Qualitative Analysis of Faculty Opinions on and Perceptions of Research Impact Metrics

Caitlin Bakker, Kristen Cooper, Allison Langham-Putrow, and Jenny McBurney

We present a qualitative analysis of the results of a survey of faculty and researchers at a large Midwestern R1 university around their understanding of and attitudes toward scholarly metrics. The survey included opportunities for participants to provide free-text responses regarding their use of metrics and concerns they have about the use of metrics for assessment. Participants indicated they understand metrics and use them in a variety of ways, but they have concerns about administrators' potentially inappropriate use of metrics in assessment. Participants expressed a desire to be involved in decision making around the use of metrics in evaluation processes. With the end goal of improving our library's research impact–related services to better support faculty and researchers across campus, this exploratory qualitative analysis offers a more nuanced understanding of the current landscape of opinion around research impact metrics. To develop tools and services that actually address faculty and researcher needs, librarians must develop a comprehensive understanding of their interests and concerns around metrics.

Introduction

Researchers and their institutions are increasingly called upon by funders, legislators, and other stakeholders to demonstrate their productivity and the subsequent impact of their research, both in the scientific community and in society. Research impact metrics, including traditional bibliometrics, have been one mechanism for assessing the quality of research for more than 60 years.[1] Bibliometrics are "a set of quantitative methods used to measure, track, and analyze print based literature" and are used by individual scholars, institutions, and funding agencies to measure the impact of research and scholarship.[2] Traditional metrics mainly focus on how often an article is cited by other scholarly articles; examples include the h-index, a measure of an author's quantity of publications and how many times they have been cited, and the Journal Impact Factor, a measure of how often a journal's articles are cited.[3]

With the dissemination of scholarly work shifting from exclusively print to electronic formats that can be shared and accessed online, a new form of metrics, "altmetrics," has been an increasing topic of conversation. Altmetrics are "new metrics based on the social web for analyzing and informing scholarship."[4] Examples of altmetrics include the number of times a link to a work has been clicked on or the number of likes, shares, and mentions on platforms such as Twitter or Facebook.[5]

Despite the growing conversation around altmetrics, researcher use of these resources remains unproven, and traditional metrics have begun playing an increasing role in research decision making, particularly outside North America. Reasons for this include the evaluation of federal or public money spent on higher education and research based on its quality and impact, as well as steps institutions have taken to develop strategies for research in the face of competition with peer institutions for students, staff, and resources.[6] Examples of the use of metrics in decision-making processes can be found across the globe. Smith, Crookes, and Crookes note that, in Australia, traditional metrics, including Journal Impact Factors, citation rates, and the h-index, are frequently used for measuring research impact.[7] Wilsdon et al. provide examples of several other countries that incorporate metrics in their national research assessment programs, such as Denmark, Italy, and the Netherlands;[8] however, there have been recent signals of movement to deemphasize the role of metrics in research evaluation.[9] The United States differs in that there is no nationwide research assessment system or program,[10] nor is there a single funding body that mandates the use of a particular metric or metrics. As Graham et al. note, "competing interests among affected stakeholders can result in a lack of consensus on what constitutes value and what should be measured in order to demonstrate impact."[11] Regardless of this lack of a nationwide strategy, research impact assessment is an emerging area of interest among American institutions and funding agencies.[12]

Literature Review

The role of metrics in decision making is the subject of debate, and perceptions of the importance and value of metrics differ across institutions, disciplines, departments, and individuals. In interviews by Abbott et al., some administrators stated that they do not consider metrics at all in hiring or promotion and tenure (P&T) decisions, choosing instead to rely on the letters of recommendation provided by experts in the candidate's field, with one respondent stating that he does not rely on impact factors, which "usually highlight trendy papers, boom fields and recently highlighted topics. We…don't want to follow boom."[13] Another respondent noted that, although his department collects data on teaching loads, output of papers, and h-indices, the data is used to guide researchers and is "not a hurdle that has to be leapt over to get a promotion." However, the authors acknowledge that the collection of these measures "could give the impression that they are being relied on heavily."[14] Supporting this comment from Abbott et al., DeSanto and Nichols also found that "[a] full 68 percent of respondents [to a faculty survey] expressed concern about university administrators tracking the scholarly metric data of their faculty"[15] and that only 5.4 percent of respondents thought that "a great deal of weight" should be placed on metrics as part of the P&T process.[16]

In contrast to the administrators' responses, faculty survey responses from the article by Abbott et al.[17] show that researchers believe metrics have a large impact on hiring and P&T decisions. More than 70 percent of 150 respondents believed that metrics were used in hiring and promotion decisions, and almost 70 percent also believed they were used in tenure decisions.[18]

In addition, 63 percent of respondents said that, overall, they were either "[n]ot satisfied" or "not very satisfied" with the way metrics are used in general, while only a quarter said they were satisfied.[19] Other concerns identified in Abbott et al. included the ability for researchers to manipulate metrics for their own gain, a concern shared by 71 percent of respondents, and the concern that metrics would shape the research behaviors of faculty rather than the other way around. In fact, half of respondents stated that they themselves had changed their behaviors, though often only in small ways, to improve the metrics they knew were used to measure themselves.[20] Aligning with findings from Abbott et al., Thuna and King found that metrics have influenced faculty across disciplines in their research choices, such as where to publish, who to hire, or when applying for or reviewing grants.[21]

