RESEARCH

Faculty Perceptions of Research Assessment at Virginia Tech

Rachel Miles, Virginia Pannabecker and Jim A. Kuypers
Virginia Tech, US
Corresponding author: Rachel Miles (ramiles@vt.edu)

In the spring of 2019, survey research was conducted at Virginia Polytechnic Institute and State University (Virginia Tech), a large, public, Carnegie-classified R1 institution in southwest Virginia, to determine faculty perceptions of research assessment as well as how and why they use researcher profiles and research impact indicators. The Faculty Senate Research Assessment Committee (FSRAC) reported the quantitative and qualitative results to the Virginia Tech Board of Visitors to demonstrate the need for systemic, political, and cultural change regarding how faculty are evaluated and rewarded at the university for their research and creative projects. The survey research and subsequent report started a gradual process to move the university to a more responsible, holistic, and inclusive research assessment environment. Key results from the survey, completed by close to 500 faculty from across the university, include: a.) the most frequently used researcher profile systems and the primary ways they are used (e.g., profiles are used most frequently for showcasing work, with results indicating that faculty prefer to use a combination of systems for this purpose); b.) the primary reasons faculty use certain research impact indicators (e.g., number of publications is frequently used but much more likely to be used for institutional reasons than personal or professional reasons); c.) faculty feel that research assessment is most fair at the department level and least fair at the university level; and d.) faculty do not feel positively towards their research being assessed for the allocation of university funding.

Keywords: survey research; research assessment; case study; research impact indicators; researcher profiles; university budgets

Introduction
In November 2018, the ad hoc Faculty Senate Research Assessment Committee (FSRAC) was formed under the direction of the Faculty Senate President at Virginia Polytechnic Institute and State University (Virginia Tech) to explore concerns and grievances regarding the evaluation of faculty research, scholarship, and creative works as well as concerns regarding faculty salary. As a result, FSRAC designed and distributed a survey for faculty, analyzed the results, and wrote a formal report that was presented to the Virginia Tech Board of Visitors, the governing authority for Virginia Tech, in June 2019. Select results and implications for practice were subsequently presented at two international conferences (Miles et al., 2019a; Miles et al., 2019b). This research paper summarizes the key results from those presentations, excepting salary concerns. The full report can be found at the Virginia Tech Board of Visitors website under the June 2–3, 2019 Minutes: II. Report: Constituent Reports (Kuypers et al., 2019: 12–107).

Literature Review
Because our survey research is based on a sample of faculty members at a single research institution in the United States, it is important to understand previous research on researcher and faculty perspectives, the scientometric literature on research impact indicators, and experts' recommendations on the use of indicators in research evaluation. Past survey and interview research helps research evaluators and practitioners understand the disconnect between faculty and scientometric perspectives on research assessment, especially as it applies to quantitative indicators.

Researcher Perceptions of Research Assessment
Previous studies that examine researchers' perceptions of research evaluation and the use of indicators are widely diverse, although one qualitative literature review attempted to systematically bring these sets of literature together (Rijcke et al., 2016). In their review, Rijcke et al. identified common elements and themes across the literature and

discussed key findings and conceptual arguments; overall, there are few studies that examine the actual effects or consequences of indicator use for research assessment, likely because the process of evaluation is mostly confidential and the production of knowledge takes place in a researcher's day-to-day life and cannot always be inferred through their publication practices. Thus, changes in publication patterns are often studied, such as via the national evaluation systems in Australia and the UK, both of which have a history of incentivizing research impact evaluation via simple quantitative measures in exchange for funding allocations. Such efforts often backfire, though, with numbers for these specific indicators increasing but other significant indicators suffering as a result. For example, in cases in which the 2007 UK Research Assessment Exercise (RAE) aimed at quantity, the number of publications rose sharply; subsequently, however, those publications were mostly unread and uncited (Moed, 2008). Similarly, in Australia, increased pressure to publish in high-impact-factor (IF) journals led to a drastic increase in such publications but an equally drastic decrease in citation counts (Butler, 2003). After criticism from the academic community, the British Research Excellence Framework (REF) replaced the RAE in 2014.

A notable improvement from the above includes broader grading criteria utilizing expert peer review. However, reviewers are required to assign ratings based on their assessments, reducing qualitative assessment to yet another quantitative measure (Gadd, 2019b). Even more worrying, reviewers are more likely to give higher ratings when researchers make grandiose claims of impact in their case reports (Brauer, 2018). In addition, the number of publications is still a problem in the current system; the REF 2014 required an average of four research outputs per submission from a unit, and the REF 2021 decreased the average to 2.5 outputs (Higher Funding Council of England, 2019). Academics in the humanities tend to produce more books, and the REF provides a relatively short period of time for one book publication, let alone two or three (Marenbon, 2018).

Such examples epitomize Goodhart's law: 'when a measure becomes a target, it ceases to be a good measure' (Strathern, 1997). As institutional pressure to attain target measures increases, the measures often become less reliable and more corrupt. A recent study of over 120 million papers published over the past two centuries found that the most trusted and well-known research impact indicators, such as publication counts, citation counts, and the journal IF (JIF), have become increasingly compromised and less meaningful as a result of the pressures to not only publish but also produce high-impact work (Fire & Guestrin, 2019).

Despite these findings, researchers' confidence in citation indicators remains strong (Blankstein & Wolff-Eisenberg, 2019; Buela-Casal & Zych, 2012; Thuna & King, 2017). However, movements for disrupting the academic incentive system are on the rise, such as the Leiden Manifesto and the Declaration on Research Assessment (DORA) (Hicks et al., 2015; The American Society for Cell Biology, 2012). Unfortunately, in recent decades, university rankings have also become increasingly central yet impractical competitions for universities to attain 'global standards' of achievement. Their influence on global funder decisions causes a trickledown effect on how universities ultimately evaluate individual researchers for tenure and promotion (E. Gadd, 2019). These global competitions can easily distract universities away from their missions and discourage locally relevant research. A promising initiative from the International Network of Research Management Societies (INORMS) intends to begin 'rating the rankers' in an effort to hold them to account for their methodology and indicators of excellence (INORMS Research Evaluation Working Group, 2019). Another significant initiative, CWTS Leiden Ranking, outlines principles to guide others in the responsible use of university rankings (Centre for Science and Technology Studies (CWTS), 2019).

Recent survey and interview studies have found that academics' negative sentiment towards these quantitative indicators has hardened. For instance, several faculty survey and interview studies from Ithaka S+R found that the majority of faculty members in the US feel pressured to alter how they communicate their research as a result of review, promotion, and tenure (RPT) requirements to publish journal articles, even when research is more suitable for other output types, or to publish only in journals that meet department standards of prestige or minimum JIF requirements (Blankstein & Wolff-Eisenberg, 2019; Cooper et al., 2017a; Cooper et al., 2017b; Cooper et al., 2018, 2019; Cooper et al., 2017c; Rutner & Schonfeld, 2012; Schonfeld & Long, 2014; Templeton & Lewis, 2015). Interestingly, the most recent Ithaka S+R US faculty survey found that younger faculty members were more likely to prefer the replacement of the subscription publishing model with an open access (OA) model but less likely to publish OA, most likely in response to RPT requirements (Blankstein & Wolff-Eisenberg, 2019).

As an example of the above, consider that without this pressure to meet metrics expectations, many researchers select journals based on how their specific area of research matches that of the journal's scope and readership, regardless of whether the journal is considered prestigious or 'high impact.' This phenomenon mostly affects researchers conducting interdisciplinary (IDR) research and researchers in smaller sub-disciplines whose research is less read but still arguably significant to the larger body of knowledge (Beets et al., 2015; Cooper et al., 2017a; Cooper et al., 2017b; Cooper et al., 2017c; Rutner & Schonfeld, 2012; Schonfeld & Long, 2014). Overall, IDR research leads to more academic breakthroughs and more long-term citation impact (Leahey et al., 2017), but academics are disincentivized from pursuing it (Bromham et al., 2016; Wang et al., 2017) and encouraged to reach more easily defined, short-term metrics, such as producing more publications more frequently and publishing in high-impact journals.

With the increasing focus on short-term quantitative indicators for research assessment and resource allocation, institutions across the world have shifted priorities and sought expert advice on bibliometric indicators and databases, especially from academic libraries (Thuna & King, 2017). A survey study found that faculty members are more likely to consult with librarians on the use of the JIF, h-index, and citation counts for purposes related to promotion, tenure, and

grants than for other purposes, highlighting the importance of these three indicators for evaluation purposes (Miles et al., 2018). A recent review of RPT documents from a representative sample of 129 public universities in the US and Canada revealed that despite the high incidence of the terms 'public' and 'community,' there were no explicit incentives for 'assessing the contributions of scholarship to the various dimensions of publicness' (17), such as the openness and accessibility of scholarship. Rather, the documents revealed that faculty are incentivized to communicate and publish scholarship through narrow academic channels, shown through the documents' explicit requirements of 'impact' across research, teaching, and service (measured for publications by citation indicators, including JIF) (Alperin et al., 2019).

The Value of Research Impact Indicators

In contrast to the academic layperson, scientometricians and bibliometricians typically study citation-based indicators on larger, macro-level scales to understand the structure of and connections between academic fields over decades and centuries or to forecast future research trends. This section will briefly review selected scientometric literature on bibliometric and altmetric indicators as well as the interpretation and appropriate uses of these indicators, with a focus on indicators featured in this paper's survey and survey results.

Interpretations and Limitations of Citation Impact Indicators

There is exhaustive literature on bibliometric indicators, with coverage ranging from their mathematical calculations and theories to their interpretations and practical applications (Waltman, 2016).

As research impact and evaluation becomes an increasingly trending topic in academia, the number of databases that produce the data and metrics also rises, and thus, understanding the source of bibliometric data takes on increased importance. The three major data providers are Web of Science (WoS) from Clarivate Analytics, Scopus from Elsevier, and Google Scholar (GS), though Microsoft Academic (MA) and Dimensions from Digital Science are examples of emerging data sources (Visser et al., 2020). Each database comes with its own set of benefits and limitations, the most obvious factors being cost and coverage across factors such as disciplines, timespan, geographical areas, and languages (Martín-Martín et al., 2018). The main difference is that GS and MA are the most comprehensive bibliographic data sources, especially for content in the humanities and in non-English languages, while Scopus, WoS, and, to some extent, Dimensions are more selective (Martín-Martín et al., 2018; Visser et al., 2020). By sacrificing some data quality and accuracy, researchers and evaluators can find, cite, and evaluate more diverse content (Harzing, 2016). However, use of the more broadly representative sources may be rejected by those who view selectivity and prestige as the gold standard in scholarly communication.

An additional appeal of some large bibliographic databases is the availability of a suite of citation impact indicators at the article, journal, and author levels for evaluative purposes. Scopus (via SciVal), WoS (via InCites), and Dimensions (as part of its higher subscription levels) also provide analytics and benchmarking tools that are marketed primarily to institutions and funders for use in research evaluation, peer comparison, and identification of collaborators, funding, and research trends for research strategy development. The ease of use of these large-scale analytical tools often results in institutions and others using them for broad institutional purposes, despite the incomplete coverage of their institution's research and the general coverage limitations of such databases and the resulting incomplete datasets that underlie any one of them. While GS and MA provide basic citation data, they do not provide research analytic tools or a wide selection of metrics beyond the h-index and citation counts.

Understanding the serious limitations of these large research databases when any one (or even all of them) is treated as providing a 'comprehensive' dataset and picture of global research is important; they underlie major, well-known indicators such as the JIF list and rankings (based on WoS data) and the Times Higher Education (THE) World University Rankings, for which Scopus is 'the exclusive data source' (Hanafi & Boucherie, 2018).

Among the most well-known citation impact indicators is the JIF. It is a proprietary metric published by Clarivate Analytics, available through *Journal Citation Reports*, and based on the WoS dataset (as far as the journals that are included and the calculation of citation metrics for articles in those journals). Research evaluators often treat the JIF as a ranking tool for finding the 'best' journals in one's field. The JIF has several drawbacks, such as its sensitivity to outliers and lack of field normalization, and the fixation on the JIF in academic culture causes a serious clogging effect in scientific publishing (Wilsdon, 2015). Scholars are willing to wait long periods of time for editorial decisions by prestigious high-impact journals, often delaying the communication of important scientific results.

