How Faculty Demonstrate Impact: A Multi-Institutional Study of Faculty Understandings, Perceptions, and Strategies Regarding Impact Metrics

Caitlin Bakker, Jonathan Bull, Nancy Courtney, Dan DeSanto, Allison Langham-Putrow, Jenny McBurney, and Aaron Nichols*

Faculty and institutions are increasingly called upon to present succinct, quantified descriptions of their research impact to administrators, funders, legislators, and academics. This project sought to explore how and what researchers think about research impact measures across disciplines and institutions. Presenters will discuss findings from a multi-institutional survey of faculty (n=1202) addressing their familiarity with metrics and altmetrics and impression of the accuracy of these metrics. We discuss how researchers use such statistical measures to demonstrate the importance of their scholarship, and their attitudes towards use of scholarly metrics by administrators. We also address possible implications for librarians supporting these faculty members.

Introduction
Bibliometrics, or the statistical measure of scholarship, has been an evolving part of the academic landscape for decades.1 Recently, scholarly metrics and altmetrics data and tools have become increasingly more visible in the academy and accessible to the average faculty member.

This study considered a number of questions related to faculty views of scholarly metrics, including:

- How familiar are faculty with scholarly metrics and altmetrics?
- How accurate do faculty perceive scholarly metrics and altmetrics to be?
- How important are scholarly metrics and altmetrics to faculty?
- How much weight do faculty believe should be given to scholarly metrics and altmetrics?
- How are scholarly metrics and altmetrics used in the promotion and tenure (P&T) process?

By gathering data from multiple institutions of various sizes and missions, the authors hope to show trends in how faculty of all ranks perceive, interpret, and use scholarly metrics. In addition, this study seeks to explore the use of scholarly metrics in the P&T process.

* Caitlin Bakker, Biomedical Research Services Librarian, University of Minnesota Health Sciences Libraries, cjbakker@umn.edu. Jonathan Bull, Scholarly Communications Librarian, Valparaiso University, jon.bull@valpo.edu. Nancy Courtney, Research Impact Librarian, Ohio State University Libraries, courtney.24@osu.edu. Dan DeSanto, Instruction Librarian, University of Vermont, daniel.desanto@uvm.edu. Allison Langham-Putrow, Scholarly Communications Librarian, University of Minnesota–Twin Cities, lang0636@umn.edu. Jenny McBurney, Research Services Coordinator & Social Sciences Librarian, University of Minnesota–Twin Cities, jmcburne@umn.edu. Aaron Nichols, Access/Media Services Librarian, University of Vermont, aaron.nichols@uvm.edu
Literature Review

A number of recently published articles provide overviews of the scholarly metrics landscape. One such review by Margaret Mering in *Serials Review* outlines in very clear language the most common impact metrics and the topics associated with statistical measures of scholarly output. Ludo Waltman's article, "A Review of the Literature on Citation Impact Indicators," also provides a good overview but includes greater detail in discussions of specific impact measures, citation databases, problems with certain impact measures, and trends associated with the topic. Waltman provides an introduction to the differing publication practices and the ensuing citation patterns between the social sciences and humanities (SSH) and the sciences. Waltman builds on the work of Hicks and Nederhof in describing large-scale differences between the natural sciences and SSH literature in publication and citation patterns, although the most comprehensive accounting of disciplinary difference in citation patterns is the chapter titled "Disciplinary Impact" in the Roemer and Borchardt volume *Meaningful Metrics.* Here, Roemer and Borchardt discuss differences in citation practice not only across larger genres like the natural sciences and SSH, but also provide detail on the specific citation practices within given fields.

As much as has been written about quantifying the impact of scholarship, nearly as much has been written about the problems this quantification brings about. Literature in many disciplines have pointed to the problematic nature of equating impact with citedness, and broadly influential publications such as *Nature,* *Science,* and *BMJ* have also weighed in. The San Francisco Declaration on Research Assessment (DORA) summarizes the problems with journal-level metrics in identifying the skewed nature of citation distributions within journals, the field-specific properties of measures like the journal impact factor (JIF), the ease with which impact metrics can be manipulated, and the proprietary nature of the data used to calculate metrics such as JIF. In an effort to address many of these problems, new metrics were introduced that uncoupled the prestige or impact of the article from the prestige or impact of the journal. With download counts, page views, and data from scholarly communities readily available, the article need no longer rely on the journal in which it was published in order to demonstrate its impact. Article-level metrics and altmetrics measure the interest in a digital scholarly work.