More recently, there has been a growing interest in considering the disciplinary differences in faculty's use of and attitudes toward research impact metrics. Faculty in the sciences and social sciences have generally expressed greater awareness of and interest in metrics and have felt that metrics played a more significant role in P&T processes than their colleagues in the arts and humanities.[22] Researchers across disciplines have also expressed increasing awareness of altmetrics.[23] However, when considering faculty awareness or familiarity with metrics, as Thuna and King noted, "awareness does not necessarily equal understanding."[24]

Librarians have long been acquainted with the Journal Impact Factor and the various citation indices through which metrics are available. Additional metrics-related services are an emerging area in librarianship, as evidenced in the recent survey from Gutzman et al. of seven health sciences libraries,[25] which reinforces findings from the 2015 ACRL SPEC Kit on Scholarly Output Assessment Activities.[26] However, a 2015 Ithaka S+R Faculty Survey found that less than 20 percent of respondents have the library assist them with assessing the impact of their publications,[27] so there is still room for growth in this area. Development and refinement of library services requires an in-depth understanding of user needs.

Bakker et al. quantitatively compared respondents across three broad disciplinary areas.[28] They found that respondents in the Arts and Humanities were less familiar with metrics and perceived metrics to be less accurate than respondents in Social Sciences and those in the Sciences and Health Sciences. Researchers in the Arts and Humanities also felt that metrics were less important in their promotion and tenure and annual review processes, and they desired that metrics would hold less weight in these processes than researchers in other discipline areas. While these quantitative data provided a broad overview of faculty attitudes and perceptions of the importance of metrics, they do not provide insight into why the respondents may have felt this way, or what related concerns or opportunities they saw in this area. Although quantitative assessments of faculty attitudes have been conducted, qualitative assessment has been less prevalent in this area. In this paper, we provide a qualitative analysis of survey data gathered to describe faculty and researcher use of and attitudes toward research impact metrics and their concerns regarding the use of these measures. We chose to focus our qualitative analysis on our institution to develop a more in-depth understanding of researcher attitudes within our context and to subsequently gain insight necessary to begin developing services to meet their needs.

It should be noted that the nature of qualitative research reflects a different framing of the research question, and these questions often tend to be more exploratory than hypothesis-driven.[29] As Corbin and Strauss note, "underlying the use of qualitative methods is the assumption that all of the concepts pertaining to a given phenomenon have not been identi-

fied, or aren't fully developed, or are poorly understood and further exploration on a topic is necessary to increase understanding."[30] This is particularly true when researchers take a grounded theory approach, as we have here, in which "[o]ne does not begin with a theory, then prove it. Rather, one begins with an area of study and what is relevant to that area is allowed to emerge."[31] Our objective in this research is not to conduct analyses in line with quantitative approaches, but instead to begin an exploration of the nuanced interpretation and perception of research impact assessment from a researcher's perspective.

Methodology

With the end goal of improving our library's research impact–related services to better support faculty and researchers across campus, we analyzed data gathered from a survey of faculty, instructors, and researchers at the University of Minnesota to gain insights into how they understand and view research impact metrics, such as those based on article citation counts. The survey included open-ended, free-text responses in which participants described when and how they use metrics and their concerns around the use of these metrics. While previous surveys have provided quantitative analysis focusing on researcher awareness of these topics,[32] this grounded theory approach to qualitative analysis provides deeper insight into researcher perceptions and opinions around these topics and offers a more nuanced understanding of the current landscape of opinion around research impact metrics.

The qualitative study described in this paper was part of a larger multisite research project authored by Bakker et al. that involved administering a survey regarding attitudes toward and use of research impact metrics across four institutions.[33] The broader, multisite study focused on the analysis of quantitative data from all four institutions, while this paper addresses qualitative aspects derived from the previously unexamined open-ended survey questions on beliefs, concerns, perceptions, and use of research metrics, focusing on the data from our institution.

Survey questions were based on those first developed by researchers at the University of Vermont with slight modifications at each institution to reflect local context.[34] Two versions of the survey were used locally, one for tenure-track faculty that referred to the tenure and promotion process and one for non–tenure-track faculty and researchers that referred to the annual review process. Appendix A contains the version of the survey for tenure-track faculty. Modifications were pretested for face validity by librarians across the four institutions involved in the multisite project. Survey responses from our institution were collected via Qualtrics between November 7 and December 8, 2017.

Survey participants were selected through convenience sampling. A list of all current researchers, including administrators, tenured, tenure-track, and nontenured faculty, instructors, lecturers, and research fellows was generated from human resources data. Participants were invited via email, and a reminder notice was distributed via email two weeks prior to survey closure. A total of 4,855 individuals were invited to participate, of whom 435 responded to at least one open-ended question. Despite what may appear to be a low response rate when viewed through the lens of quantitative research, it is important to recognize that, in this qualitative methodology, "sample adequacy, data quality, and variability of relevant events are often more important than the number of participants."[35] Thus, the sample size obtained was deemed to be acceptable. The survey began with an information page describing the purpose of the project and the voluntary nature of participation. The project was submitted for IRB approval and was determined not to be human subjects research.

Raw data were exported from Qualtrics, direct identifiers were removed, and the data were stored in the university's instance of Box. Access to raw data was limited to one researcher. De-identified data were made available to members of the research team through a separate Box folder. To protect participant privacy, departments and job classifications with small numbers of participants were collapsed into broader categories, and responses to open-ended questions were extracted as Microsoft Word documents and made available to the research team.