Regardless of the type of citation impact indicator used, many academics trust that citations are a rough proxy for research impact, probably because academic culture equates citations with credit, but also because many studies have found that citations are correlated with other assessments of impact or influence, such as awards, honors, Nobel laureates, research grants, academic rank, and peer reviews (Bornmann & Daniel, 2008). However, many factors and biases play a role in boosting citations, such as citation inflation, in which seasoned and recognized researchers have an advantage over junior researchers; differences in publication and citation behavior across fields; and a bias in favor of English-language literature (van Raan et al., 2011). Citations still serve as a rough proxy for research impact, especially when analyzed on a macro-level scale and/or longitudinally to analyze scholarly communication and impact (Raan, 2005). Therefore, while citations offer some evidence of impact, they should be used in conjunction with qualitative assessment, especially on the micro-level scale or the level of the individual researcher (Hicks et al., 2015).

Some proponents of citation impact indicators in the evaluation system argue that the expansion of accessible impact data helps democratize research evaluation by increasing the number of stakeholders involved in the process (Derrick & Pavone, 2013). Unfortunately, the increasingly wide availability of citation and publication data, along with quick and easy data visualization and benchmarking tools, also tends to overshadow their known limitations, inadvertently leading non-experts to rely on such tools for funding, hiring, RPT, and strategic planning decisions without the aid of experts who dedicate their academic lives to studying and understanding them (Rijcke et al., 2016). Unfortunately, any metric that measures human performance is subject to manipulation and corruption (Muller, 2018). Opponents of metric culture typically do not condone metrics themselves or even the use of metrics, but rather the *over-reliance* on metrics; they argue that metric use in the absence of expert and qualitative assessment is irresponsible, inadequate, and even dangerous (Bergstrom, 2010; Edwards & Roy, 2016; Moustafa, 2016; Muller, 2018).

Alternative Indicators of Research Impact

Alternative indicators of research impact, once called 'webometrics' (Thelwall, 2008), are now commonly referred to as altmetrics (Priem et al., 2010) and comprise a set of quantitative and qualitative data that provide insight into and context for the volume of online attention to research. Altmetrics primarily track mentions of research outputs in news media, social media, reference managers, patents, and public policy and on post-publication peer review platforms, blogs, and Wikipedia.

The majority of academics are unfamiliar with altmetrics in general and/or do not use them (Aung et al., 2019; Južnič et al., 2014; Miles et al., 2018; Sutton et al., 2018). Additionally, they are not widely used in RPT evaluations, and even when they are, evaluators do not usually consider them towards promotion or tenure decisions (Alperin et al., 2019). Overall, the majority of academics who are familiar with altmetrics view them with a mix of skepticism, concern that they can be easily gamed, and dismissal of their potential to demonstrate impact (Colquhoun & Plested, 2014; Holmberg, 2014).

Less than a decade into the evolving field of altmetrics, it is still difficult to pinpoint their value and meaning for the scientometric community and the larger academic community. Much of the early research into altmetrics sought to identify altmetrics as indicators for future citation impact, but most of the studies that drew promising conclusions found that only certain altmetrics sources provide useful data to predict future citation counts, specifically Mendeley and, to some extent, Twitter (Bornmann, 2015; Maflahi & Thelwall, 2016; Thelwall, 2018; Thelwall et al., 2013; Zahedi et al., 2017).

As for societal impact, altmetrics can accomplish this in some capacity, but their value lies in offering a means for tracking and interpreting the source, context, and audience of the attention; one should not rely on the aggregation of the data to imply an approximation for societal impact (Holmberg et al., 2019).

In their review, Rijcke et al. note that as the use and application of research evaluation methods, including the use of research metrics, are distributed across multiple actors and systems (policy makers, funders, research managers, librarians, researchers themselves, and others), the responsibility to change our current methods of evaluation and value assignation must be distributed and taken up by all of these actors as well. In this paper, the authors report on results of a survey conducted by a temporary Faculty Senate-appointed committee that reveals current perceptions of Virginia Tech faculty about research assessment at different institutional levels, research indicators, and the use of research assessment as part of budget allocation. The survey results also summarize types of research outputs and researcher profile systems used by Virginia Tech faculty. This survey research and subsequent report are part of a gradual process aiming to establish university and unit policies for responsible, holistic, and inclusive research assessment.

Distribution and Demographics of Virginia Tech Faculty Survey

The survey was submitted to the Institutional Review Board of Virginia Tech and was determined to not be research involving human subjects as defined by HHS and FDA regulations, resulting in approval to broadly distribute the survey to Virginia Tech faculty with expectations that responses are kept anonymous and results do not claim to be generalizable knowledge (IRB reference number 19-234). The survey was distributed through announcements via the Canvas course management platform, the Virginia Tech Daily, emails to each academic college, and other channels.

Overall, 501 faculty members responded to the survey, with 471 (10% of all full-time faculty) completing the survey. This survey focused on faculty who produce research outputs at Virginia Tech, either as part of their official responsibilities or as part of non-required duties. As a result, the majority of participants were tenure-track or tenured faculty (**Table 1**). In addition, most selected the faculty ranks of Assistant, Associate, or Full Professor, with Associate Professor being the most selected option (38%) (**Table 2**).

The gender of faculty participants closely mirrored university data (male: university: 58%, participants: 50%; female: university: 42%, participants: 39%; prefer not to answer: 10%). Virginia Tech's faculty race or ethnicity is majority White or Caucasian (77%), and the percentage was close to the same for participants (73%). Other races or ethnicities are shown in **Table 3**.

College or top-level unit reports were also produced by representatives from each college and were included in the full survey report to the Faculty Senate and Board of Visitors. Percentages for survey participants' primary college or top-level affiliation are shown in **Table 4**.

Overall, there was a high percentage of responses from the College of Liberal Arts and Human Sciences (CLAHS) compared to the representation of CLAHS at the university (15% versus 31%).

Table 1: Survey participants' faculty types.

Faculty Type	University Data N = 4560	Survey Data n = 471
Tenure-track or Tenured	31%	64%
Continued Appointment-track or Continued Appointment*	1%	9%
Research	16%	8%
Collegiate	1%	2%
Administrative and Professional Faculty	41%	10%
Other (instructors, professors of practice, clinical faculty, etc.)	9%	7%

* Continued appointment (CA) and CA-track are a faculty type specifically for university library faculty to differentiate them as service-oriented faculty and are comparable to the more traditional faculty type 'tenure-track or tenured.'

Table 2: Faculty rank.

Faculty Rank	University Data N = 4964	Survey Data n = 307
Assistant Professor	8%	24%
Associate Professor	10%	38%
Full Professor	12%	31%
Distinguished Professor	Unavailable	2%
Professor Emeritus/Emerita	Unavailable	1%
Other (instructor, lecturer, postdoc, adjunct, visiting, clinical, collegiate, unspecified)	32%	4%

Table 3: Survey participants' race or ethnicity compared to university data.

Race or Ethnicity	University Data N = 4451	Survey Data n = 401
African American or Black	5%	3%
White or Caucasian	77%	73%
Hispanic or Latino/a/x	3%	3%
Asian or Asian American	8%	4%
American Indian or Alaskan Native; Native Hawaiian or other Pacific Islander; Middle Eastern or North African (MENA); Multiple races; other*	8%	3%
Prefer not to answer	N/A	14%

* For the purposes of protecting participants' identity, some categories are combined.

Table 4: Top-level or college affiliation.

Top-level or College Affiliation	University N = 3118	Survey n = 399
College of Liberal Arts and Human Sciences	15%	31%
College of Agriculture and Life Sciences	21%	20%
College of Engineering	20%	11%
University Libraries	3%	8%
College of Science	14%	6%
College of Architecture and Urban Studies	5%	6%
Pamplin College of Business	6%	5%
College of Natural Resources and Environment	5%	3%
Virginia Tech Carilion School of Medicine and Fralin Biomedical Research Institute; Virginia-Maryland College of Veterinary Medicine	10%	1%
Corps of Cadets; Honors College; Student Affairs	0.5%	0%
Other	N/A	3%
Prefer not to answer	N/A	6%

Results

Research Outputs, Profiles, and Measures of Assessment

Types of Research Outputs

Participants could indicate the type(s) of research outputs they have produced, are producing, or will produce (**Figure 1**).

Examples of works self-described in the 'Other' category include

- Submissions and testimony to/before government entities
- Media interviews
- Blog posts
- Musical recordings
- Reports for sponsors (and other reports considered by participant as 'unpublished')
- Public art projects
- Speculative and built architecture
- Cartographic products

The responses about currently produced research outputs show presentations or lectures, publications, and grants as the most-produced types of outputs. The total number of respondents addressing currently or previously produced outputs was 422. A lower number responded regarding future outputs (n = 219). This may indicate that respondents who answered that they 'intend to produce future outputs' may consider this category to be *additional* types of outputs they may *also* produce in the future; or for some, this may represent a planned shift to new types of outputs. If so, patents, creative works, design-based works, and other types seem to be of interest to participants as future outputs, either in addition to the types they currently produce or as a shift into new formats. In addition, the only research output that has a higher percentage of 'intent to produce' than 'previously or currently produced' is patents, which may reflect Virginia Tech's initiatives in innovation, technology, and industry partnerships.
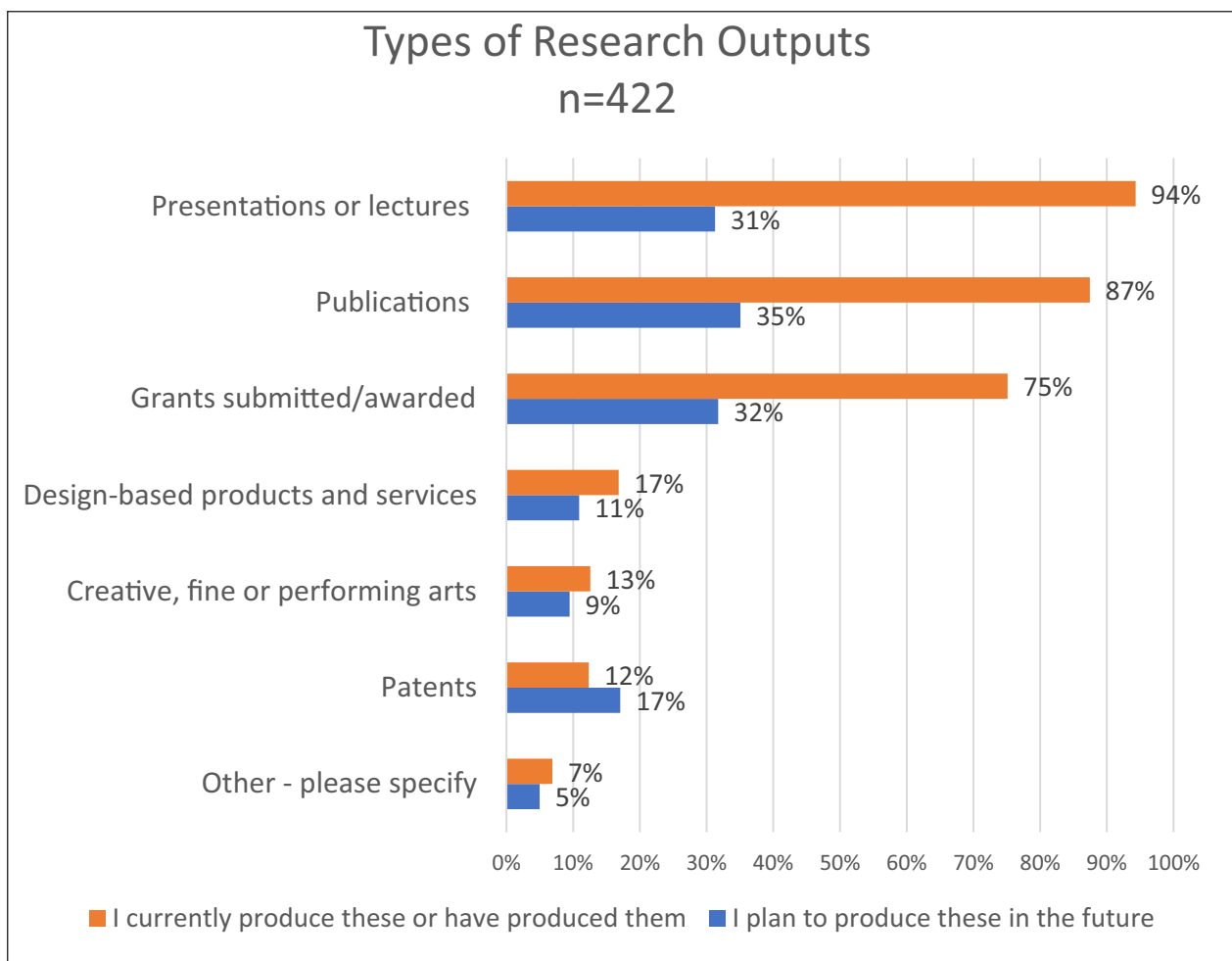

Figure 1: Types of research outputs produced or that faculty plan to produce.