Altmetrics are a relatively new and developing supplement to traditional metrics. Perhaps due to their nascent, faculty familiarity and acceptance of altmetrics is relatively low. A study by Gruzd, Staves, and Wilk found that altmetrics were not accepted as measures of scholarly impact in the P&T process for Library and Information Studies (LIS) faculty, and Sutton, Miles, and Konkell found that, while most LIS faculty have at least heard of altmetrics, only a small number consider themselves experts. We see this trend not only in LIS, but across multiple disciplines. In a survey of faculty across all disciplines at the University of Vermont, DeSanto and Nichols found that a very small number of faculty respondents had a high level familiarity with altmetrics, but that most scholars seem to be at least somewhat familiar with them. Because altmetrics are a piece of the scholarly metrics conversation that scholars, especially junior faculty, have some familiarity with, we decided to include questions about altmetrics alongside our questions about traditional metrics in this study.

In addition to new methods of demonstrating impact, there are also evidenced shifts in how faculty in different disciplines are choosing to publish, cite, and demonstrate the value of their scholarship. Recent literature suggests that scholars in the social sciences are transitioning from publication patterns that at one time more closely resembled those in the humanities, to publication patterns more similar to those affiliated with the sciences. Studies have found that scholarship in the social sciences is becoming more collaborative, more widely indexed in resources like Web of Science, and more journal-centric. The University of Vermont study found that social sciences faculty now understand, value, and use scholarly metrics in the P&T process in ways that more closely align with the publication patterns of faculty in the sciences.
And yet, for all that has been written about numerical demonstrations of scholarly impact, librarians still commonly address questions from faculty about impact metrics and spend time supporting faculty in gathering impact data for the P&T process. The authors of a 2010 study in *Nature* asked faculty about the importance of impact metrics to hiring and promotion decisions. The study found that 70% of faculty believed that metrics were used in P&T decisions and that 63% of faculty were unhappy about the ways in which metrics were used. The authors followed up the poll with interviews of administrators that illustrated the often conflicting signals faculty receive about the importance of quantifying the impact of their scholarship. Similarly, our current study addresses faculty opinions regarding the use of impact metrics in the P&T process; however, this study does not address the hiring process or identify faculty administrators as separate from faculty.

In fact, very little has been written about how impact metrics are understood, used, and valued by faculty. As the literature on impact metrics has grown, the literature about faculty reception has remained virtually silent. The DeSanto and Nichols study asked faculty respondents about their familiarity with scholarly metrics (both journal- and article-level), their use of metrics, and the role of metrics in their department’s P&T process. The study pointed to similarities and differences across disciplines and confirmed that understanding and use of impact metrics are linked to perceived importance in the P&T process, although there remained a good deal of uncertainty amongst faculty of the role impact metrics actually play in their department’s promotion and tenure process.

Thuna and King also used a single-campus study to interview 79 senior research faculty who were more than five years post-tenure. Their study highlights disciplinary differences in faculty understanding and use of metrics, leading Thuna and King to advocate for more targeted disciplinary faculty outreach, saying “There appears to be a need to fill the gap between librarians and faculty researchers with examples of disciplinary best practices in the use of impact metrics.” Vinyard and Colvin also began their 2018 study with a desire to expand library outreach. They created a LibGuide to support faculty with their own publication efforts (e.g., providing information on the journal selection process) and conducted presentations with academic departments. They followed up on these presentations with a survey that asked faculty about their awareness and use of impact metrics, the results of which again underscore the need for further faculty outreach and education.

A limitation acknowledged in the DeSanto and Nichols, Thuna and King, and Vinyard and Colvin papers is the single-campus nature of the studies. This limitation prevents the authors from drawing conclusions past the boundaries and contexts of their own campuses and, without a larger body of literature from which to draw, it remains difficult to confidently make wider conclusions about faculty knowledge, opinion, and use of impact metrics. This study seeks to begin addressing those questions in a more robust multi-campus assessment. It builds on the survey work of DeSanto and Nichols and expands the assessment of faculty across two large universities (the Ohio State University and the University of Minnesota), one medium-sized university (the University of Vermont), and one smaller university (Valparaiso University). This study represents the first large-scale effort to assess faculty interaction with impact metrics and draw conclusions with greater certainty.