NVivo 12 was used to analyze these data and the analysis was based on the principles of grounded theory.[36] Three researchers independently coded 10 responses using line-by-line open coding. This led to the development of a coding scheme (see appendix B). Researchers independently coded all responses, and the NVivo databases were merged. This triangulation of investigators and sources provided multiple perspectives on the data, creating a richer interpretation of data reflecting a broader range of experiences, thereby strengthening confidence in conclusions derived through this research.[37] NVivo's Coding Comparison function was used to determine interrater reliability and, where the kappa coefficient was less than 0.4, codes were revisited until consensus was reached. The researchers then met to identify themes emerging from the data through the identification of recurring concepts and relationships between the codes, which resulted in the theoretical framework described in this paper.

Results
Participants described complex relationships with research impact metrics, simultaneously engaging with them in evaluation practices while expressing significant concerns regarding their use and meaning. Seven themes that describe participants' attitudes toward, use of, and concerns regarding these measures were identified: (1) Disciplinary awareness is key when considering research impact metrics; (2) Metrics are used in information-seeking activities and as a means of self-assessment; (3) Metrics are used when evaluating the work of others; (4) Metrics should not be considered a proxy for the full range of researcher impact; (5) Administrator use of metrics in researcher evaluation is a concern; (6) Inappropriate use of metrics could potentially result in negative consequences; (7) Shared decision-making regarding metrics is necessary. We describe each of these themes in detail below.

Disciplinary awareness is key when considering research impact metrics
Participants strongly articulated the need for administrators and decision makers to have an in-depth understanding of the discipline to appropriately contextualize and interpret metrics. For example, citation patterns vary widely between disciplines and can vary dramatically even between subdisciplines. As one participant stated, "[l]ike most aspects of reviews and evaluations, context is very important when determining overall value." Participants felt that, without having the context surrounding a number, it is possible to undervalue certain areas of research. As one respondent noted, "[a] high quality but specialized paper may be cited less than a poorer quality but more general (or controversial) one."

A number of participants reported that they do not have concerns with metrics per se, but placed caveats on when and how they should be used, such as not using a single metric or relying solely on metrics without conducting a qualitative evaluation of one's work. As one participant noted, "…my work is interdisciplinary (on the borders of the humanities and social sciences) but it is only 'tracked' in this fashion in medical and scientific journals. This… is completely skewed and misleading."

Metrics are used in information-seeking activities and as a means of self-assessment

Metrics serve as a data point for participants in their own decision-making processes relating to their work. Deciding upon a journal for manuscript submission was a frequently noted use case for metrics. However, participants also monitored citations to their work as a means of determining interest: "I look at my metrics every time I publish something new, to see what is getting traction." One participant described this as being "particularly useful because I'm working on a topic that doesn't sit neatly within any specific discipline, so it would otherwise be hard to see how my work is traveling and to discover connections that stimulate my own research and allow it to travel even farther."

Participants expressed a desire to better understand who is interested in their research area rather than the direct impact of citation on specific metrics, noting "the metrics just come along for the ride." Participants noted that they would consult metrics sources when prompted by an external source, such as an email notification, out of curiosity, or when necessary for P&T purposes.

Metrics are used when evaluating the work of others

Participants reported using research impact metrics for a variety of reasons in the assessment of other individuals, including when making hiring decisions, serving on P&T committees, and writing letters of support for colleagues. Beyond factoring into the hiring process, metrics were described as an element of recruitment, with one scholar noting that "I check [the] scholarly outputs of people we might think about encouraging to apply to a job in our department." Although assessment of potential hires was a frequently mentioned reason for consulting metrics, some participants described hesitation, indicating that "[w]hile scholarly metrics are a useful tool in a casual sense, I'm not very comfortable with hiring and promotion decisions being made based on them."

Participants engaged with metrics to vet others' research, particularly "[w]hen I'm reading a paper far outside my field and I do not know the quality of the journal the work appears in, and I am not necessarily qualified to see potential flaws in the work." However, they were also careful to note that they used these metrics in context; as one participant said, metrics should be used "[a]s one of many filters to assess the importance of a researcher's work…."

Metrics should not be considered a proxy for the full range of researcher impact

Many participants indicated that they use metrics for self-advocacy, including promotion, tenure, performance evaluations, and salary negotiation. Despite the use of metrics for these purposes, participants felt that these were not robust representations of their impact. One noted that "we are a highly productive research center, and our work is used every day all over the country, and the reason for this [is because] it didn't get hidden away in peer reviewed journals—we turned it in to training, outreach, and practitioner skills that can be used to help people." Another participant echoed this sentiment, describing other valuable work they produce, such as "community-friendly dissemination pieces (technical reports, white papers) that are available to our community partners immediately and don't take 2 years to publish in a journal that no one but academics read."

Beyond the need to have a broad view of impact, participants emphasized the diverse roles faculty and researchers play within the university, reflecting on their teaching and service

responsibilities and the importance of balancing these responsibilities: "Encouraging me to move publication to my primary role is also requiring me to put education as my secondary role." One participant encouraged administrators to "keep a 'whole' view of the faculty the same way we use a 'whole' view when considering student applications to the university."

Administrator use of metrics in researcher evaluation is a concern

Participants expressed concern regarding the depth of understanding that others, particularly administrators, possess in their understanding of disciplinary differences and how that might influence their ability to use metrics appropriately. As one participant said, "[t]o the extent that university administrators lack a sufficiently nuanced understanding of each field and subfield on campus, I would be wary that their evaluation of such data likewise would be insufficiently nuanced." Participants were also concerned about what this means for evaluation, noting "[i]f metrics are being relied upon to track scholarly work, it means the people doing the tracking don't understand the work enough to judge it, and therefore shouldn't be. That is the fundamental problem."