To make some comparisons between disciplines, the data were separated into two groups: respondents in colleges primarily in the science, technology, engineering, and mathematics (STEM) disciplines and colleges primarily in the social sciences and humanities (SSH) disciplines.

- · Classified as in STEM:
    - ◦ College of Science
    - ◦ College of Engineering
    - ◦ College of Natural Resources and Environment
    - ◦ Carilion School of Medicine
    - ◦ College of Veterinary Medicine
- · Classified as in SSH:
    - ◦ College of Liberal Arts and Human Sciences
    - ◦ University Libraries
    - ◦ College of Architecture and Urban Studies
    - ◦ Pamplin College of Business

There were some research output discrepancies between the disciplines for patents as well as creative works (**Figure 2**). Interestingly, more faculty in the SSH have a higher likelihood of their intent to produce patents than faculty in STEM, although faculty in the STEM disciplines are still producing more patents overall. A similar comparison, though reversed, is true for creative, fine, or performing arts. This may reflect Virginia Tech's interest in transdisciplinary research.

In addition, faculty could also select which types of publications, presentations or lectures, and creative or fine arts they produce or intend to produce. This was further analyzed by separating the data into broad discipline categories, STEM and SSH, and unsurprisingly, faculty in the STEM disciplines were more likely to produce peer-reviewed articles, conference papers and abstracts, and software code, while faculty in the SSH were more likely to produce books, book chapters, books edited, translations, prefaces, and reviews of works by others. Surprising yet similar results were found regarding the types of publications faculty plan to produce. Mainly, it appears that faculty in the STEM disciplines are more interested than faculty in the SSH disciplines in producing peer-reviewed articles, book chapters, books, and
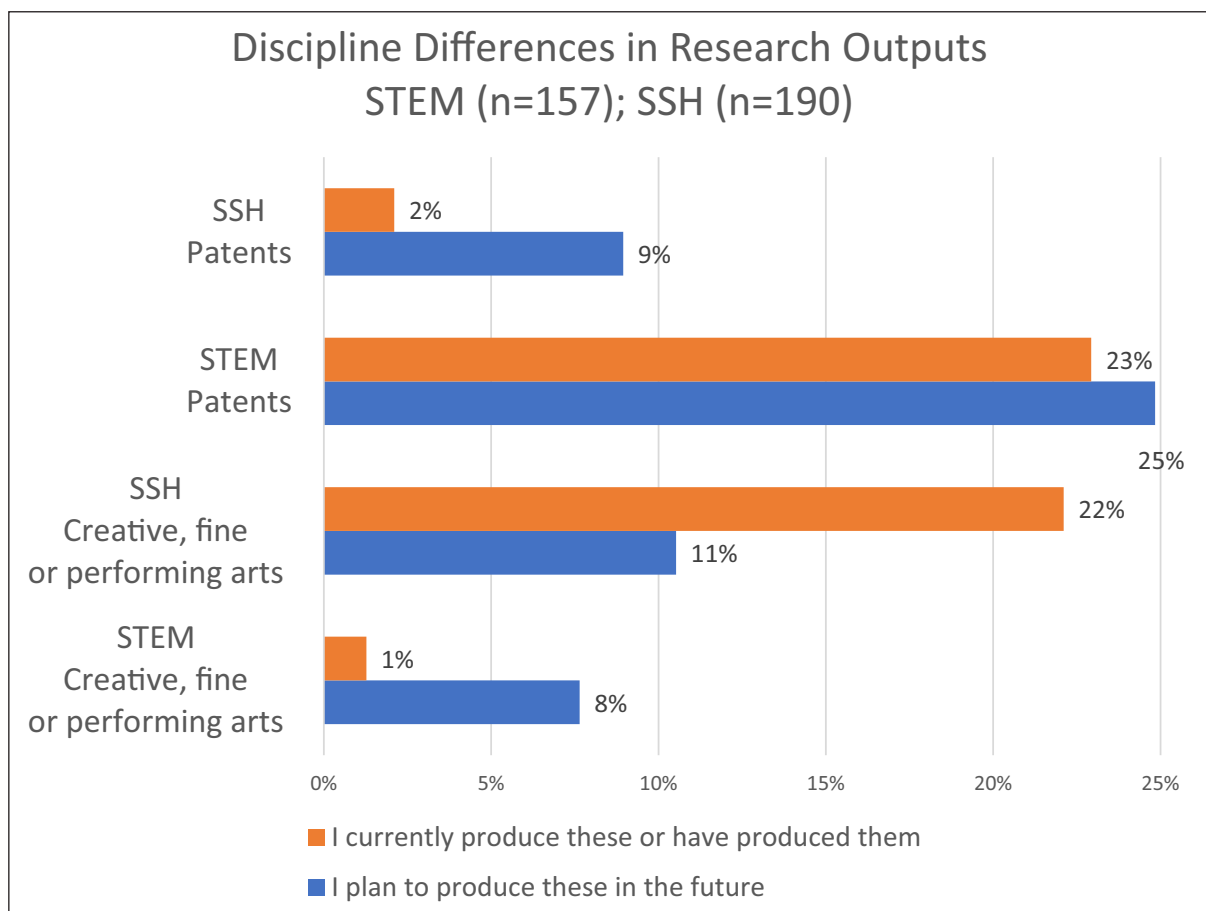

Figure 2: Discipline differences in production of patents and creative, fine, or performing arts.
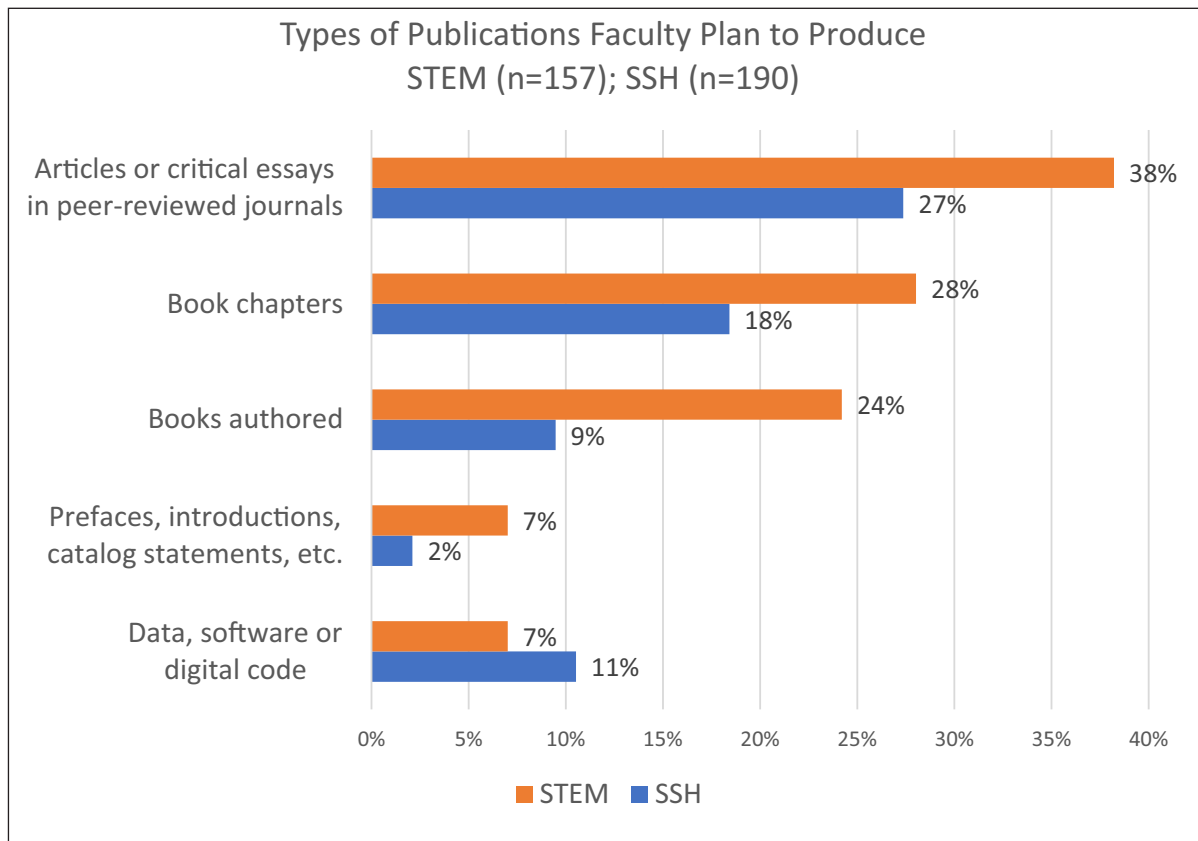
Figure 3: Types of publications faculty plan to produce by discipline.

prefaces or introductions, while faculty in the SSH are slightly more interested in producing data, software, or digital code (**Figure 3**).

There are several potential explanations for the large differences between faculty members' previously produced scholarship and their plans to produce scholarship, as well as the differences between disciplines. As discussed, Virginia Tech's initiatives to encourage transdisciplinary research could be one explanation, but other factors could play a role, such as changes in disciplinary methodologies and approaches, time commitments, or resource allocation.

Profile systems used

Participants were asked about the types of profile systems they used and then asked the primary reason for using them: for personal/professional reasons and/or institutional reasons (e.g., for reporting in annual evaluations, for promotion and tenure) (**Figure 4**). The systems used most (by 100 or more respondents) included GS, Virginia Tech's implementation of Symplectic Elements (a faculty activity data system), LinkedIn, ResearchGate, and ORCID, with several other options used by many respondents as well.

Participants were also asked how they primarily use these profile systems. The faculty activity reporting system Symplectic Elements was most used due to required use, with most participants indicating that they did not find it personally or professionally valuable. This likely relates to described frustration by faculty in survey comments as well as in other venues regarding the time factor required for initially building up one's profile in this system for annual faculty activity reporting, which includes not only publications but also teaching, grants, and service. It may be that providing faculty with additional support to enter data could help alleviate the time burden. As noted below, many survey respondents marked increased visibility and sharing of their work as a primary reason for using profile systems. Because of this, as the Elements faculty activity reporting system is integrated with other institutional systems (repositories, public profiles) that increase visibility and sharing, the work required to add or curate activity data may also be seen as a more worthwhile investment.