**Methods**

A previously developed survey was adapted and administered across three sites following slight modifications based on the authors’ experience and local context. The modified survey was pretested for face validity by librarians at each institution. The survey consisted of a combination of Likert scales, multiple choice responses, and open-ended questions. Questions sought to determine faculty perceptions of the utility of research impact metrics and the importance of metrics in promotion, tenure, and performance evaluations. We included only one question on altmetrics as the literature suggests that faculty have not widely adopted these new metrics. The
survey was administered electronically in November 2017 and was open to all faculty at the University of Minnesota, the Ohio State University, and Valparaiso University. Participants were recruited through email marketing campaigns. Responses from the previously administered survey at the University of Vermont were incorporated.

Institutional Review Board (IRB) approval was sought by each site, with each IRB finding the project to be exempt from IRB approval. Data were independently gathered at each site. The Ohio State University and the University of Minnesota administered the survey via Qualtrics, the University of Vermont used Limesurvey, and Valparaiso University administered the survey using Survey Monkey. All data were extracted as comma-separated values files to ensure interoperability. All direct identifiers were removed prior to the consolidation of responses.

Data were analyzed using R 3.4.2. Likert scales were considered continuous variables and were analyzed using analysis of variance models (ANOVAs) followed by Tukey’s post-hoc analysis. Categorical variables, such as yes and no responses, were analyzed using Chi-square analyses. P-values of less than 0.05 were considered to be significant.

**Overall Results**

There were 996 new responses from the three additional institutions, in addition to the original 206 responses collected in the previous iteration of the survey. The authors grouped responses into four broad disciplinary areas of Health Sciences (n=444), Sciences (n=343), Social Sciences (n=256), and Arts & Humanities (A&H) (n=158).

The responses by rank were Assistant Professor (n=342), Associate Professor (n=376), and Professor (n=484). Other ranks and appointments were not considered as they were not consistent across all four institutions.

**Metric Familiarity and Accuracy**

**Scholarly Metrics: Familiarity by Discipline and Rank**

Faculty in Social Sciences, Sciences, and Health Sciences departments all indicated a statistically significant higher level of familiarity than faculty in the A&H departments across all four institutions. Each department grouping averaged similarly higher scoring over A&H, with each group averaging 1.25 points or more than their A&H peers ($p<0.01$). These data suggest that A&H faculty are not as familiar with scholarly metrics as the rest of the academy.

There was no statistically significant difference by respondent faculty rank in reported familiarity with scholarly metrics ($p=0.20$).

*The University of Vermont survey included participants from the College of Nursing and Health Sciences but did not include faculty from the College of Medicine.*
Scholarly Metrics: Perceived Accuracy by Discipline and Rank

In a similar fashion to scholarly metric familiarity, there were statistically significant differences in the perceived accuracy of metrics between A&H and the other groups. Health Sciences averaged 0.92 points higher than A&H ($p<0.01$), Social Sciences averaged 0.85 points higher than A&H ($p<0.01$), and Sciences averaged 0.79 points higher than A&H ($p<0.01$).

There were no statistically significant differences between Sciences and Health Sciences, Social Sciences and Health Sciences, or Social Sciences and Sciences.

There was no statistically significant difference by respondent faculty rank in reported perceived accuracy of scholarly metrics ($p=0.09$).

Altmetrics: Familiarity by Discipline and Rank

While the majority of Science, Health Science, and Social Science faculty reported a high level of familiarity with traditional scholarly metrics, they were not as familiar with altmetrics. Social Sciences reported a higher familiarity with altmetrics, with 23.4% of respondents being either familiar or extremely familiar with altmetrics, when compared with Sciences (16.5%) and Health Sciences (16.2%). A&H was the only group reporting a majority for “Not At All Familiar” (55.1%).

There was no statistically significant difference in reported familiarity with altmetrics based on respondent faculty rank ($p=0.41$)
Metric Importance and Inclusion
When faculty were asked about the perceived importance of metrics in P&T processes, there were again statistically significant differences between A&H responses and all other broad disciplines. When compared to A&H, Health Sciences averaged 1.27 points higher in perceived importance ($p<0.01$), Social Sciences averaged 1.14 points higher ($p<0.01$), and the Sciences averaged 0.96 points higher ($p<0.01$).

