major types of data covered by the NRC indicators, the alternate sources we draw upon for program review purposes, and a brief description of our rationale for selecting the alternate source.

Table 1. Major areas of NRC study data, alternate sources, and rationale for using them.

Data type	Alternate Source	Rationale
Faculty Productivity	Academic Analytics	AA has allowed access to institutional unit-record (faculty) publication and citation data for validation purposes; timely publication of data
Student and Faculty Diversity	AAU Data Exchange, IPEDS	Timely publication of data; flexibility afforded through use of CIP taxonomy
Doctoral Cohort Analysis and Time-to-Degree	AAU Data Exchange	Timely publication of data; flexibility afforded through use of CIP taxonomy
Research Expenditures	NSF	NRC only included percent of faculty with grants

Conclusion

Considering the work leading up to the NRC study, the preparation of reports, and the subsequent conversations with faculty members at MIT, several points stand out. First, faculty members are simultaneously drawn to and skeptical of the rankings. In many instances, this skepticism has driven ultimately fruitful conversations about the availability and usefulness of program-level data. Specifically, we have advanced our own local understanding of appropriate measures of program quality, improved processes to streamline collection of data, and enhanced the sophistication with which we think about supporting programs with relevant metrics. Second, we have observed a great deal of variation in faculty members' level of engagement with understanding the data and the rankings. Some faculty members were interested in as much detail as we could provide. Others sought a high-level, "what's the main takeaway" summary, and still others were eager to disassociate their program from the ranking effort altogether.