Regarding other systems, participants responded that they use LinkedIn (171), ResearchGate (83), and Twitter (82) the most in order to network and connect with colleagues in their field. They use GS Profile (148), ResearchGate (103), and self-published websites (90) the most in order to showcase their work and increase visibility. Across all profile systems, respondents were most likely to use them to showcase work and increase their visibility. This may indicate that researchers see value in sharing highlights across many profile systems to expand their footprint and make their work available (or identifiable) as widely as possible. Although GS, ResearchGate, and ORCID may seem obvious as high-use profile systems for this area, the third most-used type is self-published websites, followed by fourth-place LinkedIn. Finally, they use the following profile systems most when tracking metrics: GS Profile (205), ORCID iD (123), and the Elements/EFARs system (97).
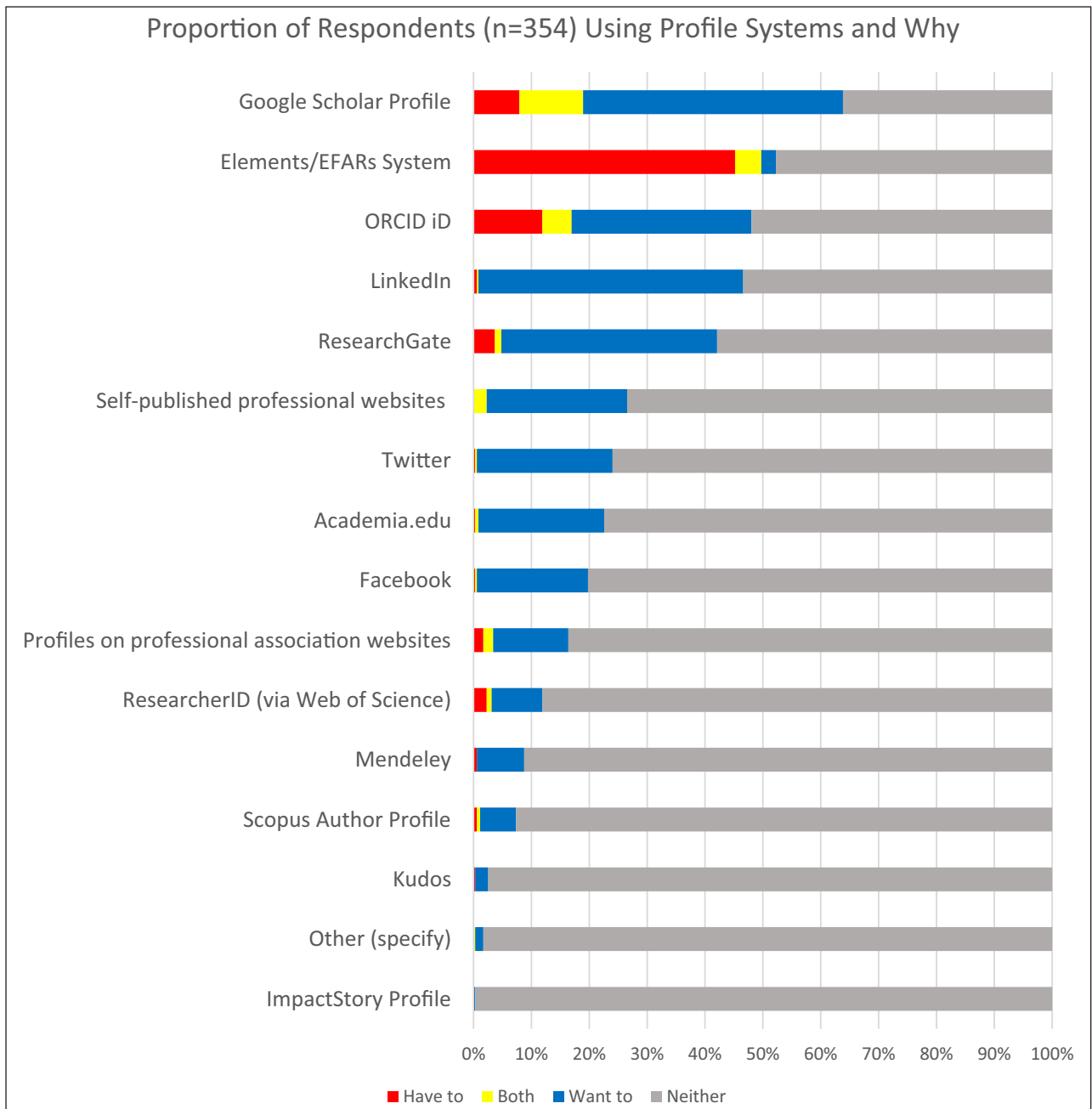
Proportion of Respondents (n=354) Using Profile Systems and Why

Figure 4: Primary reasons participants use profile systems.

Research Impact Indicators

Participants were asked about which research impact indicators they use and why. **Figure 5** shows the percent differences between these reasons, stressing which indicators are used more for institutional reasons versus personal/professional reasons.

The likelihood of using these indicators is related to both personal and institutional reasons; however, certain indicators, such as the number of publications and journal metrics, have high usage, yet both are much more likely to be used for institutional reasons. In contrast, indicators such as altmetrics, expert peer reviews, and attendance numbers are less likely to be used overall, but when they are, they are used more for personal/professional reasons (**Figure 6**). Overall, few participants indicated that they use altmetrics (n = 62, or 17%), and their use was more likely to be for personal/professional reasons, which is consistent with past research findings.

Journal reputation and journal metrics (e.g., JIF) are the most frequently used indicators. The use of the JIF for evaluation purposes in particular was criticized by respondents, with sentiment suggesting that researchers find themselves submitting to a lesser-quality journal simply because it has a higher JIF than others. For this survey, 'journal reputation' was *meant* to indicate the reputation of that journal *in its given field* rather than a journal's 'prestige' or ranking based on journal metrics; unfortunately, the survey did not provide a selection for 'journal prestige' or give a description of 'journal reputation.' Our further investigation into qualitative responses indicates that some respondents felt that journal reputation was based more on prestige and IF. For example, one respondent elaborated that if research is not
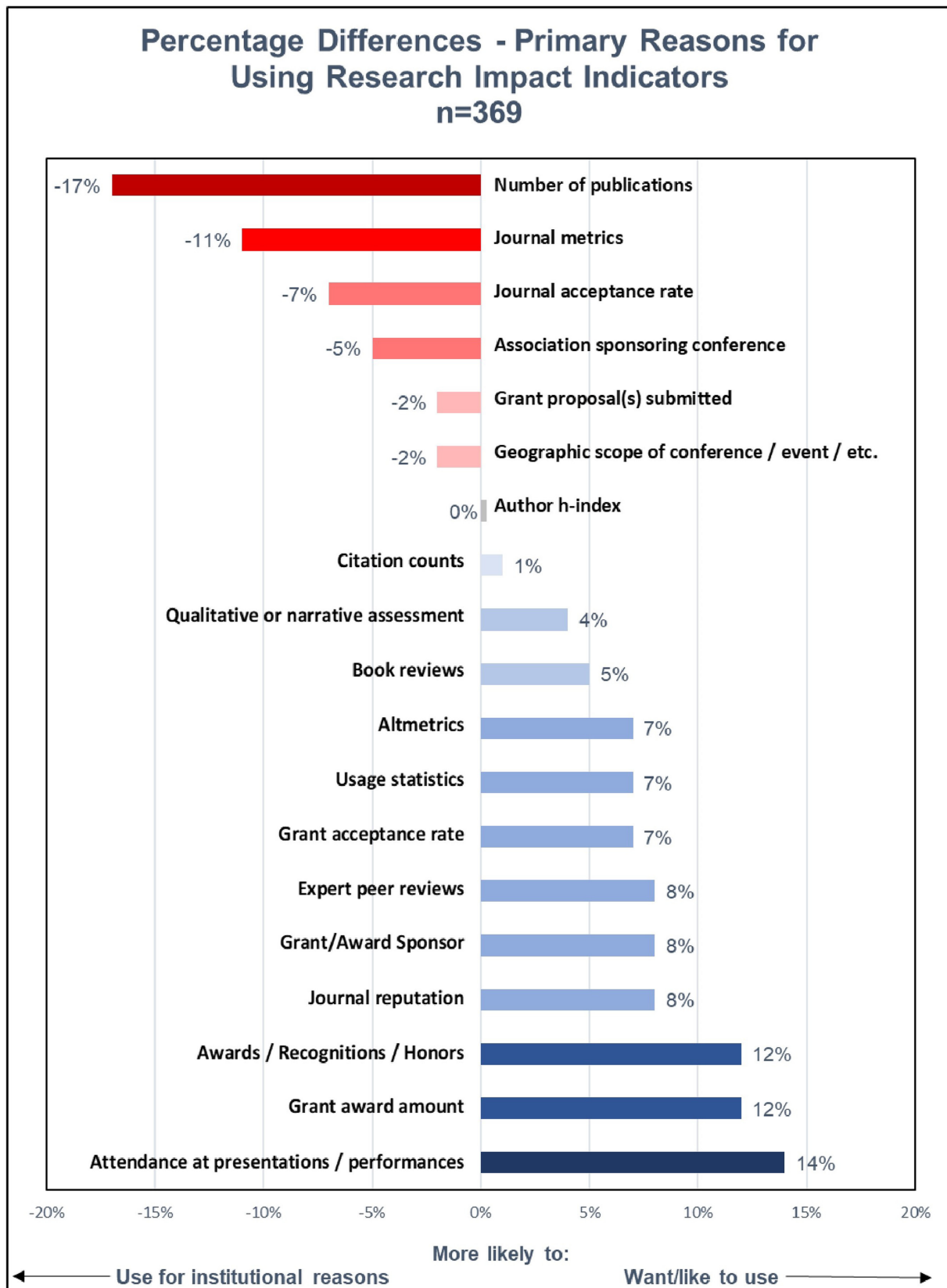
Percentage Differences - Primary Reasons for Using Research Impact Indicators
n=369

Indicator	Value
Number of publications	-17%
Journal metrics	-11%
Journal acceptance rate	-7%
Association sponsoring conference	-5%
Grant proposal(s) submitted	-2%
Geographic scope of conference / event / etc.	-2%
Author h-index	0%
Citation counts	1%
Qualitative or narrative assessment	4%
Book reviews	5%
Altmetrics	7%
Usage statistics	7%
Grant acceptance rate	7%
Expert peer reviews	8%
Grant/Award Sponsor	8%
Journal reputation	8%
Awards / Recognitions / Honors	12%
Grant award amount	12%
Attendance at presentations / performances	14%

More likely to:

← Use for institutional reasons   Want/like to use →

Figure 5: Percent differences for primary reasons for using research impact indicators.

published in one of only four on their college's designated journal list, it is assessed as having 'zero impact' without any further evaluation. Thus, it is quite possible that more participants would have exclusively indicated that they rely on 'journal reputation' for personal/professional reasons if this had been clarified in the survey. Despite this limitation, participants were still more likely to rely on journal reputation for personal or professional reasons than for institutional reasons.
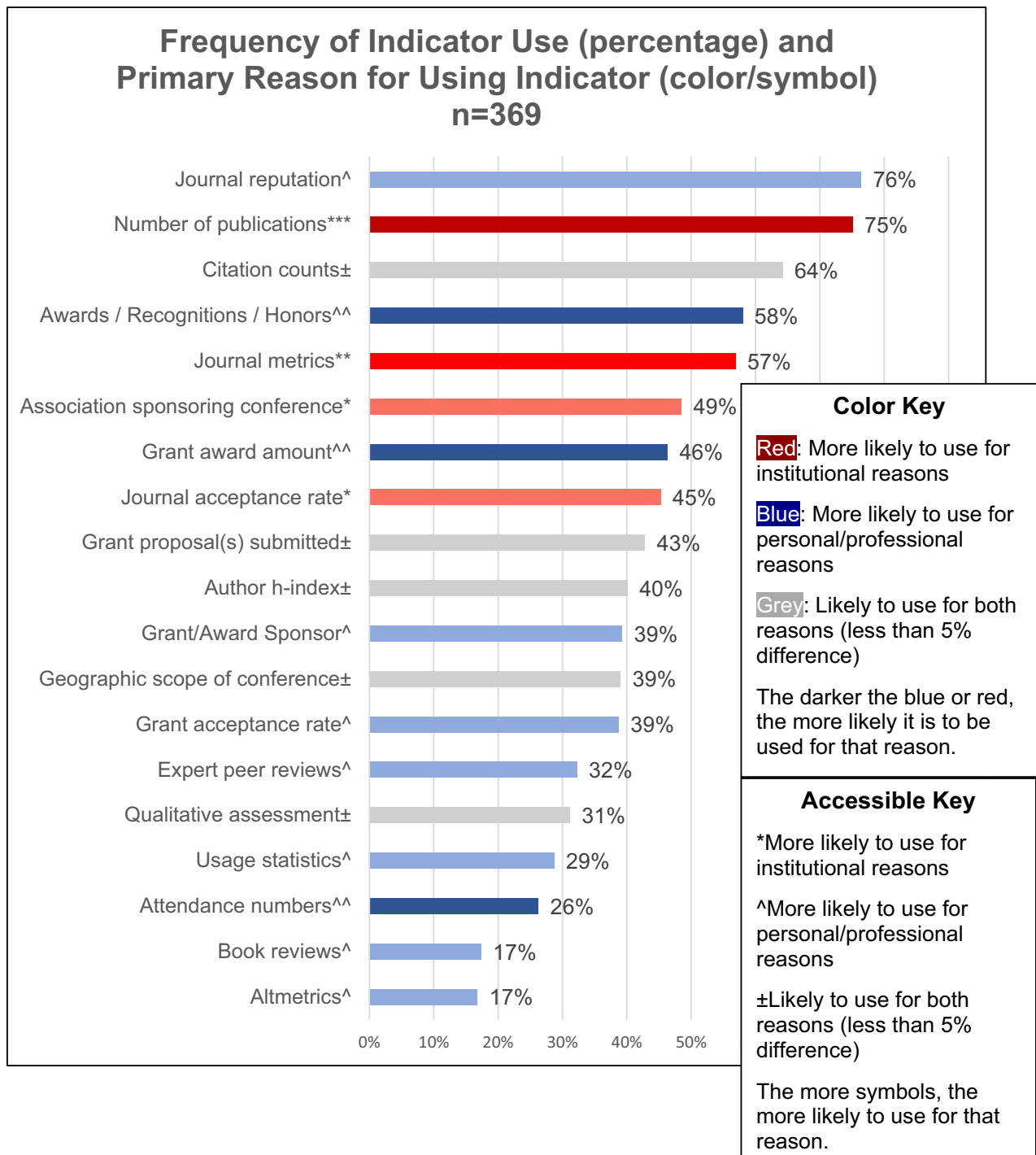
Frequency of Indicator Use (percentage) and Primary Reason for Using Indicator (color/symbol) n=369