When comparing Health Sciences, Sciences, and Social Sciences, only one statistically significant difference occurs: Health Sciences averaged a higher score than Sciences (0.32 points higher; $p=0.01$). There were no statistically significant differences between Health Sciences and Social Sciences, or Social Sciences and Sciences.

When considering perceived metric importance by rank as part of the P&T process, there were several statistically significant differences. Overall, Full Professors indicated the lowest value of perceived importance, while Assistant Professors indicated the highest value. When compared to Assistant Professors and Associate Professors, Full Professors averaged 0.79 ($p<0.01$) and 0.24 ($p=0.04$) points lower, respectively. Additionally, Associate Professors averaged 0.55 points lower than Assistant Professors ($p<0.01$).
When faculty were asked to gauge how much weight should be placed on scholarly metrics in P&T, Health Sciences, Sciences, and Social Sciences faculty all reported more weight than A&H faculty members; Health Sciences were 0.57 points higher \((p<0.01)\), Sciences were 0.43 points higher \((p<0.01)\), and Social Sciences were 0.48 points higher \((p<0.01)\).

Health Sciences also reported a statistically significant higher score (0.14 higher, \(p=0.02\)) than Sciences. No other statistically significant differences were present in the data.

Faculty members were also asked if their respective departments encourage or require inclusion of scholarly metrics in P&T documentation. Most A&H faculty members reported that they were not encouraged to include scholarly metrics in their P&T paperwork. Most faculty in the other disciplinary groupings reported that they were encouraged to include scholarly metrics.
When asked about whether inclusion was required, A&H faculty overwhelmingly reported that they were not required to include scholarly metrics in their P&T documentation, and the Sciences tend not to require the inclusion of scholarly metrics (figure 8). However, 33.2% of the total number of respondents did indicate that their departments do make metrics a requirement. Approximately equal numbers of Social Science faculty felt that their departments required and did not require including metrics in their P&T documentation, while the Health Sciences favor the requirement of scholarly metrics in P&T documentation. There were also surprisingly high rates of respondents from all disciplines who did not know if their departments required the inclusion of scholarly metrics.
It should be noted that there was significant overlap between respondents who indicated their departments required metrics and those who indicated their departments encouraged metrics. 98.2 percent of individuals whose departments required metrics also noted that the use of metrics was encouraged by their departments.

Analyzing the results by faculty rank, we saw that a larger portion of Assistant Professors reported that they did not know whether scholarly metrics were encouraged or required (figures 9 and 10).

**Discussion**

Throughout the survey results, A&H faculty proved to perceive, value, and use scholarly metrics very differently than their colleagues in the Sciences, Health Sciences, and Social Sciences. A&H faculty members reported less familiarity with both traditional metrics and altmetrics. We suspect that their unfamiliarity stems from the fact that publishing patterns in A&H disciplines tend to rely more heavily on monographs and creative works. Monographs and creative works are not represented by traditional journal-level metrics, and faculty creating work in these formats naturally would be less familiar with statistical representations of impact and less likely to view statistical metrics as an accurate representation of their scholarly output. A&H faculty also reported that P&T processes largely do not encourage or require the inclusion of scholarly metrics. This is contrasted with Sciences, Health Sciences, and Social Sciences, where faculty reported that statistical measures of impact are used during the P&T processes.

Results on questions of P&T also pointed to an inverse relationship between faculty rank and perceived importance of impact metrics. Assistant Professors placed the greatest amount of importance on impact metrics, while Full Professors placed the least amount of importance. Certainly, Assistant Professors preparing for the tenure process are seeking out ways to demonstrate the impact of their work and are often under the impression they should include statistical measures of impact as part of their tenure materials. More well-established Full Professors may not feel as much pressure to engage with statistical representations of their work’s impact given that they are no longer assessed for promotion. Full Professors may also have enough experience having gone through two P&T processes and likely having been called upon to assess the work of others to understand the limitations of statistical impact measures. They will certainly have a better understanding of what is valued in their own departmental P&T process. More follow-up work could be done with faculty to explore the reasons behind these differing views of metrics. Whether it is due to immediacy of need, better understanding of the limits of statistical impact measures, or more knowledge about the actual workings of the P&T process, our study indicates that there may be a relationship between a faculty member’s rank and the importance they place on scholarly metrics in the P&T process.
Our study also found that faculty are not nearly as familiar with altmetrics as those in the LIS field might have assumed, given the attention in LIS literature and many libraries’ outreach efforts to faculty. Many more faculty reported being “not at all” or “marginally” familiar with altmetrics than reported being “familiar” or “extremely familiar.” This finding serves as an important reminder that while altmetrics may no longer be a new concept in our field, it remains an unfamiliar concept to most faculty.