There was concern over the weighting of metrics within decision-making processes. One participant described the potential for "an over reliance on such metrics without careful and thoughtful consideration of differences among disciplines and the details of a specific scholar's career." Participants noted that metrics can offer "a false sense of objectivity," and that "people can get seduced by how clean and easy [numbers] are to use and forget the messy complexity that underlies them and what they are supposed to convey." Participants expressed the desire for administrators to engage on a deeper level with scholarly outputs, saying "[a]dministrators should read the papers themselves or discuss the importance and context of the papers with the researchers themselves."

Inappropriate use of metrics could potentially result in negative consequences

Participants described possible negative consequences that could be associated with the misapplication of metrics, both intentional and unintentional. For example, one researcher worried that administrators might apply metrics in such a way to justify not giving raises, making cuts to benefits, or "getting rid of faculty they don't like." Another described the potential for further-reaching consequences: "[p]unitive decisions might be made not only about my own career, but about my academic unit, program, college, research centers and institutes relevant for my work, etc."

Participants noted the potential effects on the direction and focus of research more broadly, particularly if metrics were to be heavily weighted. One participant compared it to unintended consequences of an increased focus on test scores in education, stating that:

> "We should be motivated to do what we think is good science, not what we think is going to get a lot of citations very quickly or, God forbid, a lot of twitter mentions. We should not ignore metrics, but giving them too much power would put us at risk of becoming like the elementary school teachers that feel they can only focus on helping their students perform on standardized tests, or the newspaper editor who eschews the important (if not exciting) news story in favor of 'click-bait'."

Others spoke about the possibility of "gaming the system," that researchers might change or feel pressure to change the direction of their research to improve their metrics. As one participant stated, any measure of success "will become the target of gamesmanship and practices will be changed to create a better score." Participants felt that a focus on metrics as the basis for evaluation may result in researchers feeling discouraged from pursuing particular lines of inquiry that might have greater impact on the field or broader impacts on society but may not result in highly cited publications.

Shared decision-making regarding metrics is necessary
Participants expressed a desire to be involved in the decision-making processes surrounding the use of metrics. Rather than a top-down approach to determining which metrics to apply and in what situations, participants advocated for a more collaborative model, as they "would want to ensure administrators had technical assistance/support, [and] shared [the metrics] with faculty to assess veracity."

Participants recognized that metrics are increasingly common and potentially beneficial. One remarked:

> "Although I want scholarly metrics to be used by administrators to track research productivity, I only want that done after a thorough vetting of the issues by the faculty. The faculty who will be judged should be involved in designing the system and metrics by which they will be judged. But they have to come up with some metrics. Saying 'we can't be judged objectively' lacks credibility as far as I'm concerned."

Participants also noted that through their involvement in the decision-making process, they would be able to identify the metrics that are most appropriate for use within their field or subfield: "[I] would hope faculty would have the chance to propose and justify [the] use of the metrics best aligned with their work rather than having one or two forced upon all types of research."

Discussion
This study was conducted to explore faculty and researcher views of the use of scholarly metrics to inform librarians at our institution on how to better provide support. Overall, participants expressed the belief that they are able to appropriately use metrics to assess others and others' work, to make decisions about where to share their work, and to advocate for themselves; and that to appropriately use metrics requires an understanding of disciplinary differences that those outside a subject area (such as administrators) may not have.

Researcher concerns regarding the use of research impact metrics by administrators largely focused on the need for a nuanced understanding of the data points, including appropriate selection of measures, and the potential negative consequences of application without this understanding. Researchers' desire to be involved in the conversation surrounding the use of metrics reflects the importance of using discipline-sensitive measures as a component of a holistic assessment of productivity and impact. Echoing sentiments found in previous research,[38] the researchers' desire to be actively engaged in decision making, to position metrics as one of many data points, and to consider the robust nature of a faculty member's

role reflects the need to avoid reductive appraisal of faculty or research. Researchers were rightfully concerned that a single measure, broadly deployed and without context, would not accurately reflect their diverse research areas and outputs. Given the role of administrators in influencing career trajectories and allocating resources, it is not surprising that the prospect of inappropriate assessment methods would be of great concern. Administrators who may be using metrics in decision-making processes should be transparent about which measures are being used and the weight of those measures within the process. Care should be taken to ensure that a range of qualitative and quantitative measures are employed in any assessment or evaluation.

The potential simplification of impact and the possibility of inappropriate use by administrators was connected to larger questions surrounding the direction and value of research. Similar to what was found by Abbott et al., researchers expressed concerns that such an incentive system may influence researchers to redirect their efforts toward areas of study that would be better served by such measures.[39] Although there was some speculation regarding the potential for "gaming the system" or other types of metric manipulation, researchers also reflected that this could be a larger issue in that such an incentive system could change the motivations for doing research and, in doing so, could have the potential for negative long-term ramifications as individuals pursue rewards rather than discovery. It is imperative to recognize that university incentive systems have the potential to influence research and publication practices of individuals and, in turn, to influence departmental and disciplinary culture and focus. The interconnection between incentive systems, publication choices, and metrics should be foregrounded in conversations regarding if, how, or when metrics are being employed by administrators, funding agencies, and other stakeholders.