Indicator	Percentage
Journal reputation^	76%
Number of publications***	75%
Citation counts±	64%
Awards / Recognitions / Honors^^	58%
Journal metrics**	57%
Association sponsoring conference*	49%
Grant award amount^^	46%
Journal acceptance rate*	45%
Grant proposal(s) submitted±	43%
Author h-index±	40%
Grant/Award Sponsor^	39%
Geographic scope of conference±	39%
Grant acceptance rate^	39%
Expert peer reviews^	32%
Qualitative assessment±	31%
Usage statistics^	29%
Attendance numbers^^	26%
Book reviews^	17%
Altmetrics^	17%

Color Key

Red: More likely to use for institutional reasons

Blue: More likely to use for personal/professional reasons

Grey: Likely to use for both reasons (less than 5% difference)

The darker the blue or red, the more likely it is to be used for that reason.

Accessible Key

*More likely to use for institutional reasons

^More likely to use for personal/professional reasons

±Likely to use for both reasons (less than 5% difference)

The more symbols, the more likely to use for that reason.

Figure 6: Likelihood of using indicator with primary reason for using the indicator.

Remarkably, but not surprisingly, participants were equally likely to use the author h-index for institutional and for personal/professional reasons (n = 148, or 40%), though experts agree that this indicator is severely limited for a number of reasons, such as the lack of accountability for complex publication patterns and its bias against early-career academics and academics who have breaks in their careers (Rowlands, 2018).

Using the same method for separating colleges into two broad disciplines (STEM and SSH), there were also discrepancies found in the data regarding the purposes behind using different research impact indicators. **Figures 7** and **8** show the discrepancies between these two broad discipline classifications in terms of what indicators they are expected to report (**Figure 7**) and what indicators they prefer to use for professional or personal reasons (**Figure 8**). Interestingly, respondents in the SSH category were 18% more likely to prefer to use the h-index for personal/professional reasons but 13% less likely to be expected to use it, whereas the opposite was true for respondents in the STEM disciplines. Respondents in the STEM category were more likely to use grant award amounts for personal/professional reasons as well as for institutional reasons (17% and 19%, respectively). In addition, the journal acceptance rate is 23% more likely
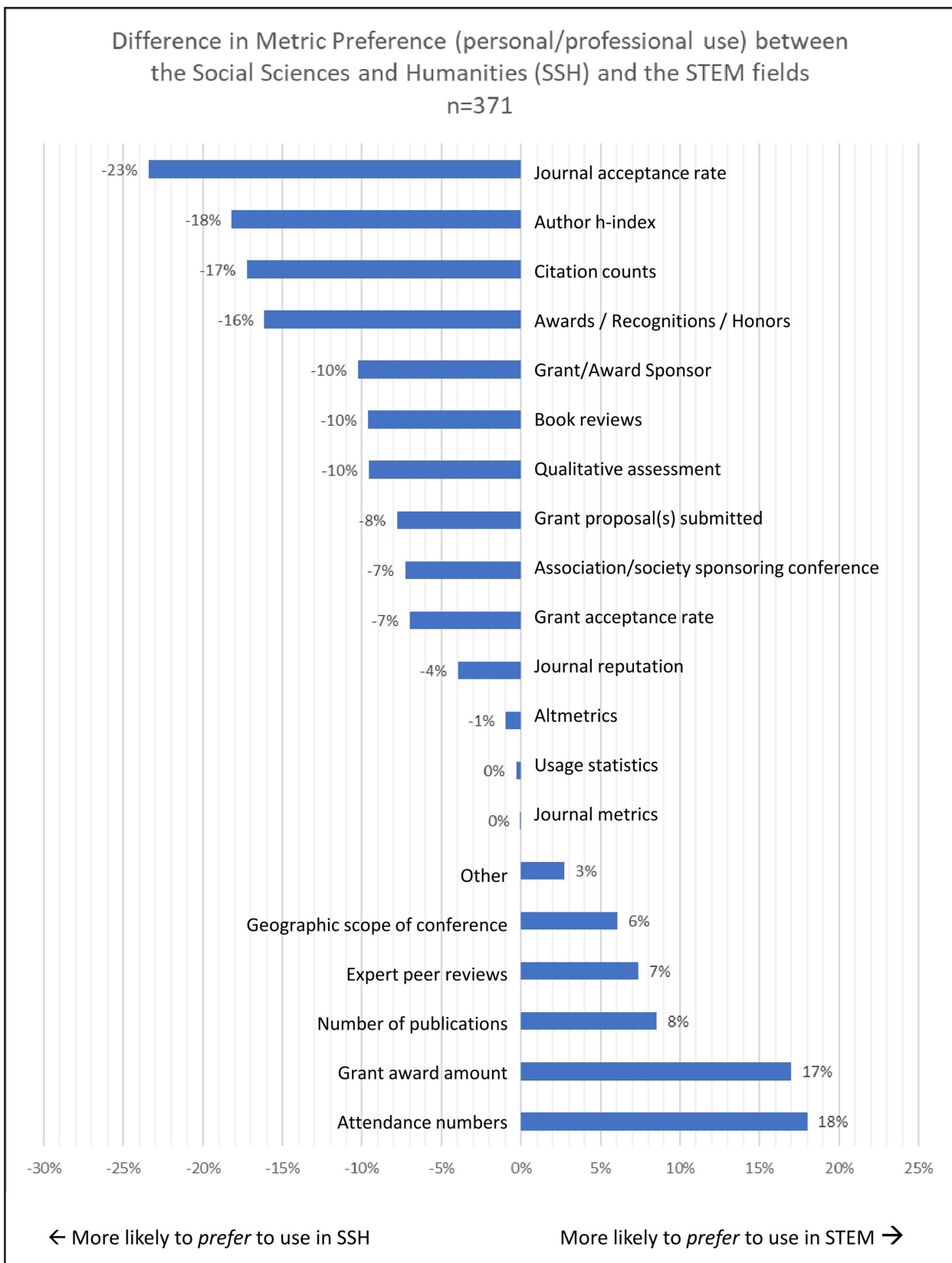
Figure 7: Differences in indicator preference by discipline.

to be used by those in the SSH for both personal/professional and institutional reasons. Journal reputation is more likely to be used for institutional reasons in the SSH (24%) but only slightly more likely to be used for personal/professional reasons (4%). Attendance numbers are much more likely to be used for personal/professional reasons in STEM (18%) but not found to be more institutionally valuable than in the SSH (1%).
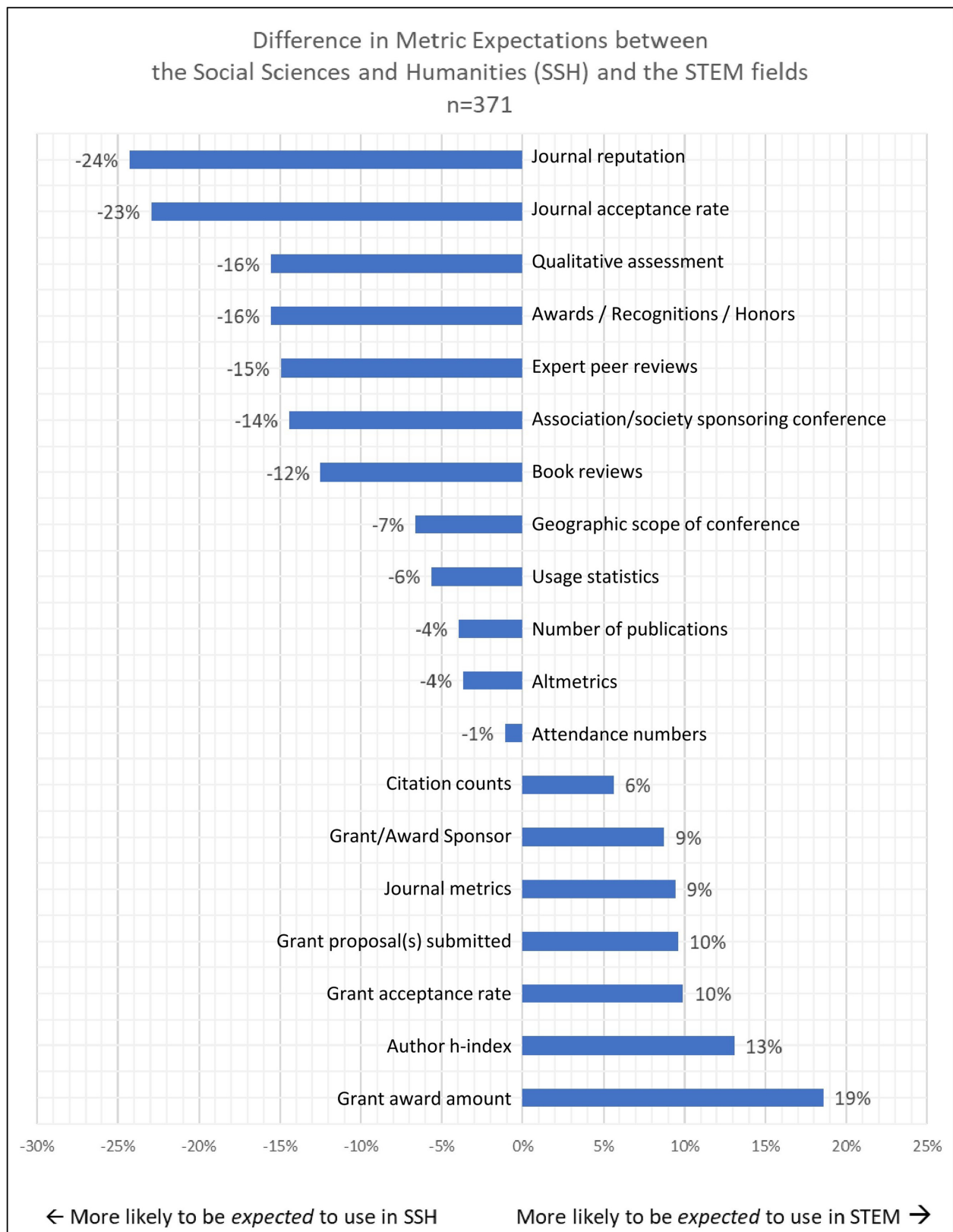
Figure 8: Differences in indicator expectations by discipline.