Finally, although there was some variance between responses in the Health Sciences, Sciences, and Social Sciences, in general these disciplinary areas tended to respond similarly to questions of familiarity with traditional metrics, perceptions of metrics’ accuracy, and the desired and perceived role of metrics in the P&T process. The similarity of the Social Sciences, Sciences, and Health Sciences confirms previous single campus studies showing that researchers in the Social Sciences may have shifted away from modes of demonstrating impact akin to A&H and may instead be choosing to demonstrate the impact of their scholarship in ways more aligned with the Sciences and Health Sciences. Further exploration is needed to determine if this finding could be extended to show that scholars in the Social Sciences are demonstrating impact in ways that more closely align with the Sciences and Health Sciences because they are choosing to publish and communicate their research in ways that align more closely with the other sciences.

Limitations and Future Research

There are some limitations to the data presented in this paper. As mentioned in the methodology section, the disciplinary divisions (Social Sciences, Health Sciences, Sciences, and A&H) are different from the original study from the University of Vermont (Social Sciences, Sciences, and A&H). Additionally, the University of Vermont study did not collect data from all of its Health Science units as part of its “Sciences” data set.

The authors plan to examine differences between their participating institutions more closely. In addition, any future iterations of this survey should include more institutions of various types/sizes. Only one smaller, liberal arts-focused institution (Valparaiso University) and one mid-sized research institution (University of Vermont) were included in this study, so certain aspects of those institutions’ survey data may not be generalizable to similar institutions. Future iterations of this survey should include more institutional diversity, especially size and mission.

In addition to the data presented in this paper, the survey also collected data related to faculty use of scholarly metrics tools and need for support. The authors plan to analyze this data in a future publication and to explore how libraries could support the use of these tools.

Conclusion

This study represents the first multi-campus attempt to address how faculty members understand, perceive, and employ statistical measures of scholarly impact. The findings extend previous single-campus studies in concluding that social science faculty now demonstrate the impact of their scholarship in ways that closely resemble faculty in other science disciplines. Likewise, social science faculty mirror faculty in other sciences in how they understand and value impact metrics. The study also reinforces previous works that found that altmetrics remain unfamiliar to a majority of faculty. These previous findings can now be claimed with greater certainty given the number of faculty respondents in this survey across four different universities.

New findings in this study point to an inverse relationship between faculty rank and the perceived importance of scholarly metrics to the P&T process. Assistant Professors responded with the greatest deal of perceived importance and Full Professors responded with the lowest amount of perceived importance to P&T. This finding presents an opportunity for more research into perceptions of pre-tenure and post-tenure faculty regarding the role that impact metrics play in the P&T process.
While much can be said anecdotally about how scholars in different disciplines and at different ranks think about and use impact metrics, up until this point, there has been little evidence in the literature. We can now draw some conclusions from this multi-site study and its alignment with other single-campus studies. In our work, we see a picture of disciplinary shifts in the social sciences, a confirmation of anecdotal ideas about demonstrating impact in the arts and humanities, a possible over-estimation on the part of librarians about the familiarity of faculty with altmetrics, and a surprising trend in the perceived valuation of impact metrics as faculty members proceed through the academic ranks.

Notes
27. Amy M. Suiter and Heather Lea Mouaison, “Supporting Scholars: An Analysis of Academic Library Websites’ Documenta-


33. DeSanto and Nichols, ”Scholarly Metrics Baseline.”

34. DeSanto and Nichols.

35. Sutton, Miles, and Konkel, ”Awareness of Altmetrics.”

36. DeSanto and Nichols, ”Scholarly Metrics Baseline.”

37. DeSanto and Nichols.

38. Vinyard and Colvin, ”How Research becomes Impact:”

### Bibliography


López-Cózar, Emilio Delgado, Nicolás Robinson-García, and Daniel Torres-Salinas. ”Science Communication: Flawed Citation Indexing.” *Science* 342, no. 6163 (December 6, 2013): 1169. https://doi.org/10.1126/science.342.6163.1169-b.