Although participants described their misgivings regarding the use of metrics in the assessment of faculty and researchers, they nevertheless employed these measures in their assessment of journals and the works of others. This echoes the findings of Thuna and King, who found that researchers in a variety of disciplines consider journal impact when selecting publication venues and research impact metrics when serving on hiring and P&T committees.[40] Among our participants, journal impact in particular was often considered a proxy for journal quality and a key component when assessing article quality, particularly in new or unfamiliar areas of study. This may indicate that, although researchers recognize the need for nuance in the interpretation of author-level metrics, such a recognition has not been fully transferred to journal assessment, despite the well-established concerns regarding gaming of Journal Impact Factor.[41] This creates an opportunity for libraries to provide education and outreach on journal assessment strategies, both for one's own work and for others' research. Liaison librarians as subject specialists have an appropriate blend of disciplinary knowledge and understanding of publication practices to provide this support and insight to their departments.

When considering their own research and research areas, participants used metrics as a means of discovery, in tracing work similar to their own through citation alerts and in keeping track of researchers in their disciplines. This reflects sentiments found in Thuna and King and in DeSanto and Nichols.[42] The metrics themselves are not the primary endpoint in these activities; instead, they are a mechanism through which researchers better understand and conceptualize their research networks.

Participants questioned how close of a proxy metrics are for the true impact of a researcher's work. Indeed, when a researcher's primary aim is to influence community behaviors,

improve practice, or teach more effectively, success in these endeavors is largely not indicated through citation-based metrics as they are producing content such as community-friendly dissemination pieces or teaching and outreach materials instead of scholarly articles. Despite these concerns, findings from the quantitative data of the larger multisite study made by Bakker et al. showed that a minority of participants were aware of altmetrics.[43] This agrees with findings from DeSanto and Nichols and also from Thuna and King about the low levels of familiarity or adoption of altmetrics by faculty.[44] This disconnect between the need for a broader representation of impact and low awareness of mechanisms to describe broader impact may be an area for future outreach activities and service expansion. Institutions and administrators may wish to consider how nonarticle research outputs are acknowledged and incentivized in the evaluation process and to ensure that the full scope of a faculty member's activities are represented in these processes.

Concerns expressed were often focused on the potential use of single measures across disciplines, although that scenario was not referenced directly or alluded to within the survey questions. The immediate association of research impact metrics with the use of a single metric is one that seems to cause significant fear among researchers. The recontextualization of metrics as a suite of data points, each of which may be deemed appropriate or inappropriate in certain contexts, is necessary when considering the development of a productive dialogue on these topics. Libraries are well-positioned to inform researchers and administrators on the appropriate use of metrics. The acquisition and use of these measures are an information literacy issue in that the effective, ethical use of these measures requires that an individual understand what these measures do and do not indicate and how they can be appropriately interpreted.

Study Limitations

This study is limited due to the number of respondents and the single-university study environment. Roughly 9 percent of the 4,855 individuals to whom the survey was sent responded to at least one open-ended question, and individuals who chose to take the time to respond to both the survey and provide text responses may have characteristics different from individuals who did not respond. However, sample sizes in qualitative research are judged differently than in quantitative analysis.[45] The 435 participants, while representing only a limited portion of the overall possible respondents, provided robust data of sufficient depth to achieve saturation.

Similarly, we recognize that researchers based at an R1 land grant institution in the United States may have different perspectives, contexts, and experiences than researchers based at other types of institutions or in other geographic areas. The issue of generalizability in qualitative research remains a contentious one and has been discussed in numerous venues.[46] Although our study may not be generalizable to a broad population, our intention is to provide an in-depth analysis of participants' thoughts and perceptions surrounding this phenomenon, which can be expanded upon in the development, implementation, and assessment of library services.

Conclusion

When considering the use of metrics by others (such as administrators) to assess their research, survey participants expressed that sufficient disciplinary and subdisciplinary knowledge combined with a collaborative and transparent approach to the selection of measures are

necessary. We found significant concern among participants that the inappropriate use of metrics, or the use of a single metric, in the evaluation of individual researchers would not only disadvantage individuals but would also have negative consequences for departments and disciplines.

Researchers described a complicated relationship with research impact metrics. Although they expressed the need for deep disciplinary knowledge when applying metrics, researchers nevertheless reported feeling confident in their own use of metrics to evaluate the work of others and the quality of journals, both in their own and other disciplines.

Given the complex landscape of multiple and often conflicting data sources, emerging measures, and potentially high stakes, it is unsurprising that researchers experience challenges with metrics. Libraries are well-positioned to support researchers and administrators in understanding the nuances of research impact metrics. However, to effectively provide this support, libraries must have a robust understanding of the researchers' knowledge, practices, and culture around research impact metrics. This study provides a view across the landscape that can be used to tailor services to address faculty and researcher concerns.

APPENDIX A. Survey

This appendix contains the full text of the survey. This paper analyzes the responses to questions 14–16. See Bakker et al. for analysis of the quantitative questions (questions 1–13):

Bakker, C. et al. 2019. "How Faculty Demonstrate Impact: A Multi-Institutional Study of Faculty Understandings, Perceptions, and Strategies Regarding Impact Metrics." In *ACRL 2019 Proceedings: Association of College and Research Libraries, Cleveland, Ohio, April 10-13, 2019*, 556–568. www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2019/HowFacultyDemonstrateImpact.pdf.