Perceptions of Research Assessment

Fairness of research assessment

Participants were asked to indicate how fairly they feel their department, college, and university assess their research. Quantitative results indicate that faculty feel more confident in their department's ability to fairly assess their research than in the college or university's ability to do so (**Figure 9**). Many participants argued that research assessment fails
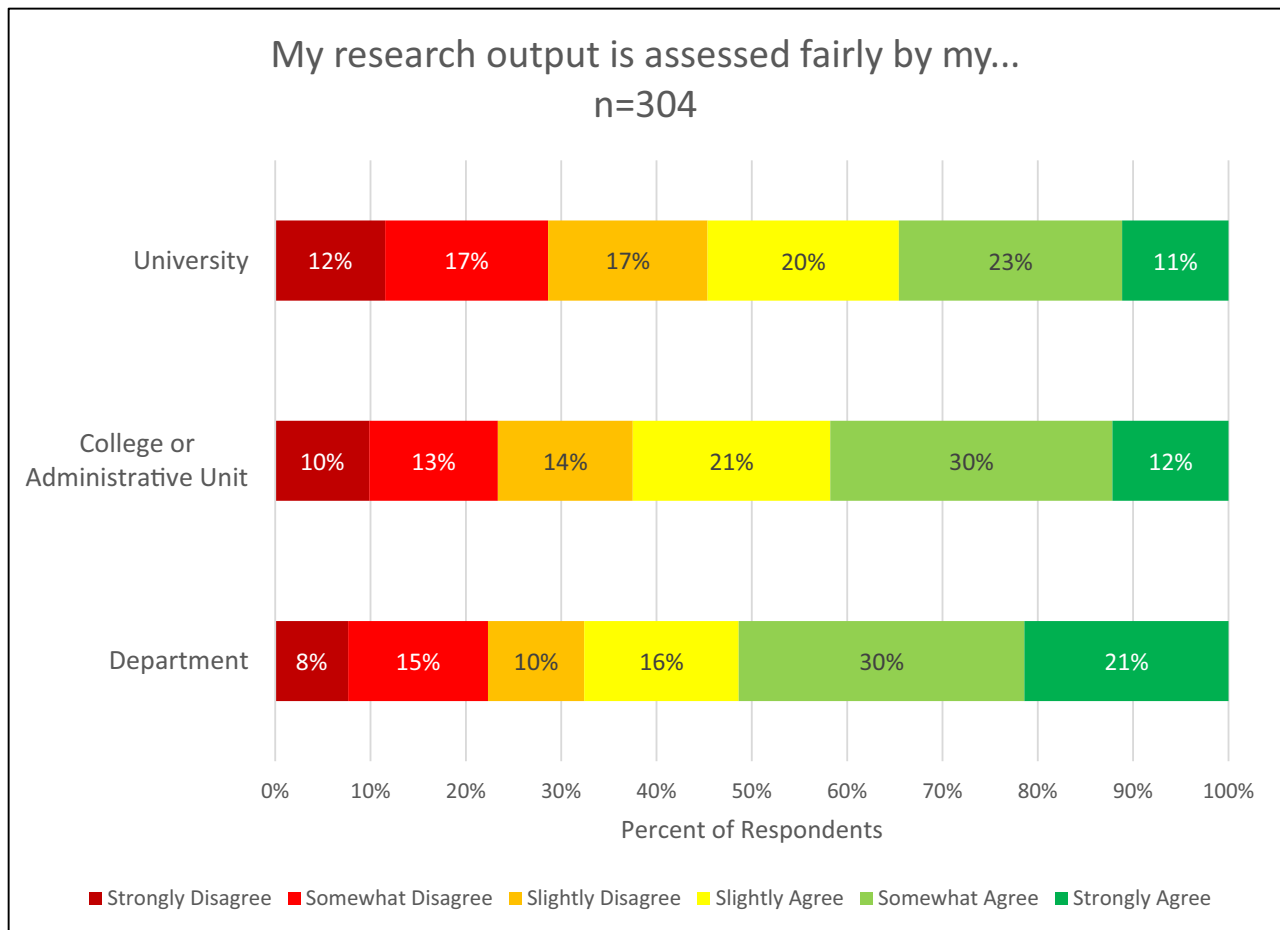
Figure 9: Perception of the fairness of research assessment on the department, college, and university levels.

to recognize disciplinary differences in output. Although primarily expressed as a concern at the college and university levels, some participants related this to be true at the department level as well. Some participants argued that the current format of annual, promotion, and tenure evaluations tends to reflect STEM forms of evaluation that do not transfer well to other areas. Numerous participants argued that different forms of research require different timelines, making it difficult to compare the output of one faculty member to the output of another faculty member. Specifically, some faculty members have larger gaps of time in between their publications because of the nature of their work, and the current model fails to account for this difference. For instance, one individual offered, 'My concern is not about how the university measures my research output when a book is published, but how the college and university measure/recognize research activity in the years between book publications.' Others noted that qualitative assessment allows for considering other aspects of research activities, such as extent and frequency of collaboration.

Integration of research into a new incentive-based budget model

Virginia Tech is currently transitioning into a new budget model, the Partnership for an Incentive Based Budget (PIBB), which is a decentralized budget model that allocates funding by focusing on incentives and metrics in the areas of instruction and enrollment, philanthropy, and sponsored research. The assessment of faculty research has not been integrated into this budget model; thus, the survey sought to gauge faculty's perceptions of this possible addition to the PIBB. Overall, faculty members' responses were overwhelmingly negative towards the notion of including assessment of their own research in the PIBB model (see **Figure 10** for a visualization of the sentiment). Most have little to no familiarity with the PIBB model, have distrust in the model's ability to fairly measure their output, and/or have concerns about the impacts of such a model on the production of their own research. A few respondents pointed out that the use of the PIBB in this manner would not incentivize or encourage more productivity; most respondents feel that alleviating their workloads, especially their commitments to service, and providing more resources will motivate them more effectively to produce and communicate more scholarship. As such, as a critical first step, FSRAC strongly recommended that a responsible research evaluation policy or statement of principles be implemented at Virginia Tech.
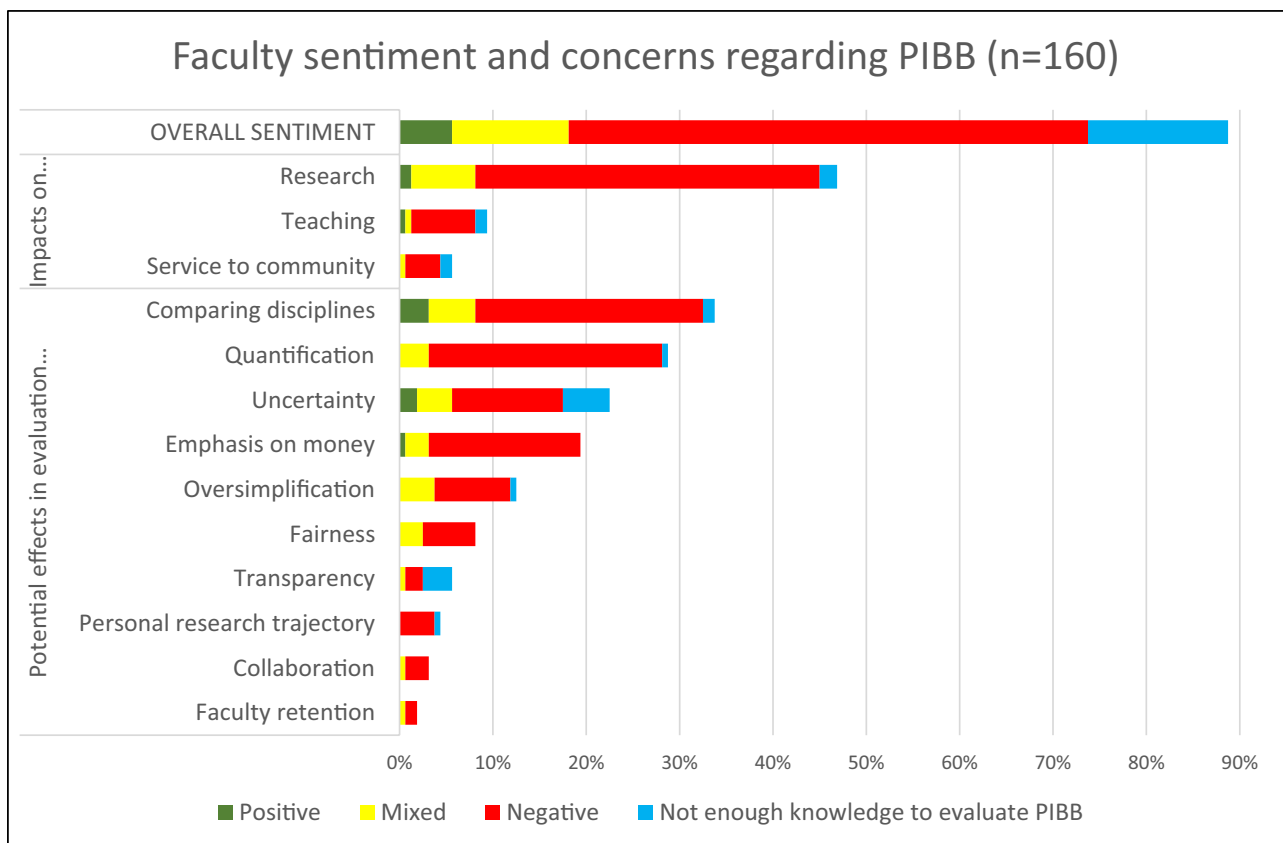
Figure 10: Faculty sentiment regarding the possible integration of research into the PIBB; response sentiment manually coded.

Limitations

Our research is based on a sample of faculty members at an individual research institution, so results cannot be generalized to the general researcher population. However, many of the results were consistent with past research, which is noteworthy. To some extent, the results of this survey can be interpreted on a more qualitative level in terms of learning from the institutional survey research and the advocacy work at Virginia Tech to promote responsible research assessment.

In addition, our survey had some limitations. The survey did not ask faculty members their specific discipline or field, and therefore, the analyses based on field discrepancies were based on college-level classifications that do not always translate well to the classification of disciplines; therefore, these results should be interpreted with care.

Given that the response rate percentages from different colleges were not commensurate with the proportion of faculty in each college in the university, the overall survey results may lean more towards some groups of faculty than if we had received representative response numbers from all areas. For example, while CLAHS faculty make up 14.85% of faculty at Virginia Tech, CLAHS survey respondents made up 30.58% of all respondents.

In general, survey fatigue typically happens near the middle to the end of a survey, and the questions at the end of the survey had fewer responses. In addition, there is typically a negative bias when participants elect to take a survey, and this shows in the text-based responses. However, the consistency of grievances across the text-based responses and across colleges also highlights the problems that need to be addressed internally.

Discussion

Most of our survey results align with past findings on faculty and researcher perceptions of research assessment and their use of research impact indicators. For example, the 2018 Ithaka S+R faculty survey results found that roughly three-quarters of faculty indicated that three factors were 'very important' in the selection of a journal: the journal's area coverage being very close to their immediate research area, the journal's IF or academic reputation, and whether the journal is highly circulated and 'well read' by scholars in the field (i.e., prestige or influence) (Blankstein & Wolff-Eisenberg, 2019). Further insight from the Ithaka faculty interview studies reveals that early-career academics in particular feel pressured to publish in well-established, discipline-specific journals regarded as prestigious for their reputations and JIFs, regardless of whether those journals' aims correspond to their own research or subfields. Many faculty across disciplines expressed grievances about pressures to publish in high-impact and/or prestigious journals despite the fact that their research would be more appropriately communicated and more likely to be read in smaller, field-specific jour-

nals or interdisciplinary journals; such journals typically have lower JIFs and would not 'count' in promotion or tenure dossiers. Our research reflects this sentiment, especially in the text-based responses, in which respondents echoed similar grievances about pressures or outright requirements to publish in specific journals for either their prestige or JIFs. In addition, respondents frequently relied upon the number of publications and journal metrics but did so more for institutional reasons. In contrast, respondents were as likely to use the number of publications (75%) as they were journal reputation (76%) but used journal reputation primarily for personal/professional reasons; this result is consistent with survey and interview findings in other studies described above. Other findings from these same studies also found that researchers maintain strong confidence in traditional citation-based indicators, such as the JIF, h-index, and citation counts, despite their grievances and the limitations of these indicators (Blankstein & Wolff-Eisenberg, 2019; Cooper et al., 2017a; Cooper et al., 2017b; Cooper et al., 2018, 2019; Cooper et al., 2017c; Rutner & Schonfeld, 2012; Schonfeld & Long, 2014; Templeton & Lewis, 2015); further research has found that researchers have little to no familiarity with altmetrics, and when they do, they typically do not find them to be valuable for promotion and tenure.

Advice and recommendations on research assessment and the use of citation impact indicators were exhaustive over the past decade. Despite the research and international initiatives to challenge irresponsible research evaluation culture, practices remain mostly the same, with researchers and administrators trusting JIFs, h-index, and bibliometric data from commercial databases, which do not comprehensively cover disciplines, especially in the humanities; non-English languages; and non-traditional research outputs. Furthermore, with some exception to rates of open research, these indicators cannot provide evidence of public impact or engagement, a purportedly significant factor for public institutions. The gap between the academic sector and the experts in research assessment remains. With growing pressures to demonstrate accountability to the public sector, this gap has been filled by commercial bibliographic database providers that incorporate research (journal, citation, and alternative) indicators (Jappe et al., 2018), regardless of their serious shortcomings. This is especially troubling considering the exclusive reliance on Scopus data for certain university rankings, such as THE World University Rankings (Hanafi & Boucherie, 2018). Rankings have become hypercompetitive environments for climbing the ladders of achievement, and though rankings serve a purpose for comparing universities internationally, they should also be used cautiously (Centre for Science and Technology Studies [CWTS], 2019). At Virginia Tech, the university's strategic plan includes specific milestone goals for achieving specific rankings in THE World University Rankings (Virginia Tech Officer for Strategic Affairs, 2020), and while this could help advance specific areas of research, it could also hurt other areas not well represented by Scopus data if not approached carefully.

Virginia Tech is in a major transition in terms of advancing its research agenda and changing its budget model. Our research team sought to understand how faculty feel about research assessment being incorporated into the new budget model, and overall, sentiment was highly negative. FSRAC wanted to address this potential issue to bring to light the drawbacks of such an approach, and the final report also pointed out the flaws and shortcomings of outside budget models that incorporate research assessment. National budget models outside the United States, such as the REF, allocate funding based on research outputs and their impact via qualitative assessment (e.g., expert peer review) and quantitative assessment (e.g., citation impact indicators). These types of formal assessment exercises are supposed to ensure accountability to the public sector, but institutional performance is difficult to measure accurately and appropriately and is often unfair and overly burdensome to the researchers and faculty being evaluated (Erickson et al., 2020; Kelchen, 2018; Muller, 2018). In addition, these expectations are often manipulated, such as in so-called 'REF poaching,' in which established researchers are headhunted immediately preceding the REF (Jump, 2013).

Perhaps most disturbing of all, the fixation on attaining prestige based on metrics can lead to severe mental health problems in higher education. Based on text-based survey responses, many faculty feel stressed and indicated that their departments, colleges, and the university expect them to attain unrealistic or inappropriate target measures in their research, teaching, and service, such as bringing in large grants every year. In extreme cases, these pressures can even lead to suicide (Flaherty, 2017; Guthrie et al., 2017; Kinman & Wray, 2013; Morrish, 2019; Parr, 2014).

To address some of these issues, senior management or administration can start by being fair and transparent about evaluation procedures and asking faculty for input during critical decisions (Arnett, 2017; Guthrie & Lichten, 2017). Longer-term goals require cultural shifts surrounding academic rewards and incentives; guiding researchers and administrators on the responsible use of research impact indicators, preferably with the support of university policies (Ayris et al., 2018); and shifting to a practice of in-house consultation and reliance on experts for gathering, analyzing, and interpreting research impact data, rather than exclusive dependence on commercial databases. Related to this approach, use of a variety of tools and data sources, such as in-house faculty EFAR systems and bibliometric and altmetric databases, has the potential to gather a more comprehensive and representative data set of faculty activities than can be found in any one data source. Expert input should also be sought by those using such data for funding, hiring, RPT, and strategic planning decisions.

Furthermore, allowing departments and units to decide how to assess themselves will allow for greater academic freedom. Quantitative measures can and should still be used, but without understanding the context of individual researchers and their fields, no metric can solely provide evidence of impact, especially across disciplines and geographic scopes. Our findings suggest that university-wide assessments are overly burdensome, and experts agree that metrics should be used to *complement* or *support* qualitative assessment, with a greater concentration on professional judgment from other colleagues in the field or the department (Bergstrom, 2010; Muller, 2018).

Conclusion

This paper provided a report of results from a survey of faculty perceptions of research assessment completed by close to 500 faculty of varying roles, ranks, and disciplines at Virginia Tech. Beyond survey results, this paper also considers national and international issues that play a role in the evaluation of research, such as world university rankings, the limitations of commercial bibliographic databases, and experts' recommendations on the use of research impact indicators and databases in formal research evaluation. Because the survey was originally intended to act as the basis for formal recommendations to the Faculty Senate and Board of Visitors at Virginia Tech, this paper also describes local issues and advocacy work towards locally relevant cultural and systemic changes; however, many of the issues are also globally relevant and found throughout the scholarly communication and scientometric literature. In the internal report, the FSRAC stressed the importance of implementing policies to address problematic evaluation practices, such as a responsible research evaluation policy that incorporates the varied disciplines represented at Virginia Tech. FSRAC also stressed that research is communicated through diverse research outputs and forms, such as inter- and transdisciplinary research endeavors, and that projects take varied time frames. FSRAC emphasized that if gathering data on faculty research outputs continues to be a central endeavor of the university, additional administrative support for assessment-related data collection and reporting is needed to improve the accuracy of faculty research data and alleviate the burden on faculty of learning a new system and recording more data.

The final report was submitted to the President of the Faculty Senate and subsequently to the Board of Visitors of the Commonwealth of Virginia for their 2 June 2019 meeting. Though FSRAC disbanded after the final report was submitted, a committee was formed in the spring of 2020 under the Faculty Senate to prioritize recommendations from the report and further advocate for university administration to respond to and act on recommendations. In addition, librarians on the committee have plans to distribute and analyze a new survey that will broaden its focus to include researchers beyond faculty, such as research staff, graduate students, and undergraduate students, and also assess researchers regarding their database use, which will help to identify internal areas of demand for training, consultations, instruction, and support services. Finally, faculty survey responses suggest that following responsible research evaluation principles, such as those from the Leiden Manifesto and DORA (Hicks et al., 2015; The American Society for Cell Biology, 2012), would be a commonsense approach, and would permit them to more easily and freely pursue their research, scholarship, and creative projects.

Data Accessibility Statement

The survey instrument and quantitative and qualitative response data analyzed for this publication are available at https://doi.org/10.7294/J1AW-SM37.

Competing Interests

The authors have no competing interests to declare.

References

Alperin, J. P., Nieves, C. M., Schimanski, L. A., Fischman, G. E., Niles, M. T., & **McKiernan, E. C.** (2019). How significant are the public dimensions of faculty work in review, promotion and tenure documents? *ELife, 8.* Scopus. DOI: https://doi.org/10.7554/eLife.42254

Arnett, A. (2017, May 2). What is the role of faculty in the higher ed business model? *Education Dive.* https://www.educationdive.com/news/what-is-the-role-of-faculty-in-the-higher-ed-business-model/441742/

Aung, H. H., Zheng, H., Erdt, M., Aw, A. S., Sin, S.-C. J., & **Theng, Y.-L.** (2019). Investigating familiarity and usage of traditional metrics and altmetrics. *Journal of the Association for Information Science and Technology, 70*(8), 872–887. DOI: https://doi.org/10.1002/asi.24162

Ayris, P., López de San Román, A., Maes, K., & **Labastida, I.** (2018). *Open Science and its role in universities: A roadmap for cultural change* (LERU Publications) [Advice paper]. The League of European Research Universities. https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change

Beets, S. D., Kelton, A. S., & **Lewis, B. R.** (2015). An assessment of accounting journal quality based on departmental lists. *Scientometrics, 102*(1), 315–332. Scopus. DOI: https://doi.org/10.1007/s11192-014-1353-0

Bergstrom, C. (2010). How to improve the use of metrics: Use ranking to help search. *Nature, 465*(7300), 870–872. DOI: https://doi.org/10.1038/465870a

Blankstein, M., & **Wolff-Eisenberg, C.** (2019). Ithaka S+R US Faculty Survey 2018 [Research Report]. *Ithaka S+R.* https://sr.ithaka.org/publications/2018-us-faculty-survey/. DOI: https://doi.org/10.18665/sr.311199

Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, *103*(3), 1123–1144. DOI: https://doi.org/10.1007/s11192-015-1565-y

Bornmann, L., & **Daniel, H.** (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80. DOI: https://doi.org/10.1108/00220410810844150

Brauer, R. (2018). *What research impact?: Tourism and the changing UK research ecosystem.* [Doctoral, University of Surrey]. http://epubs.surrey.ac.uk/846043/

Bromham, L., Dinnage, R., & **Hua, X.** (2016). Interdisciplinary research has consistently lower funding success. *Nature*, *534*(7609), 684–687. DOI: https://doi.org/10.1038/nature18315

Buela-Casal, G., & **Zych, I.** (2012). What do the scientists think about the impact factor? *Scientometrics*, *92*(2), 281–292. DOI: https://doi.org/10.1007/s11192-012-0676-y

Centre for Science and Technology Studies (CWTS). (2019). *CWTS Leiden Ranking: Responsible use.* CWTS Leiden Ranking; Centre for Science and Technology Studies (CWTS). https://www.leidenranking.com/information/responsibleuse

Colquhoun, D., & **Plested, A.** (2014, May 7). Why Altmetrics is bad for science—and healthcare. *The BMJ*. https://blogs.bmj.com/bmj/2014/05/07/david-colquhoun-and-andrew-plested-why-altmetrics-is-bad-for-science-and-healthcare/

Cooper, D., Bankston, S., Bracke, M. S., Callahan, B., Chang, H., Delserone, L. M., Diekmann, F., Dinkelman, A. L., Farmer, D., & **Young, S.** (2017a). Supporting the Changing Research Practices of Agriculture Scholars [Research Report]. *Ithaka S+R*. https://sr.ithaka.org/publications/supporting-the-changing-research-practices-of-agriculture-scholars/. DOI: https://doi.org/10.18665/sr.303663

Cooper, D., Daniel, K., Alburo, J., Banerjee, D., Bialock, T., Cheng, H., Chen, S., Cho, S., Farrell, K. S., Gabbard, R., Henchy, J. A., Hirsch, D., Hopper, M., Hughes, M. J., Islam, R., Kelly, M., Lesage, R., Li, X., Marsh, C., ... Zou, X. (2018). Supporting the Changing Research Practices of Asian Studies Scholars [Research Report]. *Ithaka S+R*. https://sr.ithaka.org/publications/supporting-the-changing-research-practices-of-asian-studies-scholars/. DOI: https://doi.org/10.18665/sr.307642

Cooper, D., Daniel, K., Bakker, C., Blanck, J., Childs, C., Gleason, A., Hanneke, S. H., Lawrence, J., Link, J., Nunez, A., Safranek, S., Smith, K. H., Twose, C., & **Woodson, S.** (2017b). Supporting the Changing Research Practices of Public Health Scholars [Research Report]. *Ithaka S+R*. DOI: https://doi.org/10.18665/sr.305867

Cooper, D., Schonfeld, R. C., Adams, R., Baker, M., Bakkalbasi, N., Bales, J. G., Bedard, R., Benda, C., Bidlack, B., Bogue, S., Burr, T., Harrison, G. C., Cohen, I., Custer, W., Dearborn, V., van Dyk, G., Estell-Holmer, S., Flynn, K., Gundry, J., ... Weiss, N. (2017c). Supporting the Changing Research Practices of Religious Studies Scholars [Research Report]. *Ithaka S+R*. DOI: https://doi.org/10.18665/sr.294119

Cooper, D., Springer, R., Benner, J. G., Bloom, D., Carrillo, E., Carroll, A., Change, B., Chen, X., Daix, E., Dommermuth, E., Figueiredo, R., Haas, J., Hafner, C. A., Henshilwood, A., Krogman, A. L. C., Zuege Kuglitsch, R., Lanteri, S., Lewis, A., Li, L., ... Yu, S. H. (2019). Supporting the Changing Research Practices of Civil and Environmental Engineering Scholars (Ithaka S+R Research Reports) [Research Report]. *Ithaka S+R*. DOI: https://doi.org/10.18665/sr.310885