1. In what discipline would you place your research?
    - □ Sciences and Health Sciences
    - □ Arts and Humanities
    - □ Social Sciences, Business, and Social Services

2. How familiar are you with scholarly metrics (Journal Impact Factor, h-index, and other metrics)?
    - □ Not at all familiar (Scholarly metrics are completely new to me)
    - □ Marginally familiar (I have heard of scholarly metrics)
    - □ Somewhat familiar (I know about scholarly metrics but have not personally used them)
    - □ Familiar (I know about scholarly metrics and have explored using them)
    - □ Extremely familiar (I track my own scholarly metrics and regularly use them to demonstrate scholarly impact)

3. How familiar are you with "altmetrics" or nontraditional means of demonstrating scholarly impact (downloads, page views, Mendeley readers, social media followers, and the like)?
    - □ Not at all familiar (This term is completely new to me)
    - □ Marginally familiar (I have heard the term altmetrics)
    - □ Somewhat familiar (I have heard of altmetrics but have not personally used them)
    - □ Familiar (I know about altmetrics and have explored gathering altmetrics on my own scholarship)
    - □ Extremely familiar (I track my own altmetrics and regularly use them to demonstrate scholarly impact)
    - □ Does your department encourage the inclusion of scholarly metrics in your promotion and tenure dossier?
    - □ Yes
    - □ No
    - □ Don't know

4. Does your department require the inclusion of scholarly metrics in your promotion and tenure dossier?*
    - □ Yes
    - □ No
    - □ Don't know

5. How important are scholarly metrics to your department's promotion and tenure process?*
    - □ Not at all important
    - □ Not very important
    - □ Somewhat important
    - □ Fairly important
    - □ Extremely important
    - □ Don't know

6. What other measures of research impact are valued in your department's promotion and tenure process?*
7. What resources do you use to find scholarly metric information?
    - □ None
    - □ Journal Citation Reports
    - □ Web of Science
    - □ Scimago Journal and Country Rank
    - □ Scopus
    - □ Google Scholar
    - □ InCites
    - □ Impact Story
    - □ ResearchGate
    - □ Mendeley
    - □ PlumX
    - □ Publish or Perish
    - □ Academic Analytics
    - □ Experts@Minnesota
    - □ SciVal
    - □ Digital Commons Dashboard
    - □ Other:
8. Where on campus would you turn for help with scholarly metrics?
9. How accurately do scholarly metrics reflect the importance of a researcher's scholarly work?
    - □ Not accurately at all
    - □ Not very accurately
    - □ Somewhat accurately
    - □ Fairly accurately
    - □ Extremely accurately
10. Why do you feel that way?
11. How much weight do you feel your department should place on scholarly metrics in their promotion and tenure processes?*
    - □ No weight
    - □ Very little weight
    - □ Some weight
    - □ A great deal of weight
12. Why do you feel that way?
13. Besides putting together your promotion and tenure dossiers,* when do you look at scholarly metrics?
14. What information regarding scholarly metrics or impact-tracking would be most helpful to you?
15. Please describe any concerns you may have about university administrators tracking the scholarly metric data of their faculty.

*__Note:__ *There were two versions of this survey: one issued to Tenure/Tenure-Track Faculty and Researchers and the other issued to Non-Tenure/Non-Tenure-Track Faculty and Researchers. The survey issued to Non-Tenure-Track Faculty and Researchers replaced references to promotion and tenure dossiers with references to annual performance reviews.*

APPENDIX B. Codebook

Code	Description
Concerns	
Accuracy of data	Concerns about whether specific data points (such as incorrect citation counts) are accurate or complete.
Administrative use: Lack of transparency	Expression of a lack of understanding of how administrators are using research impact metrics. Includes concerns regarding the lack of clarity around expectations and norms.
Administrative use: Leading to negative outcomes	Expression of concern that improper use or application of research impact metrics on the part of administrators may lead to challenges regarding resource allocation, prestige, career advancement, and so forth.
Administrative use: Misunderstanding nuance and disciplinary difference	Expression of concern that administrators will apply research impact metrics in a generalized fashion without recognizing disciplinary differences or other nuances. Include concerns that administrators should not use a single impact measure. Include "metrics are an oversimplification" here.
Administrative use: Undue emphasis	Mentions of overemphasis or overreliance by administrators.
Quantification	Includes statements to the effect that metrics are "okay" but only if qualitative information is used with quantitative. Also include statements about emphasizing quantity over quality.
Interdisciplinarity or disciplinary differences	Expression of concerns regarding disciplinary differences and their representation in research metrics, separate from administrative uses.
Issues of authorship or credit	Including author order and responsibilities.
None	Explicit statement of not having any concerns.
Potential for manipulation	Expression of concerns of the potential for metrics to be manipulated.
Tail wagging the dog	The potential influence of metrics in leading researchers to choose "trendy" research topics. Include anything about trendiness or "hot topics" as this category. Captures a sense that metrics are influencing what people research and how/where they publish. Also include the concept of "chasing numbers."
Timeliness	Impact of the citation lifecycle on publication impact (such as time needed for citations to accrue).
Desires	
Information	An expressed desire to have more information or know more about the use, background, or other aspects of metrics and associated data. Includes answers that simply state a desire to retrieve a particular metric (such as "Journal Impact Factor").
Interpersonal support	An expressed desire for increased individual support, such as the creation or gathering of metrics on behalf of the researcher.
None	Explicit statement of not wanting any information.
Resources (or tools)	An expressed desire for tools and resources through which the researcher can access metrics-related information.
Unspecified	An expressed desire for something, with no additional specification of what.

Lack of Knowledge	
Negative	A lack of knowledge of research impact metrics coupled with a lack of desire to learn more and a sense of negativity regarding use of metrics.
Neutral	A lack of knowledge of research impact metrics with no discernable opinion or attitude. Includes "unsure" or "I don't know enough to know."
Positive	A lack of knowledge of research impact metrics, but coupled with a desire for greater understanding or general positive sentiment.
Motivation	
Assessment of others	Includes stated use of metrics by researchers (not administrators) to assess others (example: exploring metrics of faculty job candidates).
Curiosity	Includes statements expressing curiosity of one's own metrics.
Employment and compensation	Includes promotion and tenure, performance evaluations, job seeking, and other incentives. Assessment of oneself (not others).
External assessment	Assessment of oneself by others.
External prompt	Upon receipt of a notification from a service like ResearchGate. Separate from external assessment.
Information and collaboration seeking	Seeking information about others or their works. Includes statements about using metrics when conducting literature reviews.
Journal evaluation	Deciding where to publish or appraisal of journals. Do not include statements that relate to publishing but are unclear (those should be uncoded).
Self-assessment or benchmarking	Use of metrics information for the purposes of evaluating one's own productivity, impact, or career trajectory
Perceived Utility	
Negative feelings	The validity and utility of research impact metrics are questioned or disparaged.
Positive feelings	Research impact metrics are described as useful, or a positive use of them is described.
Time Periods	
Ambiguous (frequent)	Includes references to checking regularly, routinely, all the time, often (in other words, adverbs).
Ambiguous (rare)	Includes almost never, on occasion (that is, adverbs).
Annually	References to checking metrics annually.
At least monthly	References to checking metrics at least once per month.
At least twice a year	Specifies some sort of unit of time (in other words, no adverbs).
Never	Statements that the respondent never looks at metrics.
Other	
Tools and resources	Reference to specific platforms through which research impact information can be gathered but not expressing a desire for tools or resources.
Alternative impacts	Includes international impact, impact of books, "real-world impact," a need for sentiment analysis, use of altmetrics, and so on.

Notes

1. Eugene Garfield, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas," *Science* 122, no. 3159 (1955): 108–11, https://doi.org/10.1126/science.122.3159.108.

2. Robin Chin Roemer and Rachel Borchardt, *Meaningful Metrics: A 21st Century Librarian's Guide to Bibliometrics, Altmetrics, and Research Impact* (Chicago, IL: Association of College and Research Libraries, 2015), 28.

3. I. Diane Cooper, "Bibliometrics Basics," *Journal of the Medical Library Association : JMLA* 103, no. 4 (October 2015): 217–18, https://doi.org/10.3163/1536-5050.103.4.013.

4. Jason Priem et al., "Altmetrics: A Manifesto," *Altmetrics* (2010), http://altmetrics.org/manifesto/.

5. Roemer and Borchardt, *Meaningful Metrics*, 106–14.

6. James Wilsdon et al., "The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management" (Bristol, UK: Higher Education Funding Council for England, n.d.), https://doi.org/10.13140/RG.2.1.4929.1363.

7. Kylie M. Smith, Ellie Crookes, and Patrick A. Crookes, "Measuring Research 'Impact' for Academic Promotion: Issues from the Literature," *Journal of Higher Education Policy and Management* 35, no. 4 (August 1, 2013): 410–20, https://doi.org/10.1080/1360080X.2013.812173.

8. Wilsdon et al., "The Metric Tide."

9. David Matthews, "A New Model for Professors in the Netherlands," *Inside Higher Ed* (December 7, 2018), https://www.insidehighered.com/news/2018/12/07/netherlands-considers-creating-faculty-positions-based-teaching-not-research-metrics.

10. Wilsdon et al., "The Metric Tide."

11. Kathryn E.R. Graham et al., "Evaluating Health Research Impact: Development and Implementation of the Alberta Innovates: Health Solutions Impact Framework," *Research Evaluation* 21, no. 5 (December 1, 2012): 354–67, https://doi.org/10.1093/reseval/rvs027, 355.

12. Steven Braun, "Supporting Research Impact Metrics in Academic Libraries: A Case Study," *portal: Libraries and the Academy* 17, no. 1 (2017): 111–27, https://doi.org/10.1353/pla.2017.0007; B. Ian Hutchins et al., "Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level," David L Vaux, ed., *PLOS Biology* 14, no. 9 (September 6, 2016): e1002541, https://doi.org/10.1371/journal.pbio.1002541.

13. Alison Abbott et al., "Metrics: Do Metrics Matter?" *Nature News* 465, no. 7300 (2010): 860–62, 861.

14. Abbott et al., "Metrics: Do Metrics Matter?" 861.

15. Dan DeSanto and Aaron Nichols, "Scholarly Metrics Baseline: A Survey of Faculty Knowledge, Use, and Opinion about Scholarly Metrics," *College & Research Libraries* 78, no. 2 (February 2017): 150–70, https://doi.org/10.5860/crl.78.2.150, 163.

16. DeSanto and Nichols, "Scholarly Metrics Baseline," 158.

17. Abbott et al., "Metrics: Do Metrics Matter?"

18. Abbott et al., "Metrics: Do Metrics Matter?"860.

19. Abbott et al., "Metrics: Do Metrics Matter?"861.

20. Abbott et al., "Metrics: Do Metrics Matter?"

21. Mindy Thuna and Pam King, "Research Impact Metrics: A Faculty Perspective," *Partnership: The Canadian Journal of Library and Information Practice and Research* 12, no. 1 (August 29, 2017): 20, https://doi.org/10.21083/partnership.v12i1.3906.

22. DeSanto and Nichols, "Scholarly Metrics Baseline"; Marc Vinyard and Jaimie Beth Colvin, "How Research Becomes Impact: Librarians Helping Faculty Use Scholarly Metrics to Select Journals," *College & Undergraduate Libraries* 25, no. 2 (April 3, 2018): 187–204, https://doi.org/10.1080/10691316.2018.1464995.

23. Htet Htet Aung, Mojisola Erdt, and Yin-Leng Theng, "Awareness and Usage of Altmetrics: A User Survey," *Proceedings of the Association for Information Science and Technology* 54, no. 1 (2017): 18–26, https://doi.org/10.1002/pra2.2017.14505401003.

24. Thuna and King, "Research Impact Metrics," 19.

25. Karen Elizabeth Gutzman et al., "Research Evaluation Support Services in Biomedical Libraries," *Journal of the Medical Library Association* 106, no. 1 (January 12, 2018): 1–14, https://doi.org/10.5195/jmla.2018.205.

26. Ruth Lewis, Cathy Sarli, and Amy M. Suiter, *Scholarly Output Assessment Activities*, SPEC Kit 346 (Washington, DC: Association of Research Libraries, 2015).

27. Christine Wolff, Alisa B. Rod, and Roger C. Schonfeld, "Ithaka S+R US Faculty Survey 2015" (2016), 52, https://sr.ithaka.org/wp-content/uploads/2016/03/SR_Report_US_Faculty_Survey_2015040416.pdf.

28. Bakker, C. et al. 2019. "How Faculty Demonstrate Impact: A Multi-Institutional Study of Faculty Understandings, Perceptions, and Strategies Regarding Impact Metrics." In ACRL 2019 Proceedings: Association of College and Research Libraries, Cleveland, Ohio, April 10-13, 2019, 556-568. www.ala.org/acrl/sites/ala.org.acrl/

files/content/conferences/confsandpreconfs/2019/HowFacultyDemonstrateImpact.pdf.

29. Kathy Charmaz, *Constructing Grounded Theory* (London, UK; Thousand Oaks, CA: Sage Publications, 2006); Juliet M. Corbin and Anselm L. Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 3rd ed.* (Los Angeles, CA: Sage Publications, 2008); Anselm L. Strauss and Juliet M. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (Newbury Park, CA; London, UK: Sage, 1990).

30. Corbin and Strauss, *Basics of Qualitative Research*, 25.

31. Strauss and Corbin, *Basics of Qualitative Research*, 23.

32. Abbott et al., "Metrics: Do Metrics Matter?"; DeSanto and Nichols, "Scholarly Metrics Baseline"; Vinyard and Colvin, "How Research Becomes Impact."

33. Bakker et al., "How Faculty Demonstrate Impact."

34. DeSanto and Nichols, "Scholarly Metrics Baseline."

35. Kirsti Malterud, Volkert Dirk Siersma, and Ann Dorrit Guassora, "Sample Size in Qualitative Interview Studies: Guided by Information Power," *Qualitative Health Research* 26, no. 13 (2016): 1753–60, https://doi.org/10.1177/1049732315617444, 1759.

36. Charmaz, *Constructing Grounded Theory*; Strauss and Corbin, *Basics of Qualitative Research*; Corbin and Strauss, *Basics of Qualitative Research*. NVivo qualitative data analysis software; QSR International Pty Ltd. Version 12, 2018.

37. Michael Quinn Patton, *Qualitative Research & Evaluation Methods* (Thousand Oaks, CA: Sage Publications, 2002).

38. DeSanto and Nichols, "Scholarly Metrics Baseline."

39. Abbott et al., "Metrics: Do Metrics Matter?"

40. Thuna and King, "Research Impact Metrics."

41. Abbott et al., "Metrics: Do Metrics Matter?"; Kai Simons, "The Misused Impact Factor," *Science* 322, no. 5899 (October 10, 2008): 165, https://doi.org/10.1126/science.1165316; Guang Yu, Dong-Hui Yang, and Wang Liang, "Reliability-Based Citation Impact Factor and the Manipulation of Impact Factor," *Scientometrics* 83, no. 1 (2010): 259–70, https://doi.org/10.1007/s11192-009-0083-1.

42. DeSanto and Nichols, "Scholarly Metrics Baseline"; Thuna and King, "Research Impact Metrics."

43. Bakker et al., "How Faculty Demonstrate Impact."

44. DeSanto and Nichols, "Scholarly Metrics Baseline"; Thuna and King, "Research Impact Metrics."

45. Malterud, Siersma, and Guassora. "Sample Size in Qualitative Interview Studies"; Margarete Sandelowski, "Sample Size in Qualitative Research," *Research in Nursing & Health* 18, no. 2 (1995): 179–83, https://doi.org/10.1002/nur.4770180211.

46. Lawrence Leung, "Validity, Reliability, and Generalizability in Qualitative Research," *Journal of Family Medicine and Primary Care* 4, no. 3 (2015): 324–27, https://doi.org/10.4103/2249-4863.161306; Denise F. Polit and Cheryl Tatano Beck, "Generalization in Quantitative and Qualitative Research: Myths and Strategies," *International Journal of Nursing Studies* 47, no. 11 (November 1, 2010): 1451–58, https://doi.org/10.1016/j.ijnurstu.2010.06.004; Sandelowski, "Sample Size in Qualitative Research."