Edwards, M. A., & **Roy, S.** (2016). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*, *34*(1), 51–61. DOI: https://doi.org/10.1089/ees.2016.0223

Erickson, M., Hanna, P., & **Walker, C.** (2020, February 17). The Senior Management Survey: Auditing the toxic university. *Impact of Social Sciences*. https://blogs.lse.ac.uk/impactofsocialsciences/2020/02/17/the-senior-management-survey-auditing-the-toxic-university/

Flaherty, C. (2017, April 25). Professor's death sparks discussion of academic mental health. *Times Higher Education (THE)*. https://www.timeshighereducation.com/news/professors-death-sparks-discussion-academic-mental-health

Gadd, E. (2019). Influencing the changing world of research evaluation. *Insights*, *32*(1), 6. DOI: https://doi.org/10.1629/uksg.456

Gadd, L. (2019, April 24). Three thoughts about REF peer review. *The Bibliomagician*. https://thebibliomagician.wordpress.com/2019/04/24/three-thoughts-about-ref-peer-review/

Guthrie, S., & **Lichten, C.** (2017, August 27). Why the mental ill health of academic researchers remains a hidden problem. *Times Higher Education (THE)*. https://www.timeshighereducation.com/blog/why-mental-ill-health-of-academic-researchers-remains-hidden-problem

Guthrie, S., Lichten, C. A., van Belle, J., Ball, S., Knack, A., & **Hofman, J.** (2017). *Understanding mental health in the research environment: A Rapid Evidence Assessment* [Product Page]. https://www.rand.org/pubs/research_reports/RR2022.html. DOI: https://doi.org/10.7249/RR2022

Hanafi, S., & **Boucherie, S.** (2018). *Discover the data behind the Times Higher Education World University Rankings.* Elsevier Connect. https://www.elsevier.com/connect/discover-the-data-behind-the-times-higher-education-world-university-rankings

Harzing, A.-W. (2016). Sacrifice a little accuracy for a lot more comprehensive coverage. *Harzing.Com*. https://harzing.com/publications/white-papers/gsbook-prologue

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & **Rafols, I.** (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520*(7548), 429–431. DOI: https://doi.org/10.1038/520429a

Higher Funding Council of England. (2019). Guidance on submissions (2019/01)—REF 2021 (Worldwide). *Higher Education Funding Council for England.* https://www.ref.ac.uk/publications/guidance-on-submissions-201901/

Holmberg, K. (2014). *How Concerned Should We Be About Gaming Altmetrics?* https://www.socialsciencespace. com/2014/07/altmetrics-may-be-able-to-help-in-evaluating-societal-reach-but-research-significance-must-be-peer-reviewed/

Holmberg, K., Bowman, S., Bowman, T., Didegah, F., & **Kortelainen, T.** (2019). What Is Societal Impact and Where Do Altmetrics Fit into the Equation? *Journal of Altmetrics, 2*(1), 6. DOI: https://doi.org/10.29024/joa.21

INORMS Research Evaluation Working Group. (2019). What makes a fair and responsible university ranking? Rating the rankings criteria Version 2. *INORMS.* https://inorms.net/wp-content/uploads/2019/09/what-makes-a-good-ranking-list-of-characteristics-version-2.pdf

Jappe, A., Pithan, D., & **Heinze, T.** (2018). Does bibliometric research confer legitimacy to research assessment practice? A sociological study of reputational control, 1972–2016. *PLOS ONE, 13*(6), e0199031. DOI: https://doi. org/10.1371/journal.pone.0199031

Jump, P. (2013, September 26). Twenty per cent contracts rise in run-up to REF. *Times Higher Education (THE).* https:// www.timeshighereducation.com/news/twenty-per-cent-contracts-rise-in-run-up-to-ref/2007670.article

Južnič, P., Vilar, P., & **Bartol, T.** (2014). What do researchers think about altmetrics and are they familiar with their abilities? *Libraries in the Digital Age (LIDA) Proceedings, 13*(0), Article 0. http://ozk.unizd.hr/proceedings/index. php/lida/article/view/128

Kelchen, R. (2018). *Higher Education Accountability.* Johns Hopkins University Press Books. https://jhupbooks.press. jhu.edu/title/higher-education-accountability

Kinman, G., & **Wray, S.** (2013). *Higher Stress: A Survey of Stress and Well-being among Staff in Higher Education.* University and College Union. https://www.yumpu.com/en/document/ read/41588860/a-survey-of-stress-and-well-being-among-staff-in-higher-ucu

Kuypers, J., Westwood, J., Wong, E., Houptman, J., Meany, K., Knapp, B., Viehland, D. D., Hicok, B., Leonard, R. H., Woods, C., Mielczarek, N., Dayer, A., Thomas, Q., Merola, J., Redican, K. J., McMillan, G., Miles, R., Pannabecker, V., Porter, N. D., & **Macdonald, A. B.** (2019). *Virginia Tech Faculty Senate Research Assessment Committee's June 2019 Report on Faculty Research and Salary* (Board of Visitors Meeting Minutes, pp. 12–107). Virginia Tech. https://bov.vt.edu/minutes.php

Leahey, E., Beckman, C. M., & **Stanko, T. L.** (2017). Prominent but Less Productive: The Impact of Interdisciplinarity on Scientists' Research*. *Administrative Science Quarterly, 62*(1), 105–139. DOI: https://doi. org/10.1177/0001839216665364

Maflahi, N., & **Thelwall, M.** (2016). When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals. *Journal of the Association for Information Science and Technology, 67*(1), 191–199. DOI: https://doi. org/10.1002/asi.23369

Marenbon, J. (2018, December 5). Research funding is short-changing humanities subjects|John Marenbon. *The Guardian.* https://www.theguardian.com/education/2018/dec/05/our-research-funding-system-is-shortchanging-the-humanities

Martín-Martín, A., Orduña-Malea, E., Thelwall, M., & **López-Cózar, E. D.** (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics, 12*(4), 1160–1177. DOI: https://doi.org/10.1016/j.joi.2018.09.002

Miles, R. A., MacDonald, A., Porter, N. D., Pannabecker, V., & **Kuypers, J. A.** (2019a). What Do Faculty Think About Researcher Profiles, Metrics, and Fair Research Assessment? A Case Study from a Research University in the Southeastern United States [Panel Presentation]. *10th Annual VIVO Conference,* Podgorica, Montenegro. https:// vtechworks.lib.vt.edu/handle/10919/93360

Miles, R., Konkiel, S., & **Sutton, S.** (2018). Scholarly Communication Librarians' Relationship with Research Impact Indicators: An Analysis of a National Survey of Academic Librarians in the United States. *Journal of Librarianship and Scholarly Communication, 6*(1). https://jlsc-pub.org/articles/abstract/10.7710/2162-3309.2212/. DOI: https://doi. org/10.7710/2162-3309.2212

Miles, R., Pannabecker, V., MacDonald, A., Kuypers, J., & **Porter, N. D.** (2019b). Faculty Perceptions on Research Impact Metrics, Researcher Profile Systems, Fairness of Research Evaluation, and Time Allocations [Panel Presentation]. *6:AM Altmetrics Conference,* Stirling, UK. https://vtechworks.lib.vt.edu/handle/10919/96929

Morrish, L. (2019). Pressure Vessels: The epidemic of poor mental health among higher education staff (No. 20; Occassional Paper, p. 76). *Higher Education Policy Institute.* https://www.hepi.ac.uk/2019/05/23/pressure-vessels-the-epidemic-of-poor-mental-health-among-higher-education-staff/

Moustafa, K. (2016). Aberration of the Citation. *Accountability in Research, 23*(4), 230–244. DOI: https://doi.org/10.1 080/08989621.2015.1127763

Muller, J. Z. (2018). *The Tyranny of Metrics.* Princeton University Press. https://www.amazon.com/Tyranny-Metrics-Jerry-Z-Muller/dp/0691174954. DOI: https://doi.org/10.23943/9781400889433

Parr, C. (2014, December 3). Imperial College professor Stefan Grimm 'was given grant income target.' *Times Higher Education (THE)*. https://www.timeshighereducation.com/news/imperial-college-professor-stefan-grimm-was-given-grant-income-target/2017369.article

Priem, J., Taraborelli, D., Groth, P., & **Neylon, C.** (2010). *altmetrics: A manifesto*. http://altmetrics.org/manifesto/

Raan, A. F. J. van. (2005). For Your Citations Only? Hot Topics in Bibliometric Analysis. *Measurement: Interdisciplinary Research and Perspectives*, *3*(1), 50–62. DOI: https://doi.org/10.1207/s15366359mea0301_7

Rijcke, S. de, Wouters, P. F., Rushforth, A. D., Franssen, T. P., & **Hammarfelt, B.** (2016). Evaluation practices and effects of indicator use—A literature review. *Research Evaluation*, *25*(2), 161–169. DOI: https://doi.org/10.1093/reseval/rvv038

Rowlands, I. (2018, March 23). Is it time to bury the h-index? [Blog]. *The Bibliomagician*. https://thebibliomagician.wordpress.com/2018/03/23/is-it-time-to-bury-the-h-index/

Rutner, J., & **Schonfeld, R.** (2012). Supporting the Changing Research Practices of Historians [Research Report]. *Ithaka S+R*. DOI: https://doi.org/10.18665/sr.22532

Schonfeld, R. C., & **Long, M. P.** (2014). Supporting the Changing Research Practices of Art Historians [Research Report]. *Ithaka S+R*. DOI: https://doi.org/10.18665/sr.22833

Sutton, S., Miles, R., & **Konkiel, S. Q.** (2018). Awareness of Altmetrics among LIS Scholars and Faculty. *Journal of Education for Library & Information Science*, *59*(1/2), 33–47. Library, Information Science & Technology Abstracts. DOI: https://doi.org/10.3138/jelis.59.1-2.05

Templeton, G. F., & **Lewis, B. R.** (2015). Fairness in the institutional valuation of business journals. *MIS Quarterly: Management Information Systems*, *39*(3), 523–539. Scopus. DOI: https://doi.org/10.25300/MISQ/2015/39.3.01

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*. DOI: https://doi.org/10.1177/0165551507087238

Thelwall, M. (2018). Early Mendeley readers correlate with later citation counts. *Scientometrics*, *115*(3), 1231–1240. DOI: https://doi.org/10.1007/s11192-018-2715-9

Thelwall, M., Haustein, S., Larivière, V., & **Sugimoto, C. R.** (2013). Do Altmetrics Work? Twitter and Ten Other svices. *PLOS ONE*, *8*(5), e64841. DOI: https://doi.org/10.1371/journal.pone.0064841

Thuna, M., & **King, P.** (2017). Research Impact Metrics: A Faculty Perspective. *Partnership: The Canadian Journal of Library & Information Practice & Research*, *12*(1), 1–25. Library, Information Science & Technology Abstracts. DOI: https://doi.org/10.21083/partnership.v12i1.3906

van Raan, A. F. J., van Leeuwen, T. N., & **Visser, M. S.** (2011). Severe language effect in university rankings: Particularly Germany and France are wronged in citation-based rankings. *Scientometrics*, *88*(2), 495–498. DOI: https://doi.org/10.1007/s11192-011-0382-1

Virginia Tech Office for Strategic Affairs. (2020). *The Virginia Tech Difference: Advancing Beyond Boundaries*. Virginia Tech. https://strategicaffairs.vt.edu/content/strategicaffairs_vt_edu/en/StrategicPlanning/the-vt-difference-advancing-beyond-boundaries.html

Visser, M., van Eck, N. J., & **Waltman, L.** (2020). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *ArXiv:2005.10732 [Cs]*. http://arxiv.org/abs/2005.10732

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*(2), 365–391. DOI: https://doi.org/10.1016/j.joi.2016.02.007

Wang, J., Veugelers, R., & **Stephan, P.** (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, *46*(8), 1416–1436. DOI: https://doi.org/10.1016/j.respol.2017.06.006

Wilsdon, J. (2015). *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. DOI: https://doi.org/10.4135/9781473978782

Zahedi, Z., Costas, R., & **Wouters, P.** (2017). Mendeley readership as a filtering tool to identify highly cited publications. *Journal of the Association for Information Science & Technology*, *68*(10), 2511–2521. DOI: https://doi.org/10.1002/asi.